Applied Numerical Mathematics 30 (1999) 225–239

An iterative method for algebraic solution to interval equations

Svetoslav Markov [1]

Bulgarian Academy of Sciences, Institute of Mathematics and Computer Sciences, Acad. G. Bonchev str. bl. 8, 1113 Sofia, Bulgaria

Abstract

The algebraic solution to systems of linear equations involving an interval square matrix and an interval right-hand side vector in terms of interval arithmetic is discussed. The basic concepts of interval arithmetic are given in a form suitable for our study. An iterative Jacobi type method is formulated and its convergence has been proved, under certain conditions on the interval matrix. In the special case when only the right-hand side is interval-valued we reduce the problem to two ordinary linear systems. An iterative numerical algorithm is proposed and numerically demonstrated. © 1999 Elsevier Science B.V. and IMACS. All rights reserved.

Keywords: Iterative methods; Interval linear system; Interval arithmetic

1. Introduction

We consider linear algebraic systems involving intervals in the $(n \times n)$-matrix $\boldsymbol{A}$ and in the right-hand side n-vector $\boldsymbol{b}$. We shall be concerned with the (*interval*) *algebraic solution* which is an interval n-vector $\boldsymbol{x}$ satisfying the system whenever the arithmetic operations are performed in interval arithmetic. We shall write the problem in the form

$$\boldsymbol{A} * \boldsymbol{x} = \boldsymbol{b}, \tag{1}$$

to emphasize that the symbol "$*$" means *interval multiplication* of the interval matrix $\boldsymbol{A}$ by the solution vector $\boldsymbol{x}$; the latter being generally an interval vector.

Using a symbolic notation based on binary variables taking values from the set $\{+, -\}$, we formulate new distributivity relations and simple rules for the transformation of algebraic expressions and equations. This leads to a powerful complete interval algebraic structure, referred to as *directed interval arithmetic*. We use the term "directed" to distinguish it from similar arithmetic theories like the "modular" arithmetic developed by Gardenes and his collaborators [2]. The directed interval arithmetic unifies both the extended interval arithmetic as developed by Kaucher [4] and the extended arithmetic for normal intervals using *inner* (*nonstandard*) operations [6]. In particular, it has been proved that the system of

[1] E-mail: smarkov@iph.bio.bas.bg.

PII: S0168-9274(98)00112-3

normal intervals with inner operations is: (i) a “projection” of the directed interval arithmetic system on the system of normal intervals, and (ii) isomorphically embedded in the directed interval algebraic system [7,8].

Some characteristic features of directed interval arithmetic:
(i) provides a simple general framework which enables us to pass from directed (proper and improper) intervals to proper intervals (with inner operations), and vice versa;
(ii) involves a set of new relations and computational rules necessary for the straightforward symbolic transformation of algebraic expressions and equations;
(iii) makes use of binary variables, respectively specific “plus–minus” notations.

The algebraic solution to (1) seems to be related to the solutions of other practically significant linear algebraic problems involving interval coefficients, such as the united, the controlled and the tolerable solutions (cf. [15–18]). For instance, it has been shown, that the algebraic solution presents an inner estimate of the united solution [5]. This shows the importance of a self-contained study of the algebraic solution to (1).

In this work we propose an interval Jacobi type iteration procedure for finding the algebraic solution of the interval system (1) and prove its convergence using new powerful tools from directed interval arithmetic. We point out to a relation between our method and a method proposed in [19] and studied in [5]. In the special case when only the right-hand side of (1) is interval-valued we give conditions for the existence of solution and formulate the solution in explicit form. A numerical algorithm based on the proposed iterative method is formulated and numerical examples are given. Parts of this work have been presented in [9,11]. Some of the basic concepts of directed interval arithmetic, necessary to formulate and prove our method cannot be found in the literature. For this reason we give a brief introduction to this arithmetic in a form suitable for our purposes.

2. Directed interval arithmetic

Given $A^-, A^+ \in \mathbb{R}$ the set $A = \{x \mid A^- \leqslant x \leqslant A^+\} = [A^-, A^+]$ is called a (normal, proper) interval (on $\mathbb{R}$). The set of all normal intervals on $\mathbb{R}$ is denoted by $I(\mathbb{R})$. The operations addition and multiplication for $A, B \in I(\mathbb{R})$ are defined by

$$A + B = \{\xi + \eta \mid \xi \in A,\ \eta \in B\}, \qquad A * B = \{\xi\eta \mid \xi \in A,\ \eta \in B\}, \tag{2}$$

and inclusion $A \subset B$ is understood in the usual set-theoretic sense. Denote

$$Z = \{A \in I(\mathbb{R}) \mid A^- \leqslant 0 \leqslant A^+\}, \qquad Z^* = \{A \in I(\mathbb{R}) \mid A^- < 0 < A^+\}.$$

The systems $(I(\mathbb{R}), +)$, $(I(\mathbb{R}) \setminus Z,\ \times)$ are Abelian cancellative semigroups. Therefore they can be isomorphically extended up to (minimal) group systems by means of the familiar extension method used for the introduction of negative and rational numbers [4]. The corresponding group systems will be further denoted by $(D, +)$, $(D^*, \times)$, where D is the set of pairs of the form $a = (A', A'')$, $A', A'' \in I(\mathbb{R})$, factorized by the equivalence relation $(A, B) = (C, D)$ if and only if $A + D = B + C$ and $D^* = D \setminus \mathcal{T}$, where $\mathcal{T}$ is the set of elements “involving” 0 to be defined below. The elements of D are called *directed intervals*. By means of the extension method one obtains the isomorphic extensions of the operations “$+$”, “$*$” and the inclusion relation “$\subset$” in D. Therefore we may assume that the algebraic system $(D, +, *, \subset)$ is well defined. All other operations in D (like subtraction, division etc.) are derived as subsidiary operations from the basic ones (2) by means of familiar algebraic constructions (isomorphic embeddings, inverses, compositions, projections, etc.) [4,8]. As a consequence of this abstract algebraic approach one can derive: (i) algebraic properties and relations in D, and (ii) corresponding formulae involving a numerical component-wise presentation of intervals. In the present work we shall make use mainly of algebraic properties of the system $(D, +, *, \subset)$ and not of the component-wise presentation of intervals (an exception of this rule are some results in Section 5 using the center-radius presentation). For convenience we include some formulae for the two familiar interval presentations—the endpoint and center-radius form in Appendix A.

Every directed interval can be presented as a pair of normal intervals, which has either the form $(A, 0)$ or the form $(0, B)$ [7,8]. Directed intervals of the form $(A, 0)$, $A \in I(\mathbb{R})$, are called proper (and if $A \in \mathbb{R}$—degenerate), the ones of the form $(0, B)$ with $B \in I(\mathbb{R}) \setminus \mathbb{R}$ are improper. The set of all proper intervals is equivalent to $I(\mathbb{R})$ and is denoted again by $I(\mathbb{R})$, the set of degenerate intervals is equivalent to $\mathbb{R}$ and is denoted by $\mathbb{R}$, and the set of improper intervals is denoted by $\overline{I(\mathbb{R})}$. We say that two directed intervals are of same type if they both belong either to $I(\mathbb{R})$ or to $\overline{I(\mathbb{R})}$. We have $D = I(\mathbb{R}) \cup \overline{I(\mathbb{R})}$, that is D consists: (i) of the set of all normal intervals, such as $[0, 1]$, $[-1, 1]$, $[1, 1] = 1$, etc., and (ii) of improper intervals, which can be presented either as $(0, A) = [A; -]$ (see, e.g., [7]), or as intervals with reverse "endpoints"—note that the endpoints of $A \in I(\mathbb{R})$ are *endpoints* of $(0, A)$ (see Appendix A). Denote

$$\overline{Z} = \{(0, A) \mid A \in I(\mathbb{R}),\ A^- \leqslant 0 \leqslant A^+\}, \qquad \overline{Z}^* = \{(0, A) \mid A \in I(\mathbb{R}), A^- < 0 < A^+\};$$

denote also $\mathcal{T} = Z \cup \overline{Z}$, $\mathcal{T}^* = Z^* \cup \overline{Z}^*$, $D^* = D \setminus \mathcal{T}$. Note that D^* is the set of pairs $(A, 0)$ or $(0, A)$, $A \in I(\mathbb{R})$, such that $0 \notin A$. The set D^* is often mentioned below, since we are able to divide by its elements. Occasionally we shall also make use of the larger set $D \setminus \mathcal{T}^*$ of elements, possibly having one of their "endpoints" equal to zero (that is, pairs $(A, 0)$ or $(0, A)$, such that A may have zero as an endpoint).

Algebraic properties and computational rules. The systems $(D, +)$ and $(D^*, *)$ are groups according to the definition of "$+$", respectively "$*$", by the above mentioned extension method; hence these systems possess inverse additive and multiplicative operators. Denote by $-_D a$ the opposite (additive inverse) of $a \in D$ and by $1 /_D a$ the inverse of $a \in D^*$ with respect to "$*$". Hence, for $a, b \in D$ the unique solution to the equation $a + x = b$ is $x = b + (-_D a) = b -_D a$. Similarly, for $a \in D^*, b \in D$ the unique solution to the equation $a * y = b$ is $y = b * (1 /_D a) = b /_D a$. Of course, $a -_D a = 0$ and $a /_D a = 1$.

Interval multiplication by scalar is special case of interval multiplication when one of the multipliers in $a * b$ is degenerate (that is, real number). Multiplication by -1 is called *negation* and is denoted $-a = (-1) * a$; we write $a - b = a + (-b)$. Interval multiplication by scalar obeys the rules: $\alpha * (\beta * u) = (\alpha\beta) * u$ and $\alpha * (u + v) = \alpha * u + \alpha * v$ for $\alpha, \beta \in \mathbb{R}, u, v \in D$. The rule $(\alpha + \beta) * u = \alpha * u + \beta * u$ does not generally hold (see below).

Dualization (*dual interval*), denoted a_-, is defined as composition of negation and opposite by $a_- = -(-_D(a))$. We have $-_D(-a) = -(-_D a) = a_-$. *Reciprocal element* is defined in D^* as $1/a = 1 /_D a_-$, we have $1 /_D (1/a) = 1/(1 /_D a) = a_-$, for $a \in D^*$. The inverse operators with respect to multiplication can be written as $1 /_D a = 1/a_-$. The inverse elements $-_D a$, $1 /_D a$ generate the *inverse operations* $a -_D b = a + (-_D b) = a + (-b_-) = a - b_-$, $a /_D b = a * (1 /_D b) = a * (1/b_-) = a/b_-$. For $a, b \in D$, $c \in D^*$ the (unique) solutions of the equations $a + x = b$, $c * y = b$, $a + c * z = b$ are respectively $x = b + (-a_-) = b - a_-$, $y = b * (1/c_-) = b/c_-$, $z = (b - a_-)/c_-$.

Denoting $a_+ = a$, we have $a_\lambda \in \{a, a_-\}$ for $\lambda \in \Lambda = \{+, -\}$. A "product" of two binary variables $\mu\nu$, $\mu, \nu \in \Lambda$, is defined by $++ = -- = +$, and $+- = -+ = -$. Using this product we may write $(a_\lambda)_\delta = a_{\lambda\delta}$, in particular, $(a_\lambda)_\lambda = a$. Note that the equations $a = b$ and $a_\lambda = b_\lambda$ are equivalent; and such are $a_\lambda = b$ and $a = b_\lambda$. The equations $a = b$ and $a - b_- = 0$ are equivalent, also $a = b$ and $a/b_- = 1$ for $b \in D^*$. We also have $(a + b)_\lambda = a_\lambda + b_\lambda$, $(a * b)_\lambda = a_\lambda * b_\lambda$, $a_\lambda - a_{-\lambda} = 0$, etc.

Define the "sign" $\sigma : D \setminus \mathcal{T}^* \to \Lambda$ of a directed interval $a = (A, 0)$, respectively, $a = (0, A)$, such that $0 \notin A$, by

$$\sigma(a) = \begin{cases} +, & \text{if } A^- \geqslant 0,\ A^+ \geqslant 0, \\ -, & \text{if } A^- < 0,\ A^+ \leqslant 0. \end{cases}$$

The *conditionally distributive* (q-*distributive*) *law* states [1,7,13]: For $a, b, c, a + b \in D \setminus \mathcal{T}^*$ we have

$$(a + b) * c_{\sigma(a+b)} = a * c_{\sigma(a)} + b * c_{\sigma(b)}, \tag{3}$$

which can be also written $(a + b) * c = a * c_{\sigma(a)\sigma(a+b)} + b * c_{\sigma(b)\sigma(a+b)}$.

In particular, for all $\alpha, \beta \in \mathbb{R}$, $c \in D$, we have

$$(\alpha + \beta) * c_{\sigma(\alpha+\beta)} = \alpha * c_{\sigma(\alpha)} + \beta * c_{\sigma(\beta)}. \tag{4}$$

Relation (4) shows that the algebraic system $(D, +, \mathbb{R}, *)$ is not a linear space; this space, called q-linear, has been studied in [8,10]. From (4) it can be noticed that the q-linear space involves a linear multiplication $\alpha \cdot c = \alpha * c_{\sigma(\alpha)}$, satisfying $(\alpha + \beta) \cdot c = \alpha \cdot c + \beta \cdot c$.

The following distributivity relation may be also useful. If $d \in \mathcal{T}^*$, then d can be splitted into two intervals a, b having the same type as d and satisfying the relations $a, b \in \mathcal{T}$, $d = a + b$ (in fact, both a and b have each one endpoint equal to zero). Then, for $c \in D^*$ we have

$$d * c = (a + b) * c = a * c + b * c. \tag{5}$$

The interval $a \in D$ is called *symmetric* if $a = -a$. Denote the set of all symmetric intervals by D_S. If $a \in D_\mathrm{S}$, then $a_- \in D_\mathrm{S}$ as well. It is easy to see that if a is symmetric, then $\gamma * a = |\gamma| * a$ for $\gamma \in \mathbb{R}$. The subsets $\mathbb{R}$ and D_S form a basis of D in the following sense. If we fix two intervals $e \in \mathbb{R}$, $s \in D_\mathrm{S}$, $e \neq 0$, $s \neq 0$, then every $a \in D$ can be uniquely presented by means of a pair of real numbers (a', a''), $a', a'' \in \mathbb{R}$, such that $a = a' * e + a'' * s_{\sigma(a'')} = a' * e + |a''| * s_{\sigma(a'')}$. We shall further fix $e = 1$, $s = j = [-1, 1]$ (note that j satisfies $j * j = j$). Thus we have $a = a' + |a''| * j_{\sigma(a'')}$, which will be symbolically denoted by $a = (a', a'')$ and called *center-radius* form. It is easy to see that in center-radius form addition is $(a', a'') + (b', b'') = (a' + b', a'' + b'')$, and interval multiplication by scalar is $\alpha * (b', b'') = (\alpha b', |\alpha| b'')$. Dualization is $(a', a'')_- = (a', -a'')$, more generally we can write $(a', a'')_\lambda = (a', \lambda a'')$. More relations for the center-radius form are given in Appendix A.

Inclusion, norm, metric. The systems $(D, +, \subset)$ and $(D^*, *, \subset)$ are isotone groups [4]. Norm in D is introduced in the usual way: $\|x\|$ is a norm in D, if: (i) $\|x\| \geqslant 0$ and $\|x\| = 0$ if and only if $x = 0$; (ii) $\|x + y\| \leqslant \|x\| + \|y\|$; and (iii) $\|k * x\| \leqslant |k|\|x\|$ for $k \in \mathbb{R}$. For $x \in D$, $\|x\| = \inf_t\{t * j_- \subset x \subset t * j\}$ is a norm [4]. Some useful properties of the norm in D are: $\|x * y\| \leqslant \|x\|\|y\|$, $\|x_-\| = \|x\|$. Other norms are given in Appendix A. The norm in D induces a metric $r(x, y)$ by means of $r(x, y) = \|x -_D y\| = \|x - y_-\|$ for $x, y \in D$.

3. Interval matrix algebra

Denote by D^n the set of all n-dimensional vectors $\boldsymbol{x} = (x_1, \ldots, x_n)$ with components $x_i \in D$, $i = 1, \ldots, n$, and by $D^{l\times k}$ the set of all $(l \times k)$-dimensional matrices $\boldsymbol{A} = \{a_{ij}\}_{i=1,\ldots,l}^{n=1,\ldots,k} = (a_{il})$ with $a_{il} \in D$. Operations between vectors and matrices of directed intervals, further called *interval vectors*, respectively *interval matrices*, are defined similarly to matrix operations involving numbers. Addition and subtraction are defined component-wise for interval vectors (matrices) of identical size. An interval vector with all components degenerate is called degenerate.

Dot (inner) product of two interval vectors $\boldsymbol{x} = (x_1, \ldots, x_n) \in D^n$, $\boldsymbol{y} = (y_1, \ldots, y_n) \in D^n$, is defined by $\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \sum_{i=1}^n x_i * y_i$; in particular, if $\boldsymbol{x} = \xi \in \mathbb{R}^n$ we have $\langle \xi, \boldsymbol{y}\rangle = \sum_{i=1}^n \xi_i * y_i$. If $\boldsymbol{A} = (a_{ij}) \in D^{m\times l}$ and $\boldsymbol{B} = (b_{ij}) \in D^{l\times n}$, then the product of the interval matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ is the matrix $\boldsymbol{C} = \boldsymbol{A} * \boldsymbol{B} = (c_{ij}) \in D^{m\times n}$ with $c_{ij} = \sum_{k=1}^l a_{ik} * b_{kj}$. This defines the expression $\boldsymbol{A} * \boldsymbol{x}$ in (1) as a product of two interval matrices: namely, $\boldsymbol{A}$, $\boldsymbol{x}$ are considered as interval matrices of order $n \times n$, $n \times 1$, respectively, and the result $\boldsymbol{A} * \boldsymbol{x}$ is an interval $(n \times 1)$-matrix. Let $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ be a real-valued (point) matrix and let $\boldsymbol{u}, \boldsymbol{v} \in D^n$. Using that the matrix $\boldsymbol{A}$ is real-valued we obtain the rule $\boldsymbol{A} * (\boldsymbol{u} + \boldsymbol{v}) = \boldsymbol{A} * \boldsymbol{u} + \boldsymbol{A} * \boldsymbol{v}$.

The norm $\|x\|$ in D is extended for vectors and matrices in the usual way. For instance, for $\boldsymbol{A} = (a_{ik}) \in D^{n\times n}$, we may define a matrix norm by

$$\|\boldsymbol{A}\| = \max_i \left\{ \sum_{k=1}^n \|a_{ik}\| \right\}.$$

Other norms are given in Appendix A. For the product of two interval matrices we have $\|\boldsymbol{A} * \boldsymbol{B}\| \leqslant \|\boldsymbol{A}\|\|\boldsymbol{B}\|$. A metric in D^n is defined by $\|\boldsymbol{x} -_D \boldsymbol{y}\| = \|\boldsymbol{x} - \boldsymbol{y}_-\|$ for $\boldsymbol{x}, \boldsymbol{y} \in D^n$. It has been proved in [4] that for $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} \in D^n$, we have $\|\boldsymbol{c} * \boldsymbol{a} -_D \boldsymbol{c} * \boldsymbol{b}\| \leqslant \|\boldsymbol{c}\|\|\boldsymbol{a} -_D \boldsymbol{b}\|$. This relation can be generalized as follows:

Proposition 1. *Let $\boldsymbol{a}, \boldsymbol{b} \in D^n$, $\boldsymbol{C} \in D^{n\times n}$. Then we have*

$$\|\boldsymbol{C} * \boldsymbol{a} -_D \boldsymbol{C} * \boldsymbol{b}\| \leqslant \|\boldsymbol{C}\|\|\boldsymbol{a} -_D \boldsymbol{b}\|.$$

Proposition 2. *Let $\boldsymbol{U} : D_1 \to D_1$, $D_1 \subseteq D^n$, be a contraction mapping in the sense that there exists $q \in \mathbb{R}$, $0 < q < 1$, such that $\|\boldsymbol{U}(\boldsymbol{x}) -_D \boldsymbol{U}(\boldsymbol{y})\| \leqslant q\|\boldsymbol{x} -_D \boldsymbol{y}\|$, for all $\boldsymbol{x}, \boldsymbol{y} \in D_1$. Then $\boldsymbol{U}$ has a fixed-point $\boldsymbol{x}^* \in D_1$, which is the limit of the sequence $\boldsymbol{x}^{(l+1)} = \boldsymbol{U}(\boldsymbol{x}^{(l)})$, $l = 0, 1, \ldots$, for any $\boldsymbol{x}^{(0)} \in D_1$.*

The proof follows the classical proof using properties of "$-_D$" such as $\boldsymbol{a} -_D \boldsymbol{a} = 0$ and $(\boldsymbol{a} -_D \boldsymbol{b}) + (\boldsymbol{b} -_D \boldsymbol{c}) = \boldsymbol{a} -_D \boldsymbol{c}$. Proposition 2 is a generalization of a fixed-point theorem from [6].

We shall consider vectors and matrices consisting of signs from $\Lambda = \{+, -\}$. The set of all n-dimensional vectors of signs $\lambda = (\lambda_1, \ldots, \lambda_n)$ with $\lambda_i \in \Lambda$ is denoted by Λ^n and the set of all $(n \times n)$-dimensional matrices of signs is denoted by $\Lambda^{n\times n}$. Matrices (vectors) of signs of same size are "multiplied" component-wise using the rules $++ = -- = +$, $+- = -+ = -$. For instance, for $\lambda = (+, -, +, -)$, $\mu = (-, +, -, +)$, we have $\lambda\mu = (-, -, -, -)$.

Similarly, the "product" of a matrix (vector) of signs by an interval matrix (vector) of the same size is defined, e.g., for $\lambda = (+, -, +, -) \in \Lambda^4$, $\boldsymbol{x} = (x_1, x_2, x_3, x_4) \in D^4$, we have $\lambda\boldsymbol{x} = (x_1, -x_2, x_3, -x_4)$. We extend further this rule for dualization as follows. Let $\lambda = (\lambda_1, \ldots, \lambda_n) \in \Lambda^n$ and $\boldsymbol{x} = (x_1, \ldots, x_n) \in D^n$. The vector $((x_1)_{\lambda_1}, \ldots, (x_n)_{\lambda_n})$ will be symbolically denoted by $\boldsymbol{x}_\lambda$. For instance, for $\lambda = (+, -, +, -) \in$

Λ^4, $\boldsymbol{x} = (x_1, x_2, x_3, x_4) \in D^4$, we have $\boldsymbol{x}_\lambda = (x_1, (x_2)_-, x_3, (x_4)_-)$. Under this convention the rules for dualization are generalized for matrices and vectors, i.e., we have that $\boldsymbol{x}_\lambda = v$ is equivalent to $\boldsymbol{x} = \boldsymbol{v}_\lambda$, $(\boldsymbol{x}_\lambda)_\mu = \boldsymbol{x}_{\lambda\mu}$, etc. According to the above we shall further assume that "products" of the form $\boldsymbol{\Gamma\Delta}$ and $\boldsymbol{\Gamma A}$, $\boldsymbol{\Gamma}$, $\boldsymbol{\Delta} \in \Lambda^{n\times n}$, $\boldsymbol{A} \in D^{n\times n}$, are well defined.

If $\boldsymbol{a} = (a_1, \ldots, a_n) \in (D^*)^n$, denote $\sigma(\boldsymbol{a}) = (\sigma(a_1), \ldots, \sigma(a_n))$ and, if $\boldsymbol{A} \in (D^*)^{n\times n}$, denote $\sigma(\boldsymbol{A}) = (\sigma(a_{ik}))$. For $\boldsymbol{x} = (x_1, \ldots, x_n) \in D^n$ we have $\boldsymbol{x}_{\sigma(\boldsymbol{a})} = ((x_1)_{\sigma(a_1)}, \ldots, (x_n)_{\sigma(a_n)})$. Using such notation we can write $\langle \boldsymbol{b}, \boldsymbol{x}_{\sigma(\boldsymbol{a})}\rangle = \sum_{i=1}^n b_i * (x_i)_{\sigma(a_i)}$, and, in particular, $\langle \boldsymbol{a}, \boldsymbol{x}_{\sigma(\boldsymbol{a})}\rangle = \sum_{i=1}^n a_i * (x_i)_{\sigma(a_i)}$.

We may further extend the above notation for a product of the form $\boldsymbol{A} * \boldsymbol{x}_{\sigma(\boldsymbol{B})}$, where $\boldsymbol{A}, \boldsymbol{B} \in (D^*)^{n\times n}$, $\boldsymbol{x} \in D^n$. To this end consider the matrix $\boldsymbol{A}$ as a system of vectors, that is $\boldsymbol{A} = (a_{ij}) = \{\boldsymbol{a}_i\}_{i=1}^n \in (D^*)^{n\times n}$, where $\boldsymbol{a}_i = (a_{i1}, \ldots, a_{in}) \in (D^*)^n$. We then have

$$\boldsymbol{A} * \boldsymbol{x}_{\sigma(\boldsymbol{B})} = \left\{\langle \boldsymbol{a}_i, \boldsymbol{x}_{\sigma(\boldsymbol{b}_i)}\rangle\right\}_{i=1}^n = \left\{\sum_{k=1}^n a_{ik} * (x_k)_{\sigma(b_{ik})}\right\}_{i=1}^n.$$

Using these notations we generalize (3) as follows:

Proposition 3. *Let $\boldsymbol{A}, \boldsymbol{B} \in (D^*)^{n\times n}$ are such that $\boldsymbol{A} + \boldsymbol{B} \in (D^*)^{n\times n}$ and let $\boldsymbol{c} \in (D^*)^n$. Then we have*

$$(\boldsymbol{A} + \boldsymbol{B}) * \boldsymbol{c}_{\sigma(\boldsymbol{A}+\boldsymbol{B})} = \boldsymbol{A} * \boldsymbol{c}_{\sigma(\boldsymbol{A})} + \boldsymbol{B} * \boldsymbol{c}_{\sigma(\boldsymbol{B})}. \tag{6}$$

Relation (6) can be also written in the form

$$(\boldsymbol{A} + \boldsymbol{B}) * \boldsymbol{c} = \boldsymbol{A} * \boldsymbol{c}_{\sigma(\boldsymbol{A})\sigma(\boldsymbol{A}+\boldsymbol{B})} + \boldsymbol{B} * \boldsymbol{c}_{\sigma(\boldsymbol{B})\sigma(\boldsymbol{A}+\boldsymbol{B})}. \tag{7}$$

Proposition 4. *Let $\boldsymbol{C} \in (D^*)^{n\times n}$ and let $\boldsymbol{a}, \boldsymbol{b} \in (D^*)^n$ are such that $\boldsymbol{a} + \boldsymbol{b} \in (D^*)^n$. Then we have*

$$\boldsymbol{C}_{\sigma(\boldsymbol{a}+\boldsymbol{b})} * (\boldsymbol{a} + \boldsymbol{b}) = \boldsymbol{C}_{\sigma(\boldsymbol{a})} * \boldsymbol{a} + \boldsymbol{C}_{\sigma(\boldsymbol{b})} * \boldsymbol{b}. \tag{8}$$

Relation (8) can be written in the form

$$\boldsymbol{C} * (\boldsymbol{a} + \boldsymbol{b}) = \boldsymbol{C}_{\sigma(\boldsymbol{a})\sigma(\boldsymbol{a}+\boldsymbol{b})} * \boldsymbol{a} + \boldsymbol{C}_{\sigma(\boldsymbol{b})\sigma(\boldsymbol{a}+\boldsymbol{b})} * \boldsymbol{b}. \tag{9}$$

Note that relations (6) and (7) are valid in particular for any $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{(n\times n)}$ and similarly, (8) and, (9) hold true for any $\boldsymbol{C} \in D^{n\times n}$ and $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^n$.

4. An iterative interval method

Given an interval matrix $\boldsymbol{A} \in D^{n\times n}$ denote $\boldsymbol{T} = \boldsymbol{T}(\boldsymbol{A}) = (t_{ij})$ with $t_{ii} = a_{ii}$, $t_{ij} = 0$, $i \neq j$. Assuming $a_{ii} \in D^*$, $i = 1, \ldots, n$, denote $\boldsymbol{T}^{-1} = \boldsymbol{T}^{-1}(\boldsymbol{A}) = (t_{ij}^*)$, $t_{ij}^* = 1/(a_{ii})_-$, $t_{ij}^* = 0$, $i \neq j$. Clearly, $\boldsymbol{T}^{-1} * \boldsymbol{T} = 1$.

Consider the following interval-valued analogue of the Jacobi iteration procedure:

$$\boldsymbol{x} := \boldsymbol{T}^{-1} * \left(\boldsymbol{b} -_D (\boldsymbol{A} -_D \boldsymbol{T}) * \boldsymbol{x}\right). \tag{10}$$

In (10), $\boldsymbol{A} -_D \boldsymbol{T} = \boldsymbol{A} - \boldsymbol{T}_-$ is the matrix $\boldsymbol{A}$ with diagonal elements replaced by zero. Using distributivity relations (3) and (5)–(9), it is easy to see that both systems of equations (1) and (10) are algebraically equivalent.

Proposition 5. *Let the matrix $\boldsymbol{A} \in D^{n\times n}$ satisfy $a_{ii} \in D^*$, $i = 1, \ldots, n$, together with $\|\boldsymbol{T}^{-1}\|\,\|\boldsymbol{A} -_D \boldsymbol{T}\| \leqslant q < 1$. Then* (1) *has a unique solution $\boldsymbol{x}_s \in D^n$ and method (10) converges to $\boldsymbol{x}_s$ for any $\boldsymbol{b} \in D^n$ and any initial approximation $\boldsymbol{x}^{(0)} \in D^n$.*

Proof. For $\boldsymbol{x} \in D^n$ denote $\boldsymbol{B}(\boldsymbol{x}) = \boldsymbol{T}^{-1} * (\boldsymbol{b} -_D (\boldsymbol{A} -_D \boldsymbol{T}) * \boldsymbol{x})$. For $\boldsymbol{x}, \boldsymbol{y} \in D^n$ we have

$$\begin{aligned} \|\boldsymbol{B}(\boldsymbol{x}) -_D \boldsymbol{B}(\boldsymbol{y})\| &= \|\boldsymbol{T}^{-1} * \big(\boldsymbol{b} -_D (\boldsymbol{A} -_D \boldsymbol{T}) * \boldsymbol{x}\big) -_D \boldsymbol{T}^{-1} * \big(\boldsymbol{b} -_D (\boldsymbol{A} -_D \boldsymbol{T}) * \boldsymbol{y}\big)\| \\ &\leqslant \|\boldsymbol{T}^{-1}\| \|\big(\boldsymbol{b} -_D (\boldsymbol{A} -_D \boldsymbol{T}) * \boldsymbol{x}\big) -_D \big(\boldsymbol{b} -_D (\boldsymbol{A} -_D \boldsymbol{T}) * \boldsymbol{y}\big)\| \\ &= \|\boldsymbol{T}^{-1}\| \|(\boldsymbol{A} -_D \boldsymbol{T}) * \boldsymbol{y} -_D (\boldsymbol{A} -_D \boldsymbol{T}) * \boldsymbol{x}\| \\ &\leqslant \|\boldsymbol{T}^{-1}\| \|\boldsymbol{A} -_D \boldsymbol{T}\| \|\boldsymbol{y} -_D \boldsymbol{x}\| < q\|\boldsymbol{y} -_D \boldsymbol{x}\|, \end{aligned}$$

using Proposition 1. The inequality $\|\boldsymbol{B}(\boldsymbol{x}) -_D \boldsymbol{B}(\boldsymbol{y})\| \leqslant q\|\boldsymbol{y} -_D \boldsymbol{x}\|$ shows that $\boldsymbol{B}$ is a contraction mapping. This combined with Proposition 2 and the fact that (1) and (10) are algebraically equivalent proves the theorem. □

In [19] the following iteration method for the solution to (1) with $\boldsymbol{b} \in D^n$ has been proposed:

$$x_i := \left(b_i -_D \sum_{j=1, j\neq i}^{n} a_{ij} * x_j\right) \Big/ (a_{ii})_-, \quad i = 1, \ldots, n. \tag{11}$$

Obviously, (11) is a component-wise form of the matrix expression (10). Kupriyanova proves convergence of the iterative process (11) to the solution of problem (1) under special (implicit) restrictions on the input data $\boldsymbol{A}, \boldsymbol{b}$ and on the initial approximation [5]. Proposition 5 requires explicit restrictions on the matrix and no restrictions on the initial approximations.

Remark. If in Proposition 5 both inequalities $\|\boldsymbol{T}(\boldsymbol{A})^{-1}\| \leqslant q < 1$, and $\|\boldsymbol{A} -_D \boldsymbol{D}(\boldsymbol{A})\| \leqslant q < 1$ are assumed, then we obtain $\|\boldsymbol{B}(\boldsymbol{x}) -_D \boldsymbol{B}(\boldsymbol{y})\| < q^2 \|\boldsymbol{y} -_D \boldsymbol{x}\|$, showing better convergence.

Our iteration method can be also applied (with some modifications) to the more general problem

$$\boldsymbol{A} * \boldsymbol{x}_{\boldsymbol{\Gamma}} + \boldsymbol{B} * \boldsymbol{x}_{\boldsymbol{\Delta}} = \boldsymbol{c}, \tag{12}$$

where $\boldsymbol{A}, \boldsymbol{B} \in D^{n\times n}$, $\boldsymbol{\Gamma}, \boldsymbol{\Delta} \in \Lambda^{n\times n}$ and $\boldsymbol{c} \in D^n$. We note that (12) is not the most general form of systems of equations of the type (1) (cf. [13]). However, in the case of real-valued (point) matrix coefficients problem (12) presents the general situation. The important special case of point matrices will be considered in the next section. We shall end this section by considering (12) in the one-dimensional case. Then we obtain a single equation of the form

$$a * y + b * y_\lambda = v, \quad \lambda \in \Lambda = \{+, -\}, \tag{13}$$

with respect to y. From (3) we see that in the case $\sigma(a)\sigma(b) = \lambda$ the expression $a * y + b * y_\lambda$ is algebraically equivalent to $y * (a + b)_{\sigma(a)\sigma(a+b)}$ and therefore the solution y of $a * y + b * y_\lambda = v$ is $y = (v/(a + b)_-)_{\sigma(a)\sigma(a+b)}$. A presentation of the expression $a * y + b * y_\lambda$ in the form $y * c$ is not possible in the case $\sigma(a)\sigma(b) = -\lambda$. The next proposition shows that (13) has a unique solution in the general case when no constraint $\sigma(a)\sigma(b) = \lambda$ is assumed.

Proposition 6. *Given are $a, b, v \in D \setminus \mathcal{T}^*$, $\lambda \in \Lambda = \{+, -\}$. Assume that $s' = a_{\sigma(a)} * a_{\lambda\sigma(b)} - b_{-\sigma(b)} * b_{-\lambda\sigma(a)} \in D^*$. Eq.* (13) *has the unique solution*

$$y = \left(a_{\lambda\delta\sigma(b)} * \frac{v}{s} - b_{-\delta\sigma(a)} * \left(\frac{v}{s}\right)_{-\lambda}\right)_{\sigma(s')}, \qquad s = s'_{-\delta\sigma(s)}, \quad \delta = \sigma(v)\sigma(y). \tag{14}$$

To determine the sign $\sigma(y)$ in (14) *consider the following two cases:*

(i) $\sigma(a)=\sigma(b)$. *In this case it can be easily seen from* (13) *that* $\sigma(y)=\sigma(v)\sigma(a)$, *hence* $\sigma(a)=\sigma(b)=\sigma(v)\sigma(y)=\delta$. *The solution* (14) *obtains the form*

$$y=\left(a_\lambda * \frac{v}{s} - b_- * \left(\frac{v}{s}\right)_{-\lambda}\right)_{\sigma(s)}, \quad s=(a*a_\lambda - b_- * b_{-\lambda})_{-\sigma(s)}.$$

(ii) $\sigma(a)=-\sigma(b)$. *Then from* (14) *we see that* $\sigma(y)=\sigma(v)\sigma(s)\sigma(a)$ *and the expression for y obtains the form*

$$y=a_{-\lambda} * \left(\frac{v}{s}\right)_{\sigma(s)} - b_- * \left(\frac{v}{s}\right)_{-\lambda\sigma(s)}, \quad s=a_- * a_\lambda - b_- * b_\lambda.$$

5. Special case: Point matrix, interval right-hand side

In the special case of point matrix and interval right-hand side we are able to find simple conditions for the existence and uniqueness of the solution and to reduce the problem to the familiar numerical case. We shall make use of the following definition (cf. [17]):

Definition. Let $\boldsymbol{A}=\{a_{i,k}\}\in\mathbb{R}^{n\times n}$ be a real-valued (point) matrix. We say that the matrix $\boldsymbol{A}$ is *completely nonsingular*, if both matrices $\boldsymbol{A}$ and $|\boldsymbol{A}|=\{|a_{i,k}|\}$ are nonsingular.

In the next proposition we make use of the center-radius presentation of an interval vector writing the components of $\boldsymbol{b}=\{b_i\}\in D^n$ in the form $b_i=(b_i', b_i'')\in D$. Denote the vector of the centers by $\boldsymbol{b}'=\{b_i'\}_{i=1}^n\in\mathbb{R}^n$ and the vector of the radii by $\boldsymbol{b}''=\{b_i''\}_{i=1}^n\in\mathbb{R}^n$, so that $\boldsymbol{b}=(\boldsymbol{b}',\boldsymbol{b}'')\in D^n$.

Proposition 7. *Let $\boldsymbol{A}=(\alpha_{i,k})\in\mathbb{R}^{n\times n}$ be completely nonsingular and $\boldsymbol{b}=(\boldsymbol{b}',\boldsymbol{b}'')=\{(b_i', b_i'')\}_{i=1}^n\in D^n$ be a given interval vector. Then* (1) *has a unique solution $\boldsymbol{x}=(\boldsymbol{x}',\boldsymbol{x}'')=\{(x_i',x_i'')\}$, such that $\boldsymbol{x}'=\{x_i'\}\in\mathbb{R}^n$ is the solution of the linear system $\boldsymbol{A}\boldsymbol{x}'=\boldsymbol{b}'$, and $\boldsymbol{x}''=\{x_i''\}\in\mathbb{R}^n$ is the solution of the linear system $|\boldsymbol{A}|\boldsymbol{x}''=\boldsymbol{b}''$.*

Proof. System (1) in component-wise form is

$$\begin{aligned}&\alpha_{11}*x_1+\alpha_{12}*x_2+\cdots+\alpha_{1n}*x_n=b_1,\\&\quad\vdots\\&\alpha_{n1}*x_1+\alpha_{n2}*x_2+\cdots+\alpha_{nn}*x_n=b_n,\end{aligned}\tag{15}$$

where $\alpha_{ij}\in\mathbb{R}$, $b_i\in D$, $i=1,\ldots,n$. Substituting in (15) the center-radius presentations for the involved intervals $x_i=(x_i',x_i'')$, $b_i=(b_i',b_i'')\in D$, we reduce system (1) to two (real-valued) linear systems for the coordinates of the unknowns. The centers $x_i'\in\mathbb{R}^n$ satisfy the linear system

$$\begin{aligned}&\alpha_{11}\cdot x_1'+\alpha_{12}\cdot x_2'+\cdots+\alpha_{1n}\cdot x_n'=b_1',\\&\quad\vdots\\&\alpha_{n1}\cdot x_1'+\alpha_{n2}\cdot x_2'+\cdots+\alpha_{nn}\cdot x_n'=b_n',\end{aligned}\tag{16}$$

and the radii $x_i'' \in \mathbb{R}$ satisfy

$$
\begin{aligned}
&|\alpha_{11}| \cdot x_1'' + |\alpha_{12}| \cdot x_2'' + \cdots + |\alpha_{1n}| \cdot x_n'' = b_1'', \\
&\quad\vdots \\
&|\alpha_{n1}| \cdot x_1'' + |\alpha_{n2}| \cdot x_2'' + \cdots + |\alpha_{nn}| \cdot x_n'' = b_n''.
\end{aligned}
\tag{17}
$$

Using that $\boldsymbol{A}$ is completely nonsingular, observe that (16) and (17) have unique solutions, hence the proposition follows. □

Note that using center-radius form we are able to prove Proposition 7 by splitting the original problem into two numerical problems of the *same* size $(n \times n)$ for the numerical components of the intervals (using center-radius form). The use of endpoint presentation of intervals generally leads to numerical problems of double size $(2n \times 2n)$ (cf., e.g., [17]).

Our technique can be applied for the general problem (12): $\boldsymbol{A} * \boldsymbol{x}_{\boldsymbol{\Gamma}} + \boldsymbol{B} * \boldsymbol{x}_{\boldsymbol{\Delta}} = \boldsymbol{c}$, involving real matrices $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{n\times n}$, and arbitrary sign matrices $\boldsymbol{\Gamma}, \boldsymbol{\Delta} \in \Lambda^{n\times n}$.

The following proposition states that problem (12) splits into two numerical linear problems $(\boldsymbol{A} + \boldsymbol{B})\boldsymbol{x}' = \boldsymbol{c}'$, $(\boldsymbol{\Gamma}|\boldsymbol{A}| + \boldsymbol{\Delta}|\boldsymbol{B}|)\boldsymbol{x}'' = \boldsymbol{c}''$.

Proposition 8. *Let $\boldsymbol{\Gamma}, \boldsymbol{\Delta} \in \Lambda^{n\times n}$, $\boldsymbol{c} \in D^n$, and let $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{n\times n}$ are such that $\boldsymbol{A} + \boldsymbol{B}$ and $\boldsymbol{\Gamma}|\boldsymbol{A}| + \boldsymbol{\Delta}|\boldsymbol{B}|$ are nonsingular. Then* (12) *has a unique solution $\boldsymbol{x} = (\boldsymbol{x}', \boldsymbol{x}'')$, where $\boldsymbol{x}'$ is solution of the linear system $(\boldsymbol{A} + \boldsymbol{B})\boldsymbol{x}' = \boldsymbol{c}'$, and $\boldsymbol{x}''$ is solution of the linear system $(\boldsymbol{\Gamma}|\boldsymbol{A}| + \boldsymbol{\Delta}|\boldsymbol{B}|)\boldsymbol{x}'' = \boldsymbol{c}''$.*

Corollary. *Let $\lambda \in \{+, -\}$, $\boldsymbol{c} \in D^n$, and let $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{n\times n}$ are such that both matrices $\boldsymbol{A} + \boldsymbol{B}$ and $|\boldsymbol{A}| + \lambda|\boldsymbol{B}|$ are nonsingular. Equation $\boldsymbol{A} * \boldsymbol{x} + \boldsymbol{B} * \boldsymbol{x}_\lambda = \boldsymbol{c}$ has a unique solution $\boldsymbol{x} = (\boldsymbol{x}', \boldsymbol{x}'')$, where $\boldsymbol{x}'$ is the solution of the linear system $(\boldsymbol{A} + \boldsymbol{B})\boldsymbol{x}' = \boldsymbol{c}'$, and $\boldsymbol{x}''$ is the solution of the linear system $(|\boldsymbol{A}| + \lambda|\boldsymbol{B}|)\boldsymbol{x}'' = \boldsymbol{c}''$.*

To demonstrate the technique for equivalent algebraic transformation of directed interval expressions we prove below the special case when the dimension of the system is one, that is the problem (12) consists of one equation only. This case can be considered as a special case of Proposition 6 with degenerate coefficients.

Proposition 9. *Let $p, q \in \mathbb{R}$, $\lambda \in \{+, -\}$, $d \in D$, $s = p^2 - q^2 \neq 0$. Then equation*

$$p * x + q * x_\lambda = d \tag{18}$$

has a unique solution

$$x = s^{-1} * (p * d - q * d_{-\lambda})_{\sigma(s)}. \tag{19}$$

Proof. Denote $\sigma = \sigma(s)$ and substitute (19) into (18):

$$
\begin{aligned}
p * x + q * x_\lambda &= \left(ps^{-1}\right) * (p * d - q * d_{-\lambda})_\sigma + \left(qs^{-1}\right) * (p * d - q * d_{-\lambda})_{\lambda\sigma} \\
&= \left(p^2 s^{-1}\right) * d_\sigma + \left(-pqs^{-1}\right) * d_{-\lambda\sigma} + \left(qps^{-1}\right) * d_{\lambda\sigma} + \left(-q^2 s^{-1}\right) * d_{-\sigma} \\
&= \left(p^2 s^{-1}\right) * d_\sigma + \left(-q^2 s^{-1}\right) * d_{-\sigma} = s^{-1} * \left(p^2 * d + \left(-q^2\right) * d_-\right)_\sigma \\
&= s^{-1} * \left(\left(p^2 - q^2\right) * d_\sigma\right)_\sigma = \left(s^{-1}\left(p^2 - q^2\right)\right) * d,
\end{aligned}
$$

where for the equivalent transformation in the last line we use the distributive relation (4). The last quantity is obviously equal to d, which proves that (19) is a solution to (18). Noticing that all above algebraic tranformations are equivalent, we see that (19) is the unique solution to (18). $\square$

The next proposition shows that in certain special cases we can write the solution to problem (1) with point matrix $\boldsymbol{A}$ in the form of an interval Cramer-type formula.

Proposition 10. *Let $A = (a_{i,k}) \in \mathbb{R}^{n\times n}$ 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 $i, k = 1, 2, \ldots, n$. Then for the solution to $\boldsymbol{A} * \boldsymbol{x} = \boldsymbol{b}$ the following Cramer-type formula holds:*

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

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

The proof is easily obtained using the properties of directed intervals. A class of matrices satisfying the conditions of Proposition 10 is the class of Vandermonde matrices, appearing in interpolation theory.

6. Numerical algorithm and experiments

The following simple numerical procedure for the solution of (step 1 is based on Proposition 6).

1. Check the condition $\|\boldsymbol{T}^{-1}\|\|\boldsymbol{A} -_D \boldsymbol{T}\| < 1$.
2. Using a random initial approximation $\boldsymbol{x}^{(0)}$ iterate according to

$$\boldsymbol{x}^{(k+1)} := \boldsymbol{T}^{-1} * \big(\boldsymbol{b} -_D (\boldsymbol{A} -_D \boldsymbol{T}) * \boldsymbol{x}^{(k)}\big), \quad k = 0, 1, \ldots. \tag{20}$$

Using the inclusion properties of the interval arithmetic operations (see, e.g., [4]) it is possible to construct a modification of the above interval algorithm, using suitable computer-arithmetic operations and delivering the result with automatic verification. Such verified algorithm was implemented and tested using an experimental package for directed interval arithmetic [14]. A discussion on the software implementation is given in [11]. For the numerical applications we need component-wise and/or center-radius presentations of the interval operations. A brief summary of these presentations is given in Appendix A.

The following examples are solved by means of the package announced in [11]. All results are obtained in several iterations using random initial approximations.

Example 1. Consider the system [3]

$$\begin{pmatrix} [0.7, 1.3] & [-0.3, 0.3] & [-0.3, 0.3] \\ [-0.3, 0.3] & [0.7, 1.3] & [-0.3, 0.3] \\ [-0.3, 0.3] & [-0.3, 0.3] & [0.7, 1.3] \end{pmatrix} * \boldsymbol{x} = \begin{pmatrix} [-14, 7] \\ [9, 12] \\ [3, 3] \end{pmatrix}.$$

The algebraic solution obtained by the iteration method (10) using random initial approximation is

$$\boldsymbol{x} = \begin{pmatrix} [-9.125, -13.0536] \\ [16.7679, 7.125] \\ [11.25, -2.67857] \end{pmatrix}.$$

Example 2. Consider the system [12]

$$\begin{pmatrix} [3.7, 4.3] & [-1.5, -0.5] & [0, 0] \\ [-1.5, -0.5] & [3.7, 4.3] & [-1.5, -0.5] \\ [0, 0] & [-1.5, -0.5] & [3.7, 4.3] \end{pmatrix} * \boldsymbol{x} = \begin{pmatrix} [-14, 14] \\ [-9, 9] \\ [-3, 3] \end{pmatrix}.$$

The solution obtained by the iteration method (10) using random initial approximation is

$$\boldsymbol{x} = \begin{pmatrix} [-2.92668, 2.92668] \\ [-0.943531, 0.943531] \\ [-0.368536, 0.368536] \end{pmatrix}.$$

Changing in the above problem only the vector $\boldsymbol{b}$, here are some results:

$$\boldsymbol{b} = \begin{pmatrix} [-14, 0] \\ [-9, 0] \\ [-3, 0] \end{pmatrix}; \qquad \boldsymbol{x} = \begin{pmatrix} [-3.46158, -0.936849] \\ [-2.3109, -1.7696] \\ [-0.903442, -0.936889] \end{pmatrix}.$$

$$\boldsymbol{b} = \begin{pmatrix} [0, 14] \\ [0, 9] \\ [0, 3] \end{pmatrix}; \qquad \boldsymbol{x} = \begin{pmatrix} [0.936849, 3.46158] \\ [1.7696, 2.3109] \\ [0.936889, 0.903442] \end{pmatrix}.$$

$$\boldsymbol{b} = \begin{pmatrix} [2, 14] \\ [-9, -3] \\ [-3, 1] \end{pmatrix}; \qquad \boldsymbol{x} = \begin{pmatrix} [0.392969, 2.86724] \\ [-1.11391, -1.09203] \\ [-0.824654, -0.181313] \end{pmatrix}.$$

$$\boldsymbol{b} = \begin{pmatrix} [2, 14] \\ [3, 9] \\ [-3, 1] \end{pmatrix}; \qquad \boldsymbol{x} = \begin{pmatrix} [1.46332, 3.54147] \\ [2.45663, 2.2762] \\ [0.111972, 0.518212] \end{pmatrix}.$$

Example 3. This is an example from [3]:

$$\begin{pmatrix} [2, 3] & [0, 1] \\ [1, 2] & [2, 3] \end{pmatrix} * \boldsymbol{x} = \begin{pmatrix} [0, 120] \\ [60, 240] \end{pmatrix}; \qquad \boldsymbol{x} = \begin{pmatrix} [0, 17.1429] \\ [30, 68.5714] \end{pmatrix}.$$

7. Conclusion

We formulate explicit conditions on the matrix in (1) and prove that under these conditions the interval Jacobi-type method (11) converges to the solution to (1) with an *arbitrary* initial approximation and *arbitrary* right-hand side. The presented techniques may be used to formulate and prove convergence of other iteration methods, which are interval analogues of familiar iteration methods for the numeric case. From our results we conclude that algebraic transformations based on directed interval arithmetic can be successfully used for the formulation of iterative procedures for the solution of the interval algebraic problem (1), respectively for the analysis of the solution (e.g., with respect to its convergence). Thereby directed interval arithmetic can be used in a way much similar to using real arithmetic.

Directed interval arithmetic is the natural arithmetic for the solution of algebraic equations with interval coefficients, since it is obtained from the arithmetic for normal intervals via algebraic completion.

Solving interval algebraic equations using only proper intervals can be compared to solving real linear algebraic equations when restricting ourselves only to positive integer numbers (Diophantine equations). Indeed, in both cases we stay within semigroup systems and make no use of inverse additive/multiplicative operations.

Acknowledgements

The present work arose from a lecture at the minisymposium organized by R. Weiss and W. Schönauer within the 15th IMACS World Congress. The author is grateful to the organizers of the minisymposium for their kind invitation to participate. The reeferee's suggestions and remarks stimulated many additional improvements. The author is also grateful to E. Popova for fruitful discussions and help with the numerical experiments. This work was partially supported by the National Science Fund under grant No. MM-521/95.

Appendix A

Below we give some formulae and expressions for the numerical component-wise presentations of intervals, using the endpoint and center-radius presentation proposed in [4].

A.1. End-point form

The set of directed intervals D is equivalent to the set of all ordered pairs of real numbers $\{[\alpha, \beta] \mid \alpha, \beta \in \mathbb{R}\}$ [4]. The first component (endpoint) of $a \in D$ is further denoted by a^-, and the second by a^+, so that $a = [a^-, a^+]$. Thus $a^\lambda \in \mathbb{R}$ with $\lambda \in \Lambda = \{+, -\}$ is the first or the second component of $a \in D$ depending on the value of λ. The directed interval $a = [a^-, a^+]$ is *proper* (*normal*) if $a^- \leqslant a^+$, *degenerate* if $a^- = a^+$, and *improper* if $a^- > a^+$. The set of all proper intervals is equivalent to $I(\mathbb{R})$, the set of degenerate intervals is equivalent to $\mathbb{R}$, and the set of improper intervals is $\overline{I(\mathbb{R})}$. We have $D = I(\mathbb{R}) \cup \overline{I(\mathbb{R})}$.

To every directed interval $a = [a^-, a^+] \in D$ corresponds a binary variable *type* or *direction*, defined by

$$\tau(a) = \begin{cases} +, & \text{if } a^- \leqslant a^+, \\ -, & \text{if } a^- > a^+. \end{cases}$$

We have

$$\overline{Z} = \{a \in \overline{I(\mathbb{R})} \mid a^+ \leqslant 0 \leqslant a^-\}, \qquad \overline{Z}^* = \{a \in \overline{I(\mathbb{R})} \mid a^+ < 0 < a^-\},$$

$\mathcal{T} = Z \cup \overline{Z}$, $\mathcal{T}^* = Z^* \cup \overline{Z}^*$, $D^* = D \setminus \mathcal{T}$.

The *sign* of a directed interval $\sigma : D \setminus \mathcal{T}^* \to \Lambda$ can be expressed by

$$\sigma(a) = \begin{cases} +, & \text{if } a^-, a^+ \geqslant 0, \\ -, & \text{if } a^- < 0,\ a^+ \leqslant 0. \end{cases}$$

Addition "+" in D can be expressed endpoint-wise as

$$a+b=\left[a^-+b^-, a^++b^+\right], \quad a,b\in D.$$

An endpoint-wise presentation of multiplication "$*$" in D is given by

$$a*b=\begin{cases}\left[a^{-\sigma(b)}b^{-\sigma(a)},\ a^{\sigma(b)}b^{\sigma(a)}\right], & a,b\in D\setminus\mathcal{T}^*,\\ \left[a^{\delta\tau(b)}b^{-\delta},\ a^{\delta\tau(b)}b^{\delta}\right], & \delta=\sigma(a),\ a\in D\setminus\mathcal{T}^*,\ b\in\mathcal{T}^*,\\ \left[a^{-\delta}b^{\delta\tau(a)},\ a^{\delta}b^{\delta\tau(a)}\right], & \delta=\sigma(b),\ a\in\mathcal{T}^*,\ b\in D\setminus\mathcal{T}^*,\end{cases} \tag{A.1}$$

$$a*b=\begin{cases}\left[\min\left\{a^-b^+,a^+b^-\right\},\ \max\left\{a^-b^-,a^+b^+\right\}\right], & a,b\in Z^*,\\ \left[\max\left\{a^-b^-,a^+b^+\right\},\ \min\left\{a^-b^+,a^+b^-\right\}\right], & a,b\in\overline{Z}^*,\\ 0, & \left(a\in Z^*,\ b\in\overline{Z}^*\right)\vee\left(a\in\overline{Z}^*,\ b\in Z^*\right).\end{cases} \tag{A.2}$$

If one of the multipliers in (A.1) is degenerate, than for any $[a,a]=a\in\mathbb{R}$, $b\in D$ we obtain $a*b=[ab^{-\sigma(a)},\ ab^{\sigma(a)}]$. This implies that the case $a\in D\setminus\mathcal{T}^*$, $b\in\mathcal{T}^*$ in (A.1) can be written as $a*b=a^{\sigma(a)\tau(b)}*b$.

Negation is expressed by $-b=(-1)*b=[-b^+,-b^-]$. The composite operation $a+(-1)*b=a+(-b)=a-b=[a^--b^+,a^+-b^-]$ is *subtraction*.

The inverse is presented component-wise as $-_D a=[-a^-,-a^+]$, for $a\in D$, and $1/_D a=[1/a^-,1/a^+]$, for $a\in D^*$. We have component-wise for the dualization: $[a^-,a^+]_-=[a^+,a^-]$, and for the reciprocal: $1/[a^-,a^+]=[1/a^+,1/a^-]$, $a\in D^*$; also $[a^-,a^+]_\lambda=[a^{-\lambda},a^\lambda]$.

Inclusion in D is $A\subset B$ if and only if $a^-\geqslant b^-$ and $a^+\leqslant b^+$.

If $x\in D$ we have

$$\|x\|=\inf_t\{t*j_-\subset x\subset t*j\}=\max\left\{\left|x^-\right|,\left|x^+\right|\right\}.$$

Other norms are $\|x\|_2=|x^-|+|x^+|$ and $\|x\|_3=(|x^-|^2+|x^+|^2)^{1/2}$. The induced metric in D is

$$r(x,y)=\|x-_D y\|=\max\left\{\left|x^--y^-\right|,\left|x^+-y^+\right|\right\},$$
$$\|x-_D y\|_2=\left|x^--y^-\right|+\left|x^+-y^+\right|,$$
$$\|x-_D y\|_3=\left(\left|x^--y^-\right|^2+\left|x^+-y^+\right|^2\right)^{1/2}.$$

If $(c_1,\dots,c_n)\in D^n$ we have

$$\left\|(c_1,\dots,c_n)\right\|=\max_i\left\{\|c_i\|\right\}_{i=1}^n=\max_i\left\{\left|c_i^-\right|,\ \left|c_i^+\right|\right\}_{i=1}^n.$$

For the other two norms we have

$$\left\|(c_1,\dots,c_n)\right\|_2=\sum_{i=1}^n\|c_i\|_2=\sum_{i=1}^n\left(\left|c_i^-\right|+\left|c_i^+\right|\right);$$
$$\left\|(c_1,\dots,c_n)\right\|_3=\left(\sum_{i=1}^n\|c_i\|_3^2\right)^{1/2}=\left(\sum_{i=1}^n\left(\left|c_i^-\right|^2+\left|c_i^+\right|^2\right)\right)^{1/2}.$$

A.2. Center-radius form

Denote the center of a by a' and the radius by a''. Both a',a'' are elements of $\mathbb{R}$; the radius of an improper intervals is negative and we have $\tau(a)=\sigma(a'')$. An interval in center-radius form is denoted by (a',a''). The transition formulae between both forms (endpoint and center-radius) are:

(i) $a' = \frac{1}{2}(a^- + a^+)$, $a'' = \frac{1}{2}(a^+ - a^-)$, and (ii) $a^+ = a' + a''$, $a^- = a' - a''$. In center-radius form addition has the form

$$(a', a'') + (b', b'') = (a' + b', a'' + b''), \quad a, b \in D.$$

Interval multiplication is given by

$$(a', a'') * (b', b'') = \begin{cases} (a'b' + \sigma(b)\sigma(a)a''b'', |b'|a'' + |a'|b''), & \text{if } a, b \in D \setminus \mathcal{T}, \\ (a'b' + \sigma(b)\tau(a)a'b'', |b'|a'' + |a''|b''), & \text{if condition C}, \\ 0, & \text{if } a, b \in \mathcal{T},\ \tau(a) = -\tau(b), \end{cases} \tag{A.3}$$

Condition C in (A.3) is: either (C1) $a \in \mathcal{T}$, $b \in D \setminus \mathcal{T}$, or (C2) $a, b \in \mathcal{T}$, $\tau(a) = \tau(b)$, $\chi(a) \leqslant \chi(b)$, where χ is defined for $a \neq 0$ by

$$\chi(a) = a^{-\delta}/a^{\delta} = (a' + \delta a'')/(a' - \delta a''), \quad \delta = \tau(a)\sigma(a).$$

The interval multiplication by scalar is a special case of (A.3):

$$\alpha * (b', b'') = (\alpha b', |\alpha| b''), \quad \alpha \in \mathbb{R},\ b \in D. \tag{A.4}$$

Below we summarize the presentations of the operators opposite, negation, dualization, reciprocal and inverse, both in endpoint and center-radius form:

$$\begin{aligned} &-_D[a^-, a^+] = [-a^-, -a^+], && -_D(a', a'') = (-a', -a''), \\ &-[a^-, a^+] = [-a^+, -a^-], && -(a', a'') = (-a', a''), \\ &[a^-, a^+]_- = [a^+, a^-], && (a', a'')_- = (a', -a''), \\ &1/[a^-, a^+] = [1/a^+, 1/a^-], && 1/(a', a'') = (a'/\Delta(a), a''/\Delta(a)), \\ &1/_D[a^-, a^+] = [1/a^-, 1/a^+], && 1/(a', a'') = (a'/\Delta(a), -a''/\Delta(a)), \end{aligned}$$

where $\Delta(a) = (a')^2 - (a'')^2 = a^- a^+$.

Inclusion can be expressed in center-radius presentation by $a \subset b$ if and only if $|b' - a'| \leqslant b'' - a''$. Note that the latter condition splits to two simultaneous restrictions $b'' - a'' = \gamma \geqslant 0$ and $|b' - a'| \leqslant \gamma$.

Norm, metric. The following functionals are norms in D:

$$\|c\|^{(r)} = \max\{|c'|, |c''|\}; \qquad \|c\|_2^{(r)} = |c'| + |c''|; \qquad \|c\|_3^{(r)} = \left(|c'|^2 + |c''|^2\right)^{1/2}.$$

In D^n we may take respectively

$$\|(c_1, \ldots, c_n)\|^{(r)} = \max_i \{\|c_i\|^{(r)}\} = \max_i \{|c_i'|, |c_i''|\};$$

$$\|(c_1, \ldots, c_n)\|_2^{(r)} = \sum_{i=1}^n \|c_i\|_2^{(r)} = \sum_{i=1}^n (|c_i'| + |c_i''|);$$

$$\|(c_1, \ldots, c_n)\|_3^{(r)} = \left(\sum_{i=1}^n \left(\|c_i\|_3^{(r)}\right)^2\right)^{1/2} = \left(\sum_{i=1}^n \left(|c_i'|^2 + |c_i''|^2\right)\right)^{1/2}.$$

References

[1] N. Dimitrova, S. Markov and E. Popova, Extended interval arithmetics: New results and applications, in: L. Atanassova and J. Herzberger, eds., *Computer Arithmetic and Enclosure Methods* (North-Holland, Amsterdam, (1992) 225–232.

[2] E. Gardenes, A. Trepat and H. Mielgo, Present perspective of the SIGLA interval system, Freiburger Intervall-Berichte 82/9, Inst. für Angewandte Mathematik, Univ. Freiburg (1982) 1–65.

[3] E. Hansen, Bounding the solution of interval linear equations, *SIAM J. Numer. Anal.* 29 (5) (1992) 1493–1503.

[4] E. Kaucher, Interval analysis in the extended interval space $I\mathbb{R}$, *Computing Suppl.* 2 (1980) 33–49.

[5] L. Kupriyanova, Inner estimation of the united solution set to interval linear algebraic system, *Reliable Computing* 1 (1) (1995) 15–31.

[6] S. Markov, Some applications of the extended interval arithmetic to interval iterations, *Computing Suppl.* 2 (1980) 69–84.

[7] S. Markov, On directed interval arithmetic and its applications, *J. UCS* 1 (7) (1995) 510–521.

[8] S. Markov, Isomorphic embeddings of interval systems, *Reliable Computing* 3 (3) (1997) 199–207.

[9] S. Markov, On iterative methods for linear algebraic equations with interval coefficients, in: J. Sydow, ed., *Proc. 15th IMACS World Congress on Scientific Computation, Modelling and Applied Mathematics*, Vol. 2, *Numerical Mathematics* (1997) 575–580.

[10] S. Markov, On the algebra of intervals and convex bodies, *J. UCS* 4 (1) (1998) 34–47.

[11] S. Markov, E. Popova and C. Ullrich, On the solution of linear algebraic equations involving interval coefficients, in: S. Margenov and P. Vassilevski, eds., *Iterative Methods in Linear Algebra II*, IMACS Series in Computational and Applied Mathematics, Vol. 3 (1996) 216–225.

[12] S. Ning and R.B. Kearfott, A comparison of some methods for solving linear interval equations, *SIAM J. Numer. Anal.* 34 (4) (1997) 1289–1305.

[13] E. Popova, Algebraic solutions to a class of interval equations, *J. UCS* 4 (1) (1998) 48–67.

[14] E. Popova and C. Ullrich, Directed interval arithmetic in mathematica: Implementation and applications, Technical Report 96-3, University of Basel, 1996, 1–56.

[15] S. Shary, Solving the linear interval tolerance problem, *Math. Comput. Simulation* 39 (1995) 53–85.

[16] S. Shary, Algebraic approach to the interval linear static identification, tolerance and control problems, or one more application of Kaucher arithmetic, *Reliable Computing* 2 (1) (1996) 1–32.

[17] S. Shary, Algebraic approach in the “outer problem” for interval linear equations, *Reliable Computing* 3 (2) (1997) 103–135.

[18] S. Shary, Controllable solution set to interval static systems, *Appl. Math. Comput.* 86 (1997) 185–196.

[19] V. Zyuzin, 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).