Milan Hladík

Understanding and Using Interval Analysis

January 13, 2022

Contents

Notation

We mostly follow the standardized interval notation from Kearfott et al. (2002); cf. Neumaier's web page `http://www.mat.univie.ac.at/~neum/software/int/`.

Sets

$\operatorname{conv} \mathcal{S}$	the convex hull of a set $\mathcal{S} \subset \mathbb{R}^n$, p. 14
$\operatorname{int} \mathcal{S}$	the interior of a set $\mathcal{S} \subset \mathbb{R}^n$
Σ	the solution set to $\boldsymbol{A}x = \boldsymbol{b}$
$\mathbb{N}$	the set of all natural numbers $1, 2, \ldots$
$\mathbb{Z}$	the set of all integer numbers
$\mathbb{Q}$	the set of all rational numbers
$\mathbb{R}$	the set of all real numbers
$\mathbb{C}$	the set of all complex numbers

Numbers

r^+	the positive part of a real number r, i.e, $r^+ = \max\{r, 0\}$
r^-	the negative part of a real number r, i.e, $r^- = \max\{-r, 0\}$
$\mathsf{Re}(z)$	the real part of a complex number z
$\mathsf{Im}(z)$	the imaginary part of a complex number z
$\overline{z}$	the complex conjugate of a complex number z (do not confuse with the upper bound of an interval)

Vectors and matrices

$v, \ldots$	column vectors
e_i	the ith standard unit vector (with convenient dimension), i.e., $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)^T$
e	the vector of ones (with convenient dimension), i.e., $e = (1, \ldots, 1)^T$
b_y	the vector $b_y \coloneqq b^c + \operatorname{diag}(y) b^\Delta$ p. 27
$\|v\|$	a vector norm, p. 11
$A, \ldots$	matrices
I_n	the identity matrix of size n
$\operatorname{diag}(v)$	the diagonal matrix with entries $v_1, \ldots, v_n$
A_{i*}	the ith row of a matrix A
A_{*j}	the jth column of a matrix A
A_{yz}	the matrix $A_{yz} = A^c - \operatorname{diag}(y) A^\Delta \operatorname{diag}(z)$, p 27
0	the zero vector or the zero matrix
$\|A\|$	a matrix norm, p. 11
A^T	the transpose of matrix A
A^{-1}	the inverse of matrix A
$\rho(A)$	the spectral radius of a matrix A, p. 12
$\lambda_i(A)$	the ith largest eigenvalue of a symmetric matrix A, p. 12
$A^\dagger$	the Moore–Penrose pseudoinverse of A
$\langle A \rangle$	the comparison matrix of A, i.e., $\langle A \rangle_{ii} = \|a_{ii}\|$ and $\langle A \rangle_{ij} = -\|a_{ij}\|$ for $i \neq j$, p. 14

The absolute value $|\cdot|$ and the relations $\leq \ldots$ of vectors and matrices are understood componentwise.

Intervals

$\boldsymbol{a}$	a real compact interval, p. 24
$\underline{a}$, $\inf(\boldsymbol{a})$	the lower bound of an interval $\boldsymbol{a}$, p. 24
$\overline{a}$, $\sup(\boldsymbol{a})$	the upper bound of an interval $\boldsymbol{a}$, p. 24
a^c, $\mathrm{mid}(\boldsymbol{a})$	the center of an interval $\boldsymbol{a}$, p. 24
a^Δ, $\mathrm{rad}(\boldsymbol{a})$	the radius of an interval $\boldsymbol{a}$, p. 24
$\mathrm{mag}(\boldsymbol{a})$	the magnitude of an interval $\boldsymbol{a}$, i.e., $\mathrm{mag}(\boldsymbol{a}) = \max\{\lvert a\rvert;\ a \in \boldsymbol{a}\}$, p. 25
$\mathrm{mig}(\boldsymbol{a})$	the mignitude of an interval $\boldsymbol{a}$, i.e., $\mathrm{mig}(\boldsymbol{a}) = \min\{\lvert a\rvert;\ a \in \boldsymbol{a}\}$, p. 25
$\boldsymbol{a} \leq \boldsymbol{b}$	$\overline{a} \leq \underline{b}$, p. 25
$\boldsymbol{a} < \boldsymbol{b}$	$\overline{a} < \underline{b}$, p. 25
$\mathrm{dist}(\boldsymbol{a}, \boldsymbol{b})$	distance of intervals, i.e., $\mathrm{dist}(\boldsymbol{a}, \boldsymbol{b}) = \max\{\lvert\underline{a} - \underline{b}\rvert, \lvert\overline{a} - \overline{b}\rvert\}$, p. 127
$\boldsymbol{A}$	an interval matrix $\{A;\ \underline{A} \leq A \leq \overline{A}\}$, p. 26
$\boldsymbol{A}^S$	a symmetric interval matrix $\{A \in \boldsymbol{A};\ A = A^T\}$, p. 27
$\mathbb{IR}$	the set of all real compact intervals, p. 25
$\mathbb{IR}^{m\times n}$	the set of all m-by-n interval matrices, p. 26
$\mathbb{IR}^n$	the set of all n-dimensional interval vectors, p. 25
$\square\mathcal{S}$	interval hull of a set $S \subset \mathbb{R}^n$, i.e., $\square\mathcal{S} \coloneqq \bigcap_{\boldsymbol{v}\in\mathbb{IR}^n:\mathcal{S}\subseteq\boldsymbol{v}} \boldsymbol{v}$, p. 25

Similar notation is used for interval vectors $\boldsymbol{v}, \ldots$ and interval matrices $\boldsymbol{A}, \ldots$

Functions

$\mathrm{sgn}(x)$	the sign of a real x, i.e., $\mathrm{sgn}(x) = 1$ if $x \geq 0$ and $\mathrm{sgn}(x) = -1$ otherwise
$\rho(A)$	the spectral radius of a matrix A
$f'(x)$	the derivative of $f\colon \mathbb{R} \to \mathbb{R}$
$f'_{x_i}(x)$, $\frac{\partial f}{\partial x_i}$	the partial derivative of $f\colon \mathbb{R}^n \to \mathbb{R}$ with respect to x_i
$\nabla f(x)$	the gradient of $f\colon \mathbb{R}^n \to \mathbb{R}$
$\nabla f(x)$	the Jacobian of $f\colon \mathbb{R}^n \to \mathbb{R}^n$, i.e., $(\nabla f(x))_{ij} = \frac{\partial f_i}{\partial x_j}$
$\nabla^2 f(\boldsymbol{x})$	the Hessian of $f\colon \mathbb{R}^n \to \mathbb{R}$, p. 14

Chapter 1

Preliminaries

Linear algebra

Most of the following can be found in Fiedler et al. (2006); Hogben (2007); Horn and Johnson (1985); Meyer (2000); Neumaier (1990).

Complex numbers. *A complex number* z can be viewed as an expression of the form $a + bi$, where $a, b \in \mathbb{R}$ and the imaginary number i satisfies $i^2 = -1$. Here, a is called *the real part* of z and is denoted by $\mathsf{Re}(z)$, and b is called *the imaginary part* of z and is denoted by $\mathsf{Im}(z)$. Arithmetic operations on complex numbers $z_1 = a + bi$ and $z_2 = c + di$ is then defined

$$\begin{aligned} z_1 \pm z_2 &= (a \pm c) + (b \pm d)i, \\ z_1 z_2 &= (ac - bd) + (cb + ad)i, \\ \frac{z_1}{z_2} &= \frac{ac + bd}{c^2 + d^2} + \frac{cb - ad}{c^2 + d^2} i, \quad z_2 \neq 0. \end{aligned}$$

Geometrically, a complex number $a + bi$ can be represented as a pair of numbers (a, b) in a plane. *The complex conjugate* of a complex number $z = a+bi$ is the complex number $\overline{z} = a-bi$. *The absolute value* of a complex number $z = a+bi$ is the distance of the point (a, b) to the origin, that is, $|z| = \sqrt{a^2 + b^2} = \sqrt{z\overline{z}}$. Some basic properties of complex numbers and operations:

- $z = \overline{z}$ if and only if z is real,
- $z + \overline{z} = 2\,\mathsf{Re}(z) \leq 2|z|$,
- $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$,
- $\overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2}$,
- $|z_1 + z_2| \leq |z_1| + |z_2|$,
- $|z_1 \cdot z_2| = |z_1| \cdot |z_2|$,
- $|z_1/z_2| = |z_1| \,/\, |z_2|$.

A complex number $z = a + bi$ can be also represented in the form $z = |z| \cdot e^{i\varphi}$, where φ is the angle of the the vector (a, b) with the positive real axis. This form is particularly convenient for multiplication since the product of complex numbers $z_1 = r_1 \cdot e^{i\varphi_1}$ and $z_2 = r_2 \cdot e^{i\varphi_2}$ is $z_1 z_2 = (r_1 r_2) \cdot e^{i(\varphi_1 + \varphi_2)}$.

Matrices and vectors. Matrices are denoted by capitals $A, B, \ldots$ For the entries of a matrix A we use A_{ij} or a_{ij}. The set of all $m \times n$ real matrices is denoted by $\mathbb{R}^{m \times n}$, and A^T stands for the transposed matrix of A.

Vectors are considered as column vectors

$$v = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = (v_1, v_2, \ldots, v_n)^T.$$

We use $\mathbb{R}^n$ for the set of all n-dimensional real vectors. It is useful to view vectors as one-column matrices. Thus, the inner product of two vectors $u, v \in \mathbb{R}^n$ can be expressed as

$$u^T v = \sum_{i=1}^{n} u_i v_i.$$

The ith row of a matrix $A \in \mathbb{R}^{m\times n}$ is denoted by

$$A_{i*} = (a_{i1}, a_{i2}, \ldots, a_{in}),$$

and the jth column by

$$A_{*j} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}.$$

Special types of vectors:

- the ith standard unit vector $e_i = (0, \ldots, 0, 1, 0, \ldots, 0)^T$,
- the vector of ones $e = (1, \ldots, 1)^T$.

Special types of matrices:

- $A \in \mathbb{R}^{m\times n}$ (not necessarily square) is *diagonal* if $a_{ij} = 0$ for $i \neq j$. The square diagonal matrix with entries $v_1, \ldots, v_n$ is denoted by $\mathrm{diag}(v)$.
- The identity matrix of size n is $I_n = \mathrm{diag}(e)$.
- $A \in \mathbb{R}^{n\times n}$ is *symmetric* if $A^T = A$.
- $A \in \mathbb{R}^{n\times n}$ is *orthogonal* if $A^T A = I_n$.

Inequalities $>, \geq, \ldots$ between vectors and between matrices are understood entrywise. For example, $A \geq 0$ means that all entries of A are nonnegative.

Determinants. THe determinant of a matrix $A \in \mathbb{R}^{n\times n}$ is denoted by $\det(A)$ and we suppose that the reader is familiar with the basic definitions and properties.

Let $A \in \mathbb{R}^{n\times n}$. Then

- a submatrix of A is obtained by removing arbitrary (possibly no) rows and columns from A,
- a principal submatrix of A is obtained by removing the same set of rows and columns from A
- a leading principal submatrix of A is obtained by removing the last k rows and columns from A $(0 \leq k < n)$.

Notice that there are $(2^n - 1)^2$ nonempty submatrices and $2^n - 1$ principal submatrices of A, but only n leading principal submatrices of A. Determinants of the various kinds of submatrices lead to the following notion:

- a minor is the determinant of a square submatrix of A,
- a principal minor is the determinant of a principal submatrix of A,
- a leading principal minor is the determinant of a leading principal submatrix of A.

The adjugate matrix of a matrix $A \in \mathbb{R}^{n\times n}$ is the matrix $\operatorname{adj}(A) \in \mathbb{R}^{n\times n}$ defined entrywise as

$$\operatorname{adj}(A)_{ij} = (-1)^{i+j} \det(A^{ji}), \quad i,j = 1,\dots,n,$$

where A^{ji} comes from A by removing its jth row and ith column. We always have

$$A \operatorname{adj}(A) = \det(A) I_n.$$

Provided A is nonsingular, it implies that $A^{-1} = \frac{1}{\det(A)} \operatorname{adj}(A)$.

Linear systems of equations and inverse matrices. Cramer's rule gives an explicit formula for the solution of a nonsingular system of linear equations. In the following, $e_i = (0,\dots,0,1,0,\dots,0)^T$ denotes the ith canonical vector, and the expression $A + (b - A_{*i})e_i^T$ says that the ith column in A is replaced by b.

Theorem 1.1 (Cramer's rule)**.** *Let $A \in \mathbb{R}^{n\times n}$ be nonsingular and $b \in \mathbb{R}^n$. Then the linear system of equations $Ax = b$ has a unique solution and its entries are*

$$x_i = \frac{\det(A + (b - A_{*i})e_i^T)}{\det(A)}, \quad i = 1,\dots,n.$$

Theorem 1.2 (Sherman–Morrison formula)**.** *Let $A \in \mathbb{R}^{n\times n}$ be nonsingular and $b, c \in \mathbb{R}^n$. If $c^T A^{-1} b = -1$, then $A + bc^T$ is singular, otherwise*

$$(A + bc^T)^{-1} = A^{-1} - \frac{1}{1 + c^T A^{-1} b} A^{-1} b c^T A^{-1}.$$

Definition 1.3. The Moore–Penrose pseudoinverse of $A \in \mathbb{R}^{m\times n}$ is a matrix $A^\dagger \in \mathbb{R}^{n\times m}$ satisfying

(1) $A = AA^\dagger A$,
(2) $A^\dagger = A^\dagger A A^\dagger$,
(3) $AA^\dagger$ is symmetric,
(4) $A^\dagger A$ is symmetric.

Theorem 1.4. *The Moore–Penrose pseudoinverse exists for every $A \in \mathbb{R}^{m\times n}$, is unique, and also satisfies, among others,*

(1) $A = A^{-1}$ when A is nonsingular,
(2) $A^\dagger = (A^T A)^{-1} A^T$ when A has linearly independent columns,
(3) $A^\dagger = A^T (AA^T)^{-1}$ when A has linearly independent rows.

Norms. *A vector norm* is a function $\|\cdot\| \colon \mathbb{R}^n \to \mathbb{R}$ satisfying for every $\alpha \in \mathbb{R}$ and $u, v \in \mathbb{R}^n$

1. $\|v\| \geq 0$ and $\|v\| = 0$ if and only if $v = 0$,
2. $\|\alpha v\| = |\alpha| \cdot \|v\|$,
3. $\|u + v\| \leq \|u\| + \|v\|$.

The well-known examples of norms are

- *the Euclidean norm:* $\|v\|_2 \coloneqq \sqrt{v^T v}$,
- *the maximum norm:* $\|v\|_\infty \coloneqq \max(|v_1|,\dots,|v_n|)$,
- *the Manhattan norm:* $\|v\|_1 \coloneqq \sum_{i=1}^n |v_i|$,
- *the general p-norm with $p \geq 1$:* $\|v\|_p \coloneqq (\sum_{i=1}^n |v_i|^p)^{1/p}$.

A matrix norm is a function $\|\cdot\| \colon \mathbb{R}^{m\times n} \to \mathbb{R}$ satisfying for every $\alpha \in \mathbb{R}$ and $A, B \in \mathbb{R}^{m\times n}$

1. $\|A\| \geq 0$ and $\|A\| = 0$ if and only if $A = 0$,

2. $\|\alpha A\| = |\alpha| \cdot \|A\|$,
3. $\|A + B\| \leq \|A\| + \|B\|$,
4. $\|AB\| \leq \|A\| \cdot \|B\|$.

Thus, a matrix norm is a vector norm on the linear space $\mathbb{R}^{m\times n}$ with the additional consistency property 4. An example of a matrix norm is

- *the Frobenius norm:* $\|A\|_F \coloneqq \sum_{i=1}^m \sum_{j=1}^n a_{ij}^2$.

Another important matrix norms are those induced by a vector norm. Let $\|\cdot\|$ be a vector norm on $\mathbb{R}^n$, then the *induced matrix norm* is defined

$$\|A\| \coloneqq \max_{x:\|x\|=1} \|Ax\|.$$

By considering the Euclidean, maximum and Manhattan vector norms, we obtain the induced matrix norms

- $\|A\|_2 = \sigma_1(A)$, the maximal singular value,
- $\|A\|_\infty = \max_{i=1,\dots,m} \sum_{j=1}^n |a_{ij}|$,
- $\|A\|_1 = \|A^T\|_\infty = \max_{j=1,\dots,n} \sum_{i=1}^m |a_{ij}|$.

Subordinate matrix norms are consistent with the norms that induce them, giving rise to the useful inequality

$$\|Av\| \leq \|A\| \cdot \|v\|.$$

Eigenvalues. Let $A \in \mathbb{R}^{n\times n}$. The eigenvalue of A is any $\lambda \in \mathbb{C}$ such that $Ax = \lambda x$ for some nonzero $x \in \mathbb{R}^n$. The vector x is called the eigenvector corresponding to λ. The spectral radius of A, denoted as $\rho(A)$, is the largest absolute value of all (even complex) eigenvalues of A. For each matrix norm we have $\rho(A) \leq \|A\|$.

Theorem 1.5. *Let $A \in \mathbb{R}^{n\times n}$, then $\rho(A) \leq \|A\|$ for any matrix norm.*

Theorem 1.6. *Let $A \in \mathbb{R}^{n\times n}$ and $\varepsilon > 0$. Then there exists a matrix norm such that $\rho(A) \leq \|A\| \leq \rho(A) + \varepsilon$.*

Theorem 1.7 (Perron–Frobenius)**.** *Let $A \in \mathbb{R}^{n\times n}$ be nonnegative. Then $\rho(A)$ is an eigenvalue of A and the corresponding eigenvector is nonnegative.*

Theorem 1.8. *Let $A, B \in \mathbb{R}^{n\times n}$. If $|A| \leq B$, then $\rho(A) \leq \rho(|A|) \leq \rho(B)$.*

Theorem 1.9. *Let $A \in \mathbb{R}^{n\times n}$ be nonnegative, $x \in \mathbb{R}^n$ positive and $\mu \in \mathbb{R}$. If $Ax < \mu x$, then $\rho(A) < \mu$.*

Theorem 1.10 (Neumann Series)**.** *Let $A \in \mathbb{R}^{n\times n}$, then the following are equivalent*

(1) $\rho(A) < 1$,
(2) $\lim_{k\to\infty} A^k = 0$,
(3) $I + A + A^2 + \dots$ *converges.*

In which case, $(I_n - A)^{-1} = \sum_{k=0}^\infty A^k = I + A + A^2 + \dots$

Eigenvalues of symmetric matrices. If A is symmetric, then it has only real eigenvalues, so we may suppose they are sorted non-increasingly

$$\lambda_1(A) \geq \dots \geq \lambda_n(A).$$

Theorem 1.11 (Rayleigh–Ritz)**.** *For any symmetric matrix $A \in \mathbb{R}^{n\times n}$ we have*

$$\lambda_1(A) = \max_{x:\|x\|_2=1} x^T Ax, \quad \lambda_n(A) = \min_{x:\|x\|_2=1} x^T Ax.$$

Theorem 1.12 (Weyl). *For any symmetric matrices $B, C \in \mathbb{R}^{n\times n}$ we have*

$$\lambda_i(B) + \lambda_n(C) \le \lambda_i(B+C) \le \lambda_i(B) + \lambda_1(C), \quad i = 1, \dots, n.$$

Theorem 1.13 (Bendixson). *For any eigenvalue $\lambda + i\mu \in \mathbb{C}$ of matrix $A \in \mathbb{R}^{n\times n}$ we have*

$$\lambda_n\big(\tfrac{1}{2}(A+A^T)\big) \le \lambda \le \lambda_1\big(\tfrac{1}{2}(A+A^T)\big),$$
$$\lambda_n\big(\tfrac{1}{2i}(A-A^T)\big) \le \mu \le \lambda_1\big(\tfrac{1}{2i}(A-A^T)\big).$$

Proof. See Marcus and Minc (1964); Stoer and Bulirsch (2002). □

SVD decomposition and singular values. Each matrix $A \in \mathbb{R}^{m\times n}$ can be decomposed as $A = U\Sigma V^T$, where $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$ are orthogonal and $\Sigma \in \mathbb{R}^{m\times n}$ is diagonal with nonnegative entries. This decomposition is called *SVD decomposition* (singular value decomposition) and the diagonal entries of Σ are called *singular values*. They can be sorted as $\sigma_1 \ge \cdots \ge \sigma_n \ge 0$, where the number of positive singular values is equal to the rank of A. The largest singular value σ_1 provides the value of the matrix 2-norm (spectral norm) and the smallest singular value σ_n provides the distance of A to the nearest singular matrix in 2-norm (provided A is square).

Theorem 1.14 (Relation of eigenvalues and singular values). *Let $A \in \mathbb{R}^{m\times n}$ with singular values $\sigma_1, \dots, \sigma_n$. Then*

(1) matrix A^TA has eigenvalues $\sigma_1^2, \dots, \sigma_n^2$;

(2) matrix $\left(\begin{smallmatrix} 0 & A \\ A^T & 0 \end{smallmatrix}\right)$ has eigenvalues $\pm\sigma_1, \dots, \pm\sigma_n$ and the others are zero.

Positive (semi)definiteness. Let $A \in \mathbb{R}^{n\times n}$ be symmetric. The matrix A is called *positive semidefinite* if $x^TAx \ge 0$ for each $x \in \mathbb{R}^n$. The matrix A is called *positive definite* if $x^TAx > 0$ for each nonzero $x \in \mathbb{R}^n$. Equivalent characterizations follow.

Theorem 1.15. *Let $A \in \mathbb{R}^{n\times n}$ be symmetric. The following are equivalent:*

(1) A is positive semidefinite,

(2) A has nonnegative eigenvalues,

(3) A can be expressed as $A = BB^T$ for a certain matrix B,

(4) all principal minors of A are nonnegative.

Theorem 1.16. *Let $A \in \mathbb{R}^{n\times n}$ be symmetric. The following are equivalent:*

(1) A is positive definite,

(2) A has positive eigenvalues,

(3) A can be expressed as $A = BB^T$ for a certain nonsingular matrix B,

(4) all leading principal minors of A are positive.

Moreover, a symmetric matrix $A \in \mathbb{R}^{n\times n}$ is positive definite if and only if it can be factorized as $A = LL^T$, where $L \in \mathbb{R}^{n\times n}$ is lower triangular with positive diagonal entries. This is called *Cholesky decomposition* and L is uniquely determined.

Special matrices.

Definition 1.17. A matrix $A \in \mathbb{R}^{n\times n}$ is called an *M-matrix*, if $a_{ij} \le 0$ for $i \ne j$ and $A^{-1} \ge 0$.

Theorem 1.18. *Let $A \in \mathbb{R}^{n\times n}$ be such that $a_{ij} \le 0$ for $i \ne j$. Then the following are equivalent:*

(1) A is an M-matrix,

(2) there is $v > 0$ such that $Av > 0$,

(3) $A^{-1}e > 0$,

(4) there is a splitting $A = M - N$, where $N \ge 0$, $M^{-1} \ge 0$ and $\rho(M^{-1}N) < 1$,

(5) each splitting $A = M - N$, where $N \geq 0$, $M^{-1} \geq 0$ satisfies $\rho(M^{-1}N) < 1$,

(6) all principal leading minors of A are positive,

(7) the real part of all eigenvalues of A is positive,

(8) the real eigenvalues of A are positive.

Theorem 1.19. *Let $A \in \mathbb{R}^{n\times n}$ be nonnegative. Now, $\rho(A) < 1$ if and only if $(I_n - A)^{-1} \geq 0$.*

Proof. "⇒" By Theorem 1.10, $(I_n - A)^{-1} = \sum_{k=0}^{\infty} A^k \geq 0$.
"⇐" By Definition 1.17, $I_n - A$ is an M-matrix, and by Theorem 1.18(5), $\rho(A) < 1$. □

Definition 1.20. A matrix $A \in \mathbb{R}^{n\times n}$ is called an *H-matrix*, if the so-called *comparison matrix* $\langle A\rangle$ is an M-matrix, where $\langle A\rangle_{ii} = |a_{ii}|$ and $\langle A\rangle_{ij} = -|a_{ij}|$ for $i \neq j$.

Obviously, each M-matrix is an H-matrix. Further, each H-matrix is nonsingular.

Definition 1.21. A matrix $A \in \mathbb{R}^{n\times n}$ is called a *P-matrix*, if every principal minor is positive.

There are $2^n - 1$ principal minors of A, which is hard to overcome and it makes the problem of checking P-property to be computationally expensive (Coxson, 1994). Nevertheless, as long as A is symmetric, P-property is equivalent to positive definiteness, and thus easily checked. P-matrices arises in the theory of linear complementarity problems providing existence and uniqueness of solutions (Cottle et al., 2009; Murty and Yu, 1997). Efficiently recognizable subclasses of P-matrices are positive definite matrices, M-matrices, H-matrices with positive diagonal entries and totally positive matrices.

Calculus

Let $f\colon \mathbb{R}^n \to \mathbb{R}$. By $f'_{x_i}(x)$ we denote the partial derivative of f with respect to x_i, and by $\nabla f(x) \coloneqq (f'_{x_1}(x), \ldots, f'_{x_n}(x))^T$ the gradient of $f\colon \mathbb{R}^n \to \mathbb{R}$. The Jacobian of a vector valued function $f\colon \mathbb{R}^n \to \mathbb{R}^n$ is denoted by $\nabla f(x)$, too; its ith row is formed by the gradient of the ith function $f_i(x)$. The Hessian of f is denoted by $\nabla^2 f(x)$, so $(\nabla^2 f(x))_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}(x)$.

The following results are found, e.g., in Petersen and Pedersen (2012); Rudin (1987).

Theorem 1.22 (Mean value theorem for single-valued functions)**.** *Let $f\colon \mathbb{R} \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b). Now, there is $c \in (a, b)$ such that $f(b) - f(a) = f'(c)(b - a)$.*

Theorem 1.23 (Mean value theorem for vector-valued functions)**.** *Let $\mathcal{U} \subseteq \mathbb{R}^n$ be an open set and $f\colon \mathbb{R}^n \to \mathbb{R}$ differentiable on $\mathcal{U}$. For any $a, b \in \mathcal{U}$ there is c on the line segment connecting them such that $f(b) - f(a) = \nabla f(c)^T(b - a)$.*

Theorem 1.24 (Taylor's theorem)**.** *Let $f\colon \mathbb{R} \to \mathbb{R}$ be $m+1$ times differentiable on (a, x) and continuous on $[a, x]$. For each $x \in \mathbb{R}^n$ there is $c \in (a, x)$ such that*

$$f(x) = f(a) + f'(a)(x - a) + \frac{f''(a)}{2!}(x - a)^2 + \cdots + \frac{f^{(m)}(a)}{m!}(x - a)^m + \frac{f^{(m+1)}(c)}{(m + 1)!}(x - a)^{m+1}.$$

Theorem 1.25. *Let $A \in \mathbb{R}^{n\times n}$ nonsingular, $b \in \mathbb{R}^n$ and x^* the solution of $Ax = b$. Then*

(1) $\dfrac{\partial (A^{-1})_{k\ell}}{\partial a_{ij}} = -(A^{-1})_{ki}(A^{-1})_{j\ell}$,

(2) $\dfrac{\partial x_k}{\partial a_{ij}} = -(A^{-1})_{ki}x^*_j, \quad \dfrac{\partial x_k}{\partial b_i} = (A^{-1})_{ki}$.

Convexity. A set $\mathcal{M} \subseteq \mathbb{R}^n$ is *convex* if for each $x_1, x_2 \in \mathcal{M}$ and for each $\lambda_1, \lambda_2 \geq 0$, $\lambda_1 + \lambda_2 = 1$, we have $\lambda_1 x_1 + \lambda_2 x_2 \in \mathcal{M}$. *A convex hull* of a set $\mathcal{M} \subseteq \mathbb{R}^n$ is the smallest convex set containing $\mathcal{M}$, or more formally,

$$\operatorname{conv} \mathcal{M} = \cap\{\mathcal{S};\ \mathcal{M} \subseteq \mathcal{S} \text{ and } \mathcal{S} \text{ is convex}\}.$$

Let $\mathcal{M} \subseteq \mathbb{R}^n$ be convex. Then a function $f\colon \mathcal{M} \to \mathbb{R}$ is called *convex* on $\mathcal{M}$ if for each $x_1, x_2 \in \mathcal{M}$ and each $\lambda_1, \lambda_2 \geq 0$, $\lambda_1 + \lambda_2 = 1$, we have

$$f(\lambda_1 x_1 + \lambda_2 x_2) \leq \lambda_1 f(x_1) + \lambda_2 f(x_2).$$

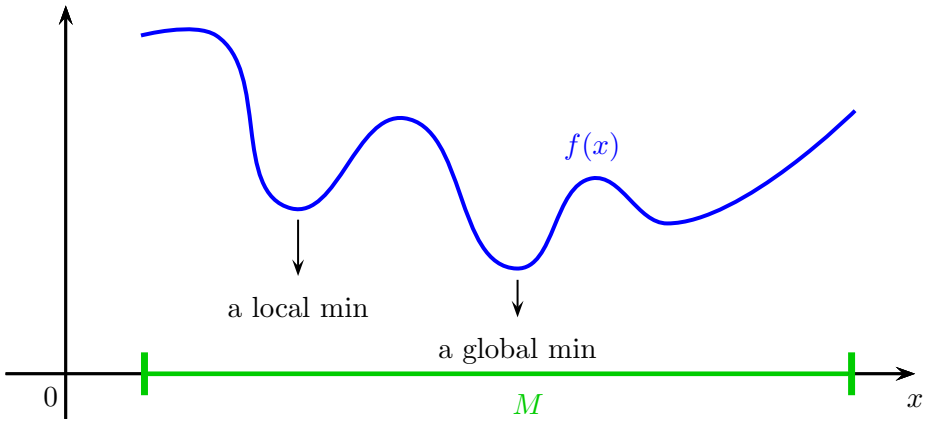


Figure 1.1: Local and global minima.

Topology

Theorem 1.26 (Brouwer's fixed-point theorem)**.** *Let $\mathcal{U} \subseteq \mathbb{R}^n$ be a convex compact set and $f\colon \mathcal{U} \to \mathcal{U}$ continuous. There is $x \in \mathcal{U}$ such that $f(x) = x$.*

Theorem 1.27. *Let $v \in \mathbb{R}^n$ and $U, V \in \mathbb{R}^{n\times n}$ such that $\rho(UV) < 1$. Then the iterations*

$$x \mapsto O(x) \coloneqq U(v + Vx)$$

converge to a unique fixed point for any initial vector.

Proof. For any $x, y \in \mathbb{R}^n$, any sufficiently small $\varepsilon > 0$ and any matrix norm from Theorem 1.6, we have

$$\|O(x) - O(y)\| = \|UV(x - y)\| \le \|UV\| \cdot \|x - y\| \le (\rho(UV) + \varepsilon)\|x - y\| < \|x - y\|.$$

That is, the mapping $O(x)$ is a contraction, and for any initial vector, the iterations converge to the fixed point $x = (I_n - UV)^{-1}Uv$. □

Optimization

An optimization problem reads (Bazaraa et al., 2006; Nocedal and Wright, 2006)

$$\min\ f(x)\ \text{ subject to }\ x \in M,$$

where $f\colon \mathbb{R}^n \to \mathbb{R}$ is *the objective function* and $M \subseteq \mathbb{R}^n$ is *the feasible set*. We consider just minimization problems since a maximization problem is transformed to a minimization problem by $\max_{x\in M} f(x) = -\min_{x\in M} -f(x)$.

The goal is to find *a global minimum*, which is a feasible point $x^* \in M$ such that $f(x^*) \le f(x)$ for every $x \in M$. There can be multiple minima, for example, the sine function has infinitely many minima on the real line. It may also happen that there is no minimum, that is, the minimal value of objective function $f(x)$ on M is not attained. Consider, for example, the problem $\min_{x\in\mathbb{R}} x$, which is unbounded from below, or the problem $\min_{x\in\mathbb{R}} e^x$, which is bounded from below.

Global minima are hard to compute and many solver get stuck in local minima. A point $x^* \in M$ is *a local minimum* if there is $\varepsilon > 0$ such that $f(x^*) \le f(x)$ for every $x \in M \cap N_\varepsilon(x^*)$, where $N_\varepsilon(x^*) = \{x;\ \|x - x^*\| \le \varepsilon\}$ is a neigborhood of x^*.

Linear programming

Linear programming is the class of optimization problems, where the feasible set is described by linear constraints and the objective function is linear as well. There are many books on linear programming, including Padberg (1999); Schrijver (1998).

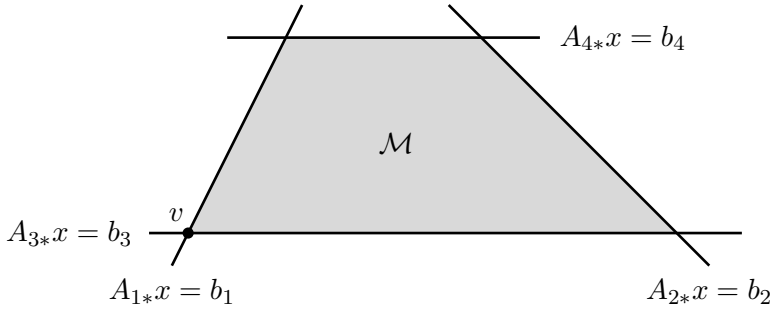


Figure 1.2: A convex polyhedron $\mathcal{M}$ as the intersection of four half-spaces.

Convex polyhedra. Let $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. The set $\mathcal{M}$ described by the system of linear inequalities $Ax \le b$ is called a *convex polyhedron*. Geometrically, it is an intersection of half-spaces; one inequality determines a half-space, and a system of inequalities corresponds to the intersection of the particular half-spaces. Concretely, if $A_{i*} \neq 0$, then the ith inequality $A_{i*}x \le b_i$ describes a half-space in $\mathbb{R}^n$ and A_{i*} is the normal vector, which is perpendicular to the half-space. Thus a convex polyhedron constitutes an intersection of half-spaces; see Figure 1.2 provides an illustration.

A *face* of a convex polyhedron $\mathcal{M}$ described by $Ax \le b$ is a nonempty set described by the system $Ax \le b$, in which some of the inequalities are replaced by equations. One-dimensional face is an *edge* and 0-dimensional face is a *vertex*. The *active set* of a feasible point $x^* \in \mathcal{M}$ is the index set of inequalities that are satisfied as equations for x^*. For example, the active set of the vertex v is $\{1, 3\}$ in Figure 1.2. The active set of an interior point of $\mathcal{M}$ is empty.

The Minkowski–Weyl theorem claims that any bounded convex polyhedron $\mathcal{M}$ can be equivalently represented as the convex hull of its vertices. Thus we have two alternative representations of a convex polyhedron – by a system of linear inequalities or by a list of vertices.

LP problem in equality form. Linear programming problems seek for a point of a convex polyhedral set that attains the minimal (or maximal) value of a linear function. An LP problem in equality form reads

$$\min c^T x \;\text{ subject to }\; Ax = b,\; x \ge 0, \tag{P}$$

where $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. We assume without loss of generality that $m \le n$.

A vector $x \in \mathbb{R}^n$ is called *a feasible solution* if it satisfies the constraints, that is, if $Ax = b$, $x \ge 0$. If there exists at least one feasible, then the LP problem is *feasible*; otherwise it is *infeasible*. A feasible solution $x^* \in \mathbb{R}^n$ is *optimal* if $c^T x^* \le c^T x$ for each feasible solution x. If x^* is an optimal solution, then $c^T x^*$ is *the optimal value*.

For an index set $B \subseteq \{1, \dots, n\}$, we denote by A_B the restriction of A to the columns indexed by B. We say that B is *a basis* if it has cardinality m and A_B is nonsingular. A basis B is *feasible* if $A_B^{-1}b \ge 0$.

Geometrically, the feasible set $Ax = b$, $x \ge 0$ describes a convex polyhedral set in $\mathbb{R}^n$. For each feasible basis B, there is associated *a basic solution* x with entries $x_B = A_B^{-1}b$ and $x_N = 0$, where $N := \{1, \dots, n\} \setminus B$. The corresponding objective value is $c^T x = c_B^T x_B = c_B^T A_B^{-1} b$. Each basic solution geometrically forms a vertex of the feasible set, but it is not unique – one vertex can sometimes be represented by basic solutions with mutually different bases. A basis is called *optimal* if the associated solution is optimal. The well-known simplex method for s olving LP problems is based on moving from one feasible basis to another until an optimal one is reached; one step corresponds to a change of the current basis to a neighboring one having one basic index changed.

Theorem 1.28. *For a given LP problem, exactly one of the following three possibilities occurs*

(1) the problem is infeasible (there is no feasible solution),

(2) the problem is optimal (there is an optimal solution),

(3) the problem is unbounded (for each $\gamma \in \mathbb{R}$, there is a feasible x such that $c^T x \le \gamma$).

By convention, the optimal value of the linear program is ∞ in the first case. In the second case, it is a real value, and in the third case, it is $-\infty$. We use this convention generically: the minimum value of an infeasible problem is ∞ and an unbounded problem has the optimal value $-\infty$. Analogously for a maximization problem.

Theorem 1.29 (Optimality criterion)**.** *Let B be a feasible basis. Then B is an optimal basis if and only if $c_N^T - c_B^T A_B^{-1} A_N \geq 0^T$.*

Moreover, if $c_N^T - c_B^T A_B^{-1} A_N > 0^T$, then the corresponding optimal solution is unique.

Duality. From now on, we will call (P) to be the primal problem. The dual problem associated with it is the LP problem

$$\max b^T y \;\text{ subject to }\; A^T y \leq c. \tag{D}$$

Theorem 1.30 (Weak duality)**.** *Let x and y be feasible solutions of (P) and (D), respectively. Then $c^T x \geq b^T y$.*

Moreover, if $c^T x = b^T y$, then x and y are optimal.

As long as $c^T x \leq b^T y$ for some primal and dual feasible x and y, then they are optimal as well because $c^T x < b^T y$ cannot happen.

Corollary 1.31. *We have:*

(1) If (P) is unbounded, then (D) is infeasible.

(2) If (D) is unbounded, then (P) is infeasible.

(3) Let (P) be feasible. Then (P) is unbounded if and only if (D) is infeasible.

(4) Let (D) be feasible. Then (D) is unbounded if and only if (P) is infeasible.

Notice that it may happen that both (P) and (D) are infeasible. Consider, for instance, the one-dimensional LP problem

$$\min \; -x \;\text{ subject to }\; 0x = 1,\; x \geq 0.$$

Theorem 1.32 (Strong duality)**.** *The following are equivalent:*

(1) (P) has an optimal solution,

(2) (D) has an optimal solution,

(3) both (P) and (D) are feasible.

If any of the above holds, then (P) and (D) have the same optimal value.

Other forms of LP problems. The feasible set can be described not only by the system $Ax = b$, $x \geq 0$, but also by systems of other type. Two other canonical forms that are used are:

$$\min c^T x \;\text{ subject to }\; Ax \leq b,$$
$$\min c^T x \;\text{ subject to }\; Ax \leq b,\; x \geq 0.$$

All these forms are mutually transformable to each other. That is why usually the LP theory is derived for one of the canonical forms.

The duals of the above LP forms are, respectively,

$$\max b^T x \;\text{ subject to }\; A^T y = c,\; y \leq 0,$$
$$\max b^T x \;\text{ subject to }\; A^T y \leq c,\; y \leq 0.$$

Farkas' lemma. It is a very useful technique that transforms a solvable linear system into an unsolvable one and vice versa. We state its versions for linear systems of various types. They are mutually transformable to each other; the only exception is the second one, which deals with a simple system of linear equations. Notice also that the counterpart systems can be expressed in a slightly different form. For instance, $A^T u \geq 0$, $b^T u \leq -1$ is equivalent to the system $A^T u \geq 0$, $b^T u < 0$ with respect to solvability.

Theorem 1.33 (Farkas' lemma)**.** *Let $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. Then we have*

(1) $Ax = b$, $x \geq 0$ is solvable iff $A^T u \geq 0$, $b^T u \leq -1$ is not solvable,

(2) $Ax = b$ is solvable iff $A^T u = 0$, $b^T u \leq -1$ is not solvable,

(3) $Ax \leq b$, $x \geq 0$ is solvable iff $A^T u \geq 0$, $b^T u \leq -1$, $u \geq 0$ is not solvable,

(4) $Ax \leq b$ is solvable iff $A^T u = 0$, $b^T u \leq -1$, $u \geq 0$ is not solvable.

For our purposes, it is convenient to formulate Farkas' lemma for a general system of mix linear equations and inequalities.

Corollary 1.34. *Exactly one of the linear systems*

$$Ax + By = b,\ Cx + Dy \leq d,\ x \geq 0 \tag{1.1}$$

and

$$A^T u + C^T v \geq 0,\ B^T u + D^T v = 0,\ b^T u + d^T v \leq -1,\ v \geq 0, \tag{1.2}$$

is solvable.

Proof. Rewriting (1.1) into the inequality form

$$Ax + By \leq b,\ -Ax - By \leq -b,\ Cx + Dy \leq d,\ -x \leq 0,$$

and calling Theorem 1.33(4), we obtain the alternative system

$$\begin{aligned} A^T u^1 - A^T u^2 + C^T v - w &= 0,\\ B^T u^1 - B^T u^2 + D^T v &= 0,\\ b^T u^1 - b^T u^2 + d^T v \leq -1,\ u^1, u^2, v, w &\geq 0. \end{aligned}$$

Substituting $u := u^1 - u^2$ and eliminating w, we get (1.2). □

Degeneracy. In the LP problem (P), a basic solution x corresponding to the basis B is (primal) *degenerate* if some of its basic entries vanish, that is, there is one or more zeros in the vector $A_B^{-1} b$. The basic solution is *dual degenerate* if there is at least one zero in the vector $c_B^T - c_B^T A_B^{-1} A_N$.

Primal degeneracy is usually caused by the situation when there is a solution corresponding to several bases; for an illustration, consider a peak of a pyramid.

Dual degeneracy means that there is an edge of the feasible polyhedron that is perpendicular to the objective vector c, which implies that moving along this edge does not change the objective value. Therefore, if an optimal basic solution x^* is dual nondegenerate, then the optimal solution is unique.

Algorithms. One of the first practically usable algorithms for solving linear programming problems was the simplex method by Dantzig (1963). However, the method may take exponential number of steps. The provably polynomial time algorithms are the ellipsoid method by Khachiyan (1979), which is mostly of theoretical importance, and various types of interior point methods; see Padberg (1999); Roos et al. (2006); Schrijver (1998).

Computational complexity

Running time is an important aspect of each algorithm. We often measure it by the number of instructions (assignments, arithmetic operations, etc.), but from the theoretical viewpoint, we will measure it by the number of operations on an abstract Turing machine. This is a substantially different model since, for instance, the Gaussian elimination of a matrix of size n takes an order of n^3 arithmetic operations, but the number of bit operations depends on the concrete entries. It is still polynomial with respect to the input size, but the proof is not straightforward as the values might increase exponentially during the elimination.

An algorithm is called *polynomial* if it takes a polynomial number of operations with respect to the input size. A problem is *polynomial* if there is a polynomial algorithm for it.

A decision problem, answering "yes" or "no", belongs to the class P if it is a polynomial problem. It belongs to the class NP if there is a nondeterministic polynomial algorithm for it. That is, informally speaking, whenever we have a choice of several branches, we choose the right one. Precise definition and more on computational complexity is given, e.g., in Papadimitriou (1994). The class *co-NP* consists of problems whose negative is in NP. The hardest problems in the class NP are called NP-complete, and similarly for *co-NP*. Formally, a problem $\mathcal{P}$ is NP-complete if it is in NP and it is NP-hard, where NP-hardness means that every problem $\mathcal{P}'$ in NP can be polynomially reduced to this one (i.e., we can answer $\mathcal{P}'$ be solving a specific instance or instances of $\mathcal{P}$ and the transformation takes a polynomial number of operations). An example of a polynomial problem is checking solvability of linear equations and inequalities over rationals (Schrijver, 1998), an example of an NP-complete problem is checking satisfiability of a Boolean formula, and determining whether a given Boolean formula is a tautology is an *co-NP*-complete problem.

A problem in NP can be solved by a deterministic algorithm, but the computational time may increase exponentially. The question whether it can be done polynomially, that is, whether $P = NP$, is one of the most challenging open problems in computer science. All known algorithms for NP-hard problems have exponential complexity in the worst case.

When we talk about NP-completeness of non-decision problems, we mean their decision formulation analogy. For instance, the optimization problems $\min_{x\in M} f(x)$ are formulated as decision problems $\exists x \in M : f(x) \leq c$ for a parameter $c \in \mathbb{Q}$. Thus, linear programming is polynomial, while Travelling salesman problem is NP-complete.

An ε-approximation algorithm of an optimization problem $\min_{x\in M} f(x)$ finds $x^* \in M$ such that $|f(x^*) - f^*| \leq \varepsilon \cdot \max\{|f(x^*)|, |f^*|\}$, where $f^* \coloneqq \min_{x\in M} f(x)$ is the optimal value and $\varepsilon > 0$ is a given parameter.

Not only there exist problems that cannot be solved effectively, there are problems that cannot be solved at all! Such problems are called *undecidable*, and it can be mathematically proved that there is no algorithm (running on a Turing machine) to solve them. The prototype example is the halting problem, which should decide whether a given algorithm finishes computation on given input data. Another example is Hilbert's tenth problem of deciding whether a multivariable polynomial equation has a solution in integers, or the mortal matrix problem of determining whether a given finite collection of matrices can be multiplied (with repetition) in some order to yield the zero matrix.

Working example of an *NP*-complete problem. The fundamental NP-complete problem that we will employ in the book is the following by (Rohn, 2006a, Thm. 2.3), and also a variant by Hladík (2012b).

Theorem 1.35 (Rohn, 2006a)**.** *Checking solvability of*

$$-e \leq Ax \leq e,\ e^T|x| \geq 1 \tag{1.3}$$

is NP-complete in the set of non-negative positive definite rational matrices.

Theorem 1.36. *Checking solvability of*

$$-e \leq Ax \leq e,\ e^T|x| > 1 \tag{1.4}$$

is NP-complete in the set of non-negative positive definite rational matrices.

***P*-completeness.** In some sense, P-complete problems are the hardest problems in class P. Formally, a problem is P-complete if it lies in class P and and every problem in P can be reduced to it. When we consider NC-reduction, we require that it works in polylogarithmic time on a parallel computer with a polynomial number of processors. Even though P-complete problems are efficiently solvable, they are hard to parallelize, to solve in a limited space, and one also cannot hope for having a simple closed form expression for solving the problem.

An example of a P-complete problem is a linear program and its decision formulations, for example, checking solvability of a linear system $Ax = b$, $x \geq 0$. It is a challenging open problem in computational complexity theory whether P-complete problems form a strict subset of P.

Chapter 2

Introduction

Interval analysis, roughly speaking, is concerned with interval-valued problems. We first motivate the reader and show that interval data naturally appear in many different fields. Then, we remind some well-known results and introduce interval and non-interval notation used throughout the book. As a first step, we define the basic operations on intervals (interval arithmetic) and discuss some properties. Evaluation of general functions on intervals is considered as one of the fundamental problems – we give only an introductory insight, and study it more thoroughly in chapter 6.

2.1 Why intervals?

There are many scientific areas, where intervals naturally appear. They can be categorized into several classes according to their main purpose

- to deal with rounding errors
- to process a set of states
- to model uncertainty

2.1.1 Numerical issues – roundings in floating point arithmetic

Roundings can make problems – even small value changes in numerical calculations may have a large effect on the resulting value, as the following example illustrates.

Example 2.1. Consider the famous example by Rump (1994, 2010)

$$f = 333.75b^6 + a^2(11a^2b^2 - b^6 - 121b^4 - 2) + 5.5b^8 + \frac{a}{2b},$$

with $a = 77617$ and $b = 33096$. The true value is $f = \frac{a}{2b} - 2 = -0.827386\ldots$ In single, double and extended precision, the calculations from 80s gave

$$\begin{aligned} \text{single precision} \quad & f \approx 1.172603\ldots \\ \text{double precision} \quad & f \approx 1.1726039400531\ldots \\ \text{extended precision} \quad & f \approx 1.172603940053178\ldots \end{aligned}$$

The same behavior is observable for the expression in the form

$$f = 21 \cdot b \cdot b - 2 \cdot a \cdot a + 55 \cdot b \cdot b \cdot b \cdot b - 10 \cdot a \cdot a \cdot b \cdot b + \frac{a}{2b},$$

As another Rump's example Hyvönen and De Pascale (1996), let

$$a \coloneqq 10864, \quad b \coloneqq 18817, \quad c \coloneqq 9a^4 - b^4 + 2b^2.$$

The exact value of c is 1, but evaluation in floating point arithmetic usually yields $c = 2$. □

In some cases, instable expressions can be stabilized, e.g., $\sqrt{x+1} - \sqrt{x}$ is not stable for large x, but the equivalent expression $\sqrt{x+1} - \sqrt{x} = \frac{1}{\sqrt{x+1}+\sqrt{x}}$ is stable. Nevertheless, such a rearrangement is not always possible.

Representation of non-representable values

Another problem with floating point arithmetic is that many values cannot be represented exactly. One way out of it is to enclose these values into intervals the endpoints of which are representable, e.g.,

$$\sqrt{2} \in [1.4142135623730950488,\ 1.4142135623730950489],$$
$$\pi \in [3.1415926535897932384,\ 3.1415926535897932385].$$

This represents the values indirectly, but with mathematical and numerical rigor – the values lie somewhere in the respective intervals.

Disasters

For maybe a more convincing and impressive real world problems caused by round-off errors see

`http://www.math.psu.edu/dna/disasters`

Some of such disasters are:

- *The Patriot Missile failure*

  This happened on February 25, 1991, during the Gulf War. The American Patriot Missile battery in Dharan, Saudi Arabia, inaccurately calculated the time – one hundred hours of being up, a small rounding error of the binary representation of $\frac{1}{10}$ expanded to 0.34 seconds. As a consequence, the battery failed to intercept an incoming Iraqi Scud missile, which killed 28 soldiers.

[source: Wikimedia Commons, File:Patriot_missile_launch_b.jpg]

- *The sinking of the Sleipner A offshore platform*

  The Sleipner A platform producing oil and gas in the North Sea in Gandsfjorden, Norway, sprang a leak and needed to be sunk under a controlled operation on 23 August, 1991. The loss was caused by a failure in a cell wall, which was a result of an inaccurate finite element approximation of the linear elastic model – the shear stresses were underestimated by 47%.

- *The explosion of the Ariane 5 rocket*

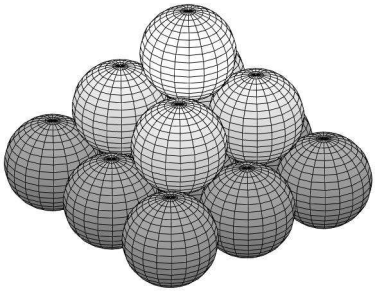
  On 4 June, 1996, the flight of the Ariane 5 launcher exploded just 40 seconds after its lift-off from Kourou, French Guiana. The damage was evaluated at $500 million. The failure was caused by an overflow when converting a 64-bit floating point number to a 16-bit signed integer value.

Computer-assisted proofs

Some mathematical proofs were carried out with the aid of computers. Employing interval arithmetic here is necessary to obtain verified results. Thus, if the software and hardware work according to the specifications, we can be sure that the results are mathematically correct. Interval analysis was utilized in proving the famous theorems and long-standing conjectures (Frommer, 2001; Neumaier, 2007) such as those below. Due to its discrete mathematical nature, the proof of the famous four color theorem avoids interval computation.

- **Kepler conjecture**

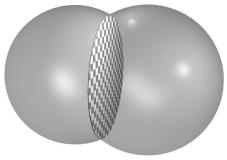
  A long-standing conjecture addressing what is the the densest packing of equally sized balls. It dates back to Johannes Kepler (1611), and was also stated as Hilbert's 18th problem. The conjecture claimed that the face-centered cubic packing, which is depicted in the figure (it is the way how greengrocers stack oranges), is the optimal one. Another optimal packing is the so-called hexagonal close packing, both having the density of $\pi/(3\sqrt{2}) \approx 74.048\%$.

  The conjecture was proved by Hales (2005). In a part of the original proof, one needs to classify about 5000 graphs by a scoring function, and the scores must not exceed a value of 8. Interval computations numerically verified this upper bound, which was indeed approached by a score ≈ 7.9997.

- **Double bubble**

  The double bubble problem asks what is the minimal surface of two given volumes. It was conjectured that the double bubble is made up by two pieces of spheres, meeting at an angle of 120°, and which are separated by a spherical surface.

  [source: Wikimedia Commons, File:Double_bubble.png]

  Hass and Schlafly (2000) proved the special case of two equally sized volumes by first showing that the minimal surface is made up by piecewise smooth surfaces, and then by showing optimality of the double bubble. Interval computation played an important role in verifying that certain value ranges of surface curvatures and meeting angles contain no optimal solution. Hutchings et al. (2002) proved the general case.

- **Goldbach's weak conjecture**

  It states that every odd number greater than 5 can be expressed as the sum of three primes. (Repetition of primes possible.)

  The conjecture was proved by Helfgott (2014). It used interval-arithmetic packages for rigorous calculations, verifying the conjecture for odd numbers up to about 10^{28}. For larger numbers, the proof went by using estimation techniques from analytic number theory.

- **The Dirac–Schwinger conjecture**

  This conjecture gives an asymptotic characterization of the ground-state energy of an atom in the quantum mechanical model. The role of interval computation was to rigorously enclose the solution of a specific ODE, and to verify an inequality related to the minimal energy property.

- There are also applications of interval computation in dynamical systems related to chaos theory, for instance, proving the number of digits of some constants appearing there.

  Computing a desired accuracy of Newton's constant of gravitation is discussed in Holzmann et al. (1996).

For related results see, e.g., Einarsson (2005); Meyer and Schmidt (1991).

2.1.2 Processing a set of states

Constraint programming and global optimization are the typical fields in which wide intervals are dealt with. We seek for an (optimal) solution of a set of constraints inside given ranges, so the intervals here represent the area to be completely processed.

- **Constraint programming**

  In constraint programming with continuous domains, interval analysis is utilized to shrink the domains, to remove interval boxes that contain no solution, and to prove in a given box there is a unique solution. As a result, we have a set of boxes, and each solution is guaranteed to lie in one of them. More in Section 8.1.

- **Global optimization**

  Interval approach to global optimization follows a basis branch & bound scheme. For each box, we can calculate lower and upper bound on the optimal value, which enables to remove the idle boxes. Moreover, constraint programming can be used to further process boxes. Again, this approach gives us guaranteed enclosures for global optima. More in Section 8.3.

Next, intervals are useful for calculating verified error bounds of various approximations.

- visualization of three dimensional surfaces (Stolte and Kaufman, 1998)

2.1.3 Modelling uncertainty by intervals

- **Mass number of ten elements**

  In December 2010, the International Union of Pure and Applied Chemistry decided to state the mass numbers of ten chemical elements (hydrogen, lithium, boron, carbon, nitrogen, oxygen, silicon, sulfur, chlorine and thallium) as intervals; for instance, the carbon has the mass interval number $[12.0096, 12.0116]$. It is not due to inaccuracy of measurement, but because there are several stable isotopes (the other elements have just one stable isotope). So far, there were considered average values, which turned out to be insufficient for reliable computations.

  See also Wieser et al. (2013).

- **Physical constants**

  Concerning gravitational acceleration, at different points on Earth, it ranges in about $[9.78,\ 9.82]\, ms^{-2}$. For example, Jakarta has $9.781\, ms^{-2}$ and Oslo has $9.819\, ms^{-2}$. Therefore, computing with the standardized value of $9.80665\, ms^{-2}$ may yield misleading results.

- **Robotics** *(Interval methods help a robot succeed, 1996)*

  The team of the University of Texas at El Paso succeeded in the International AAAI 1996 and 1997 Robot Competitions (the third and the first place, respectively). In AAAI 1996, robots had to deal with uncertainty in both sensors and actuators, and the robot designed by the team beaten the traditional statistical approach based robots by using the interval methods.

Notes and further reading. In his book, Bridger (2019) constructs the real numbers using intervals of rational numbers and interval arithmetic. The analogy to scientific measurement with known accuracy is particularly appealing for computer scientists. A justification of using intervals for modelling uncertainty is provided in Kreinovich (1995, 2007).

2.2 Getting started with intervals

Intervals. Not surprisingly, a real *interval* is defined as a set

$$\boldsymbol{a} \coloneqq [\underline{a}, \overline{a}] = \{a;\ \underline{a} \le a \le \overline{a}\},$$

where the left and right endpoints $\underline{a}, \overline{a} \in \mathbb{R} \cup \{-\infty, \infty\}$ are given. We will often use the notion of the midpoint and the radius of $\boldsymbol{a}$, defined respectively, as

$$a^c \coloneqq \frac{1}{2}(\underline{a} + \overline{a}), \quad a^\Delta \coloneqq \frac{1}{2}(\overline{a} - \underline{a}).$$

By definition, we also allow infinite values as the interval endpoints . An bounded interval can be defined also by its midpoint and radius

$$\boldsymbol{a} = [a^c - a^\Delta, a^c + a^\Delta] = \{a \in \mathbb{R};\ |a - a^c| \leq a^\Delta\}.$$

We sometimes use the left and right endpoints, midpoint and radius as functions with an interval argument

$$\inf(\boldsymbol{a}) \coloneqq \underline{a} = a^c - a^\Delta, \qquad \operatorname{mid}(\boldsymbol{a}) \coloneqq a^c = \frac{1}{2}(\underline{a} + \overline{a}),$$
$$\sup(\boldsymbol{a}) \coloneqq \overline{a} = a^c + a^\Delta, \qquad \operatorname{rad}(\boldsymbol{a}) \coloneqq a^\Delta = \frac{1}{2}(\overline{a} - \underline{a}).$$

When $a^\Delta = 0$, then the interval $\boldsymbol{a}$ consists of one real; we call such an interval *point* (or degenerate). Likewise, we will associate any real value $a \in \mathbb{R}$ with an interval $\boldsymbol{a} = [a, a]$.

The magnitude (the largest absolute value) and the *mignitude*[1] (the smallest absolute value) of $\boldsymbol{a}$ are defined as follows

$$\operatorname{mag}(\boldsymbol{a}) \coloneqq \max\{|a|;\ a \in \boldsymbol{a}\} = \max(|\underline{a}|, |\overline{a}|) = |a^c| + a^\Delta,$$
$$\operatorname{mig}(\boldsymbol{a}) \coloneqq \min\{|a|;\ a \in \boldsymbol{a}\} = \begin{cases} 0 & \text{if } 0 \in \boldsymbol{a}, \\ \min(|\underline{a}|, |\overline{a}|) & \text{otherwise.} \end{cases}$$

We will employ the following relation on intervals:[2]

$$\boldsymbol{a} \leq \boldsymbol{b} \ \Leftrightarrow\ \overline{a} \leq \underline{b} \ \Leftrightarrow\ \forall a \in \boldsymbol{a}\, \forall b \in \boldsymbol{b} : a \leq b,$$
$$\boldsymbol{a} < \boldsymbol{b} \ \Leftrightarrow\ \overline{a} < \underline{b} \ \Leftrightarrow\ \forall a \in \boldsymbol{a}\, \forall b \in \boldsymbol{b} : a < b.$$

Inclusion between intervals can be easily characterized as

$$\boldsymbol{a} \subseteq \boldsymbol{b} \ \Leftrightarrow\ \underline{b} \leq \underline{a} \leq \overline{a} \leq \overline{b} \ \Leftrightarrow\ |a^c - b^c| \leq b^\Delta - a^\Delta. \tag{2.1}$$

The set of all real intervals is denoted by $\mathbb{IR}$. Further set properties of intervals are listed in Table A.1 in the appendix.

Interval vectors. We use similar notation for interval vectors. Let $\underline{v}, \overline{v} \in \mathbb{R}^n$ such that $\underline{v} \leq \overline{v}$ entrywise. *An interval vector* is the set

$$\boldsymbol{v} \coloneqq [\underline{v}, \overline{v}] = \{v \in \mathbb{R}^n;\ \underline{v} \leq v \leq \overline{v}\}.$$

We extend the notation of the midpoint and radius to the vectors

$$v^c \coloneqq \frac{1}{2}(\underline{v} + \overline{v}), \quad v^\Delta \coloneqq \frac{1}{2}(\overline{v} - \underline{v}).$$

The set of all interval vectors of size n is denoted by $\mathbb{IR}^n$.

Interval vectors are sometimes also called as *boxes* because geometrically they represent rectangular parallelepipeds in space $\mathbb{R}^n$.

Since we work with intervals, it is often useful to intervalize non-interval sets. Let a set $\mathcal{S} \subset \mathbb{R}^n$ be a given bounded set, *the interval hull* $\square\mathcal{S}$ of $\mathcal{S}$ is the smallest interval vector containing $\mathcal{S}$, that is,

$$\square\mathcal{S} \coloneqq \bigcap_{\boldsymbol{v} \in \mathbb{IR}^n :\ \mathcal{S} \subseteq \boldsymbol{v}} \boldsymbol{v}.$$

As we will see later, determining the interval hull is often a computationally hard problem, and we focus more on computing an enclosure instead. *An enclosure* of a bounded set $\mathcal{S} \subset \mathbb{R}^n$ is any interval vector $\boldsymbol{v} \in \mathbb{IR}^n$ of such that $\mathcal{S} \subseteq \boldsymbol{v}$; see Figure 2.1. Naturally, one seeks for as small as possible enclosures.

[1] Introduced by Neumaier (1984).

[2] Other relations and orders on intervals exist, too; see Zapata et al. (2013)

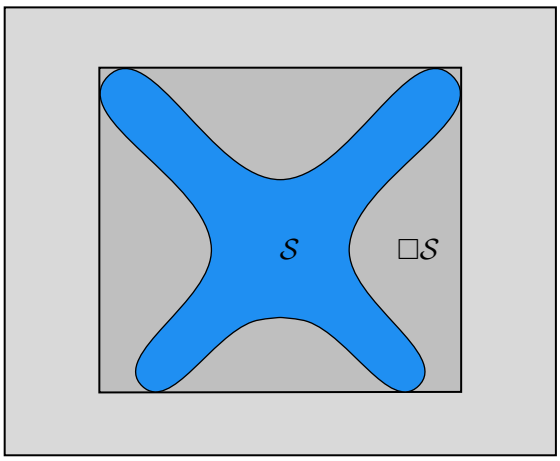


Figure 2.1: An enclosure and the interval hull of a set $\mathcal{S}$.

Interval matrices. Interval matrices generalize intervals and interval vectors and the notation is analogous. *An interval matrix* is the family of matrices

$$\boldsymbol{A} \coloneqq [\underline{A}, \overline{A}] = \{A \in \mathbb{R}^{m\times n};\ \underline{A} \le A \le \overline{A}\},$$

where $\underline{A}, \overline{A} \in \mathbb{R}^{m\times n}$ are given. The inequality $\le$ for matrices and vectors is understood entrywise throughout this book. The corresponding midpoint and radius matrices are

$$A^c \coloneqq \frac{1}{2}(\underline{A} + \overline{A}), \quad A^\Delta \coloneqq \frac{1}{2}(\overline{A} - \underline{A}).$$

The set of all interval matrices of size $m \times n$ is denoted by $\mathbb{IR}^{m\times n}$. An interval matrix with zero radius is called *point*.

We associate an interval matrix with a matrix with interval entries (and similarly for interval vectors). Thus, the (i,j)th entry of an interval matrix $\boldsymbol{A}$ is the interval $\boldsymbol{a}_{ij} = [\underline{a}_{ij}, \overline{a}_{ij}]$. For example, we can write

$$\boldsymbol{A} = \begin{pmatrix} [1,2] & 3 & [4,5] \\ 6 & [7,8] & 9 \end{pmatrix},$$

for which $\boldsymbol{a}_{12} = 3$ and $\boldsymbol{a}_{13} = [4,5]$, for instance.

Interval norms and other functions. *The magnitude* and *mignitude* of an interval matrix $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$ are applied entrywise, that is, $\operatorname{mag}(\boldsymbol{A}) \in \mathbb{IR}^{n\times n}$ and $\operatorname{mag}(\boldsymbol{A})_{ij} = \operatorname{mag}(\boldsymbol{a}_{ij})$ for all i,j; analogously for the mignitude. *The comparison matrix* is extended to the interval matrix input $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$ as follows: $\langle \boldsymbol{A} \rangle_{ii} = \operatorname{mig}(\boldsymbol{a}_{ii})$ and $\langle \boldsymbol{A} \rangle_{ij} = -\operatorname{mag}(\boldsymbol{a}_{ij})$ for $i \ne j$.

The norm of an interval vector $\boldsymbol{x} \in \mathbb{IR}^n$ or an interval matrix $\boldsymbol{A} \in \mathbb{IR}^{m\times n}$ is defined (as opposed to other functions) as the maximal norm over all realizations, so

$$\|\boldsymbol{x}\| \coloneqq \max_{x\in\boldsymbol{x}} \|x\|,$$
$$\|\boldsymbol{A}\| \coloneqq \max_{A\in\boldsymbol{A}} \|A\|.$$

Vector norms are easily computable by reduction to the real case,

$$\|\boldsymbol{x}\|_1 = \|\operatorname{mag}(\boldsymbol{x})\|_1, \quad \|\boldsymbol{x}\|_2 = \|\operatorname{mag}(\boldsymbol{x})\|_2, \quad \|\boldsymbol{x}\|_\infty = \|\operatorname{mag}(\boldsymbol{x})\|_\infty.$$

Most of the matrix norms are also easily reducible,

$$\|\boldsymbol{A}\|_F = \|\operatorname{mag}(\boldsymbol{A})\|_F, \quad \|\boldsymbol{A}\|_1 = \|\operatorname{mag}(\boldsymbol{A})\|_1, \quad \|\boldsymbol{A}\|_\infty = \|\operatorname{mag}(\boldsymbol{A})\|_\infty.$$

On the other hand, computing $\|\boldsymbol{A}\|_2$ is an *NP*-hard problem (Nemirovskii, 1993).

The spectral radius of an interval matrix $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$ is also defined as the largest possible value

$$\rho(\boldsymbol{A}) \coloneqq \max_{A\in\boldsymbol{A}} \rho(A).$$

Matrices A_{yz} and vectors b_y. In the theory of interval matrices, there often appear matrices of the special form. Let $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$, $y, z \in [-1,1]^n$ and denote

$$A_{y,z} \coloneqq A^c - \operatorname{diag}(y)A^\Delta \operatorname{diag}(z). \tag{2.2}$$

In the same fashion, we define

$$b_y \coloneqq b^c + \operatorname{diag}(y)b^\Delta. \tag{2.3}$$

It is easy to see that $A_{yz} \in \boldsymbol{A}$, but not every matrix from $\boldsymbol{A}$ has this form. Similarly for $b_y \in \boldsymbol{b}$. The reason why there is a subtraction in the definition of $A_{y,z}$ and addition in b_y is that usually $A_{y,z}$ stays on the left-hand side of a linear system, while b_y stays on the right-hand side.

Observation 2.2. *$A_{yz} \in \boldsymbol{A}$.*

Proof. We show that $|A_{yz} - A^c| \le A^\Delta$:

$$|A_{yz} - A^c| = |A^c - \operatorname{diag}(y)A^\Delta \operatorname{diag}(z) - A^c| \le |\operatorname{diag}(y)|A^\Delta|\operatorname{diag}(z)| = I_n A^\Delta I_n = A^\Delta. \qquad \square$$

Notice that if $y, z \in \{\pm 1\}^n$, then $(A_{yz})_{ij} \in \{\underline{a}_{ij}, \overline{a}_{ij}\}$, but not conversely. There are 2^{n^2} so-called *vertex matrices* that have the property that $a_{ij} \in \{\underline{a}_{ij}, \overline{a}_{ij}\}$, however, only 2^{2n} of them has the form of A_{yz} for certain $y, z \in \{\pm 1\}^n$.

Example 2.3. If $y = e$ and $z = e$, then $A_{yz} = \underline{A}$. If $y = -e$ and $z = e$, then $A_{yz} = \overline{A}$. However, the matrix

$$\begin{pmatrix} \overline{a}_{11} & \overline{a}_{12} \\ \underline{a}_{21} & \overline{a}_{22} \end{pmatrix}$$

can be written in the form A_{yz} for no $y, z \in \{\pm 1\}^2$. $\square$

Symmetric interval matrices. In some situations, symmetric interval matrices appear. When $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$, then the corresponding *symmetric interval matrix* is defined as $\boldsymbol{A}^S \coloneqq \{A \in \boldsymbol{A};\ A = A^T\}$. Thus, it consists of the symmetric matrices from $\boldsymbol{A}$. Without loss of generality we assume that $\underline{A}$ and $\overline{A}$ (and also A^c and A^Δ) are symmetric; otherwise we can shrink intervals in $\boldsymbol{A}$ without losing any symmetric matrix. (We will need this assumption, e.g., in a characterization of positive definiteness.)

Notice that a symmetric interval matrix is not an interval matrix by definition!

Range of a function. One of the fundamental problems in interval analysis is computation the range of functions over intervals. For simplicity, consider a single-valued real function $f\colon \mathbb{R}^n \to \mathbb{R}$ and an interval vector $\boldsymbol{x} \in \mathbb{IR}^n$. *The range* of f over $\boldsymbol{x}$ is defined as

$$f(\boldsymbol{x}) \coloneqq \{f(x);\ x \in \boldsymbol{x}\}. \tag{2.4}$$

As long as f is continuous, its range $f(\boldsymbol{x})$ is a compact interval. In general, however, the range $f(\boldsymbol{x})$ need not be an interval, and also determining $\square f(\boldsymbol{x})$ can be computationally demanding. Fortunately, for continuous and monotone functions, the range is easy to determine. If f is non-decreasing on $\boldsymbol{x} \in \mathbb{IR}$, then $f(\boldsymbol{x}) = [f(\underline{x}), f(\overline{x})]$ and likewise for non-increasing functions or piecewise monotone functions. Thus we have, for example,

- $\exp(\boldsymbol{x}) = [\exp(\underline{x}), \exp(\overline{x})]$,
- $\sin(\boldsymbol{x}) = [\sin(\underline{x}), \sin(\overline{x})]$ provided $\boldsymbol{x} \subseteq [-\frac{\pi}{2}, \frac{\pi}{2}]$,
- $\frac{1}{\boldsymbol{x}+1} = [\frac{1}{\overline{x}+1}, \frac{1}{\underline{x}+1}]$ provided $\underline{x} > -1$,
- $\boldsymbol{x}^2 = [\min(\underline{x}^2, \overline{x}^2), \max(\underline{x}^2, \overline{x}^2)]$ provided $0 \notin \boldsymbol{x}$, and $\boldsymbol{x}^2 = [0, \max(\underline{x}^2, \overline{x}^2)]$ otherwise.

Another simple functions are arithmetic operations, discussed more in detail in the next section.

2.3 Interval arithmetic

The basic tool to deal with intervals is by interval arithmetic. There are still people that believe that interval analysis is *just* interval arithmetic. This is not true! Since the sixties, a lot of sophisticated theoretical properties and practical methods have been developed that makes interval analysis an efficient technique for solving many problems. Notice that a blind use of interval arithmetic leads to poor results; one has to apply it more carefully.

Interval arithmetic is defined naturally as a range of values over interval domains. Let $\circ$ be a basic operation – addition, subtraction, multiplication or division. For $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{IR}$ we define

$$\boldsymbol{a} \circ \boldsymbol{b} \coloneqq \{a \circ b;\ a \in \boldsymbol{a}, b \in \boldsymbol{b}\},$$

with $0 \notin \boldsymbol{b}$ in case of division. It is not hard to see that for particular operations the interval arithmetic reads

$$\begin{aligned}
\boldsymbol{a} + \boldsymbol{b} &= [\underline{a} + \underline{b}, \overline{a} + \overline{b}], \\
\boldsymbol{a} - \boldsymbol{b} &= [\underline{a} - \overline{b}, \overline{a} - \underline{b}], \\
\boldsymbol{a}\boldsymbol{b} &= [\min(\underline{a}\underline{b}, \underline{a}\overline{b}, \overline{a}\underline{b}, \overline{a}\overline{b}), \max(\underline{a}\underline{b}, \underline{a}\overline{b}, \overline{a}\underline{b}, \overline{a}\overline{b})], \\
\boldsymbol{a}/\boldsymbol{b} &= [\min(\underline{a}/\underline{b}, \underline{a}/\overline{b}, \overline{a}/\underline{b}, \overline{a}/\overline{b}), \max(\underline{a}/\underline{b}, \underline{a}/\overline{b}, \overline{a}/\underline{b}, \overline{a}/\overline{b})].
\end{aligned}$$

Looking at real numbers as degenerate intervals, interval arithmetic generalizes the classical one, and we can make mixed expressions like $2 + 3[4, 5]$ with the meaning $[2, 2] + [3, 3][4, 5]$.

Looking at the signs of the endpoints of $\boldsymbol{a}$ and $\boldsymbol{b}$, the interval multiplication and division can be performed more economically than by checking all four combinations. For more details see Table A.2.

Remark 2.4 (Implementation)**.** When implemented in a programming language, one has to pay attention to the rounding modes. Interval arithmetic must be implemented in such a way that the output interval always contains all possible results of realizations of intervals. For example, for addition of two intervals $\boldsymbol{a} + \boldsymbol{b}$, the lower bound $\underline{a} + \underline{b}$ has to be computed in downward rounding while the upper bound $\overline{a} + \overline{b}$ in the upward rounding mode. Basically, this book mostly handles intervals in the same way as we could compute standard operations exactly. Nevertheless, we give implementation comments where necessary.

Some of the algebraic laws remain valid for interval arithmetic, some of them hold true only in a weaker form, and some do not hold. For instance, if $\boldsymbol{a}$ is a nondegenerate interval, then $\boldsymbol{a} - \boldsymbol{a} \neq 0$. If $0 \notin \boldsymbol{a}$ in addition, then $\boldsymbol{a}/\boldsymbol{a} \neq 1$. Properties of interval arithmetic are listed in detail in Tables A.3–A.11 in the appendix.

Theorem 2.5 (Basic properties of interval arithmetic)**.**

- *Interval addition and multiplication are commutative and associative.*
- *Interval addition and multiplication are not distributive in general, but they are sub-distributive instead, that is,*

$$\forall \boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c} \in \mathbb{IR} : \boldsymbol{a}(\boldsymbol{b} + \boldsymbol{c}) \subseteq \boldsymbol{a}\boldsymbol{b} + \boldsymbol{a}\boldsymbol{c}. \tag{2.5}$$

Proof. Commutativity and associativity are easy to see. To prove sub-distributivity, write

$$\begin{aligned}
\boldsymbol{a}(\boldsymbol{b} + \boldsymbol{c}) &= \{a(b + c);\ a \in \boldsymbol{a}, b \in \boldsymbol{b}, c \in \boldsymbol{c}\} \\
&= \{ab + ac;\ a \in \boldsymbol{a}, b \in \boldsymbol{b}, c \in \boldsymbol{c}\} \\
&\subseteq \{ab + a'c;\ a, a' \in \boldsymbol{a}, b \in \boldsymbol{b}, c \in \boldsymbol{c}\} \\
&= \boldsymbol{a}\boldsymbol{b} + \boldsymbol{a}\boldsymbol{c}.
\end{aligned}$$

□

The cases in which distributivity holds were characterized by Ratschek (1971); Spaniol (1970); see also Mayer (2017). For instance, distributivity $\boldsymbol{a}(\boldsymbol{b} + \boldsymbol{c}) = \boldsymbol{a}\boldsymbol{b} + \boldsymbol{a}\boldsymbol{c}$ holds if $a^{\Delta} = 0$ (i.e., $\boldsymbol{a}$ is real) or if $\underline{b} \geq 0$ and $\underline{c} \geq 0$.

Example 2.6. To see concretely that interval arithmetic is not distributive in general, consider, e.g., the following example with $\boldsymbol{a} = [1, 2]$, $\boldsymbol{b} = 1$, $\boldsymbol{c} = -1$:

$$\begin{aligned}\boldsymbol{a}(\boldsymbol{b} + \boldsymbol{c}) &= [1, 2] \cdot (1 - 1) = [1, 2] \cdot 0 = 0,\\ \boldsymbol{ab} + \boldsymbol{ac} &= [1, 2] \cdot 1 + [1, 2] \cdot (-1) = [1, 2] - [1, 2] = [-1, 1].\end{aligned}$$ □

The intuitive reason why (2.5) does not hold as equation in general is that the interval $\boldsymbol{a}$ appears twice in the right-hand side. Since interval arithmetic has no memory and considers intervals independently, the resulting interval can overestimate the true range of the expression over intervals.

This is called the *dependence problem*. Multiple appearances of the same interval quantities in interval expressions cause overestimations. This is one of the fundamental difficulties in interval analysis. However, (2.5) also shows that various rearrangements of interval expressions lead to various resulting intervals. Thus, we can try to find such a rearrangement that has possibly a small overestimation.

Example 2.7. Consider the expression $x^2 - x$ with $\boldsymbol{x} = [-1, 2]$. The expression has three equivalent forms $x^2 - x = x(x - 1) = (x - \frac{1}{2})^2 - \frac{1}{4}$, but the interval evaluations differ:

$$\begin{aligned}\boldsymbol{x}^2 - \boldsymbol{x} &= [-1, 2]^2 - [-1, 2] = [0, 4] - [-1, 2] = [-2, 5],\\ \boldsymbol{x}(\boldsymbol{x} - 1) &= [-1, 2]([-1, 2] - 1) = [-1, 2][-2, 1] = [-4, 2],\\ (\boldsymbol{x} - \tfrac{1}{2})^2 - \tfrac{1}{4} &= ([-1, 2] - \tfrac{1}{2})^2 - \tfrac{1}{4} = [-\tfrac{3}{2}, \tfrac{3}{2}]^2 - \tfrac{1}{4} = [0, \tfrac{9}{4}] - \tfrac{1}{4} = [-\tfrac{1}{4}, 2].\end{aligned}$$

The last expression is optimal, meaning that no other arrangement of the expression leads to a better (tighter) interval. This is because the interval $\boldsymbol{x}$ appears just once in this case (cf. Theorem 2.16) □

As the example below shows, the cumulative overestimation may grow exponentially fast.

Example 2.8. Consider a discrete dynamical system, where the state x_n at time n is given by a recursive formula $x_n = Ax_{n-1}$. For concreteness, let A be a rotation matrix and the initial state x_0 known only to belong to an interval vector as follows (Lohner, 2001)

$$A = \begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix}, \quad x_0 \in \boldsymbol{x} = \begin{pmatrix} [1-\varepsilon, 1+\varepsilon] \\ [1-\varepsilon, 1+\varepsilon] \end{pmatrix}.$$

Tight bounds for x_n are computed easily as

$$x_n \in A^n \boldsymbol{x},$$

but if we calculate the state bounds sequentially as

$$x_n \in A(A \dots A(A\boldsymbol{x}) \dots),$$

then the radius d_n of the right-hand side interval vector grows exponentially as

$$d_n = |A| d_{n-1} = \begin{pmatrix} |\cos\varphi| & |\sin\varphi| \\ |\sin\varphi| & |\cos\varphi| \end{pmatrix} d_{n-1} = (|\sin\varphi| + |\cos\varphi|)^n \begin{pmatrix} \varepsilon \\ \varepsilon \end{pmatrix}.$$ □

Other useful properties of intervals are mentioned below.

Proposition 2.9. *Let* $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{IR}$. *Then*

(1) $(\boldsymbol{a} \pm \boldsymbol{b})^c = a^c \pm b^c$,

(2) $(\boldsymbol{a} \pm \boldsymbol{b})^\Delta = a^\Delta + b^\Delta$,

(3) $\boldsymbol{a} \subseteq \boldsymbol{b}$ *if and only if* $|a^c - b^c| \le b^\Delta - a^\Delta$,

(4) $\boldsymbol{a} \cap \boldsymbol{b} \neq \emptyset$ *if and only if* $0 \in \boldsymbol{a} - \boldsymbol{b}$ *and if and only if* $\overline{a} \ge \underline{b} \wedge \underline{a} \le \overline{b}$.

Proof. Items (1), (2), (4) are trivial. To show (3), write $\boldsymbol{a} \subseteq \boldsymbol{b}$ as $\overline{a} \le \overline{b}$, $-\underline{b} \le -\overline{b}$, or as

$$a^c + a^\Delta \le b^c + b^\Delta, \quad -a^c + a^\Delta \le -b^c + b^\Delta,$$

which is equivalent to $\pm(a^c - b^c) \le b^\Delta - a^\Delta$. □

Interval matrix operations. Interval arithmetic naturally extends to interval matrix operations. For $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{IR}^{m\times n}$, we define $\boldsymbol{A} \pm \boldsymbol{B} \in \mathbb{IR}^{m\times n}$ as

$$(\boldsymbol{A} \pm \boldsymbol{B})_{ij} \coloneqq \boldsymbol{a}_{ij} \pm \boldsymbol{b}_{ij},$$

and for $\boldsymbol{A} \in \mathbb{IR}^{m\times p}, \boldsymbol{B} \in \mathbb{IR}^{p\times n}$, we define $\boldsymbol{AB} \in \mathbb{IR}^{m\times n}$ as

$$(\boldsymbol{AB})_{ij} \coloneqq \sum_{k=1}^{p} \boldsymbol{a}_{ik}\boldsymbol{b}_{kj}.$$

Not surprisingly, the resulting interval matrices contain all results of those operations for all realizations of intervals. It is the componentwise tightest enclosure.

Proposition 2.10. *Let $\boldsymbol{A} \in \mathbb{IR}^{m\times p}$ and $\boldsymbol{B} \in \mathbb{IR}^{p\times n}$. Then*

$$\boldsymbol{AB} = \square\{AB;\ A \in \boldsymbol{A},\ B \in \boldsymbol{B}\}, \tag{2.6}$$

but in general

$$\boldsymbol{AB} \neq \{AB;\ A \in \boldsymbol{A},\ B \in \boldsymbol{B}\}. \tag{2.7}$$

Proof. Equation (2.6) follows from the basic properties of interval arithmetic; see also Theorem 2.16 later on.

To show (2.7), consider interval matrices

$$\boldsymbol{A} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \boldsymbol{B} = ([1, 2]).$$

Then

$$\boldsymbol{AB} = \begin{pmatrix} [1, 2] \\ [1, 2] \end{pmatrix},$$

but matrix $\binom{1}{2} \in \boldsymbol{AB}$ is not attained as the product AB for particular realizations $A \in \boldsymbol{A}$ and $B \in \boldsymbol{B}$. □

Matrix operations inherit several nice properties such as commutativity and associativity of addition and subdistributivity.

Another useful property of matrix interval operations is that all vectors in the product $\boldsymbol{A}b$ are achieved as Ab for some $A \in \boldsymbol{A}$. By Proposition 2.10, this is not true for the interval matrix product $\boldsymbol{AB}$ in general.

Proposition 2.11. *Let $\boldsymbol{A} \in \mathbb{IR}^{m\times n}$ and $b \in \mathbb{R}^n$. Then*

$$\boldsymbol{A}b = [A^c b - A^\Delta|b|,\ A^c b + A^\Delta|b|] = \{Ab;\ A \in \boldsymbol{A}\}. \tag{2.8}$$

Proof. "Second equality." For each $A \in \boldsymbol{A}$ we have

$$Ab = A^c b + (A - A^c)b \leq A^c b + |A - A^c||b| \leq A^c b + A^\Delta|b|,$$

and similarly from below, which proves "$\supseteq$". To show "$\subseteq$", let d be any vector such that $|d| \leq A^\Delta|b|$. Define $z \coloneqq \operatorname{sgn}(b)$ and $y \in [-1, 1]^m$ as follows

$$y_i = \begin{cases} \frac{d_i}{(A^\Delta|b|)_i} & \text{if } (A^\Delta|b|)_i > 0, \\ 1 & \text{otherwise.} \end{cases}$$

Then

$$A^c b + d = A^c b + \operatorname{diag}(y)A^\Delta|b| = A^c b + \operatorname{diag}(y)A^\Delta \operatorname{diag}(z)b = A_{-yz}b.$$

"First equality." To show the first equation in (2.8), write

$$\begin{aligned} \{Ab;\ A \in \boldsymbol{A}\} \subseteq \boldsymbol{A}b &= A^c b + [-A^\Delta, A^\Delta]b \\ &\subseteq A^c b + [-A^\Delta|b|, A^\Delta|b|] = [A^c b - A^\Delta|b|,\ A^c b + A^\Delta|b|]. \end{aligned}$$

□

Further properties of operations with interval matrices are recored in Tables B.1–B.10 in the appendix.

Table 2.1: Division in extended interval arithmetic when $0 \in \boldsymbol{b}$.

$\boldsymbol{a}/\boldsymbol{b}$	$\boldsymbol{b} = 0$	$\underline{b} < \overline{b} = 0$	$\underline{b} < 0 < \overline{b}$	$0 = \underline{b} < \overline{b}$
$\overline{a} < 0$	$\emptyset$	$[\overline{a}/\underline{b}, \infty)$	$(-\infty, \overline{a}/\overline{b}] \cup [\overline{a}/\underline{b}, \infty)$	$(-\infty, \overline{a}/\overline{b}]$
$\underline{a} \leq 0 \leq \overline{a}$	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$	$\mathbb{R}$
$\underline{a} > 0$	$\emptyset$	$(-\infty, \underline{a}/\underline{b}]$	$(-\infty, \underline{a}/\underline{b}] \cup [\underline{a}/\overline{b}, \infty)$	$[\underline{a}/\overline{b}, \infty)$

Extended interval arithmetic. Sometimes, it is useful to extend interval arithmetic to enable division by an interval containing zero. Suppose $\boldsymbol{b} \neq 0$. We define the extended division by

$$\boldsymbol{a}/\boldsymbol{b} \coloneqq \{a/b;\ a \in \boldsymbol{a},\ 0 \neq b \in \boldsymbol{b}\},$$

so it contains all possible divisions except division by zero. It is either an interval or a union of two intervals. The explicit formula for the extended division is given in Table 2.1 for the singular case $0 \in \boldsymbol{b}$.

We cannot utilize extended interval arithmetic as is, but we have to pay attention whether it is appropriate to use for a given problem. We will employ the extended interval arithmetic in the interval Gauss–Seidel method (Remark 3.33) and in the Interval Newton method (Section 8.1.1).

Extended interval arithmetic was introduced independently by Hanson (1968); Kahan (1968b) and later elaborated by Ratz (1996), among others.

2.4 Evaluation of functions over intervals

Let us come back to the problem of computing the range $f(\boldsymbol{x})$ of a function $f\colon \mathbb{R}^n \to \mathbb{R}$ over an interval vector $\boldsymbol{x} \in \mathbb{R}^n$. Since the range $f(\boldsymbol{x})$ need not be an interval, we can focus on computations of its interval hull $\Box f(\boldsymbol{x})$ instead. Nevertheless, calculating $\Box f(\boldsymbol{x})$ is still a very difficult problem in general. Indeed, we will see later in Chapter 8 that the problem is undecidable – there provably exists no algorithm for computing it.

In view of this result, we have to resign on computing $\Box f(\boldsymbol{x})$, and turn our attention to calculate as tight as possible enclosure to $f(\boldsymbol{x})$. Even calculating a tight enclosure is still a computationally and theoretically hard problem.

In order to develops enclosure methods for $f(\boldsymbol{x})$, consider an interval function $\boldsymbol{f}\colon \mathbb{IR}^n \to \mathbb{IR}$. What properties should $\boldsymbol{f}$ have?

Definition 2.12 (Inclusion isotonicity)**.** A function $\boldsymbol{f}\colon \mathbb{IR}^n \to \mathbb{IR}$ is *inclusion isotonic* if for every $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{IR}^n$:

$$\boldsymbol{x} \subseteq \boldsymbol{y} \Rightarrow \boldsymbol{f}(\boldsymbol{x}) \subseteq \boldsymbol{f}(\boldsymbol{y}).$$

Definition 2.13 (Interval extension)**.** A function $\boldsymbol{f}\colon \mathbb{IR}^n \to \mathbb{IR}$ is *an interval extension* of $f\colon \mathbb{R}^n \to \mathbb{R}$ if for every $x \in \mathbb{R}^n$:

$$f(x) = \boldsymbol{f}(x).$$

These two properties, inclusion isotonicity (also called inclusion monotonicity) and interval extension, are enough to get a proper enclosure for the range of $f(\boldsymbol{x})$. The following theorem was already given by Moore (1966).

Theorem 2.14 (Fundamental theorem of interval analysis)**.** *If $\boldsymbol{f}\colon \mathbb{IR}^n \to \mathbb{IR}$ is inclusion isotonic and is an interval extension of $f\colon \mathbb{R}^n \to \mathbb{R}$, then for every $\boldsymbol{x} \in \mathbb{IR}^n$:*

$$f(\boldsymbol{x}) \subseteq \boldsymbol{f}(\boldsymbol{x}).$$

Proof. For every $x \in \boldsymbol{x}$, one has by interval extension and inclusion isotonicity that $f(x) = \boldsymbol{f}(x) \subseteq \boldsymbol{f}(\boldsymbol{x})$, whence $f(\boldsymbol{x}) \subseteq \boldsymbol{f}(\boldsymbol{x})$. □

We now have the desired properties, but have no example yet. We show that a simple example is obtained by a direct engagement of interval arithmetic. Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be a function that can be written as an arithmetic expression using only a finite number of arithmetic operations. Fix such an expression, and the corresponding *natural interval extension* $\boldsymbol{f}$ of f is defined by that expression when replacing the real arithmetic by the interval one. Herein, we assume that there is no division by an interval containing zero. The notion of natural interval extension is attributed to Moore (1966).

Theorem 2.15. *Natural interval extension of an arithmetic expression is both an interval extension and inclusion isotonic.*

Proof. It is easy to see that interval arithmetic is both an interval extension and inclusion isotonic. Next, proceed by mathematical induction. □

For concreteness, reconsider Example 2.7. The function $f(x) = x^2 - x$ has the natural interval extension $\boldsymbol{f}_1(\boldsymbol{x}) = \boldsymbol{x}^2 - \boldsymbol{x}$. However, $f(x) = x(x-1)$ and $f(x) = (x - \frac{1}{2})^2 - \frac{1}{4}$ are mathematically equivalent formulations of f that yield natural interval extensions $\boldsymbol{f}_1(\boldsymbol{x}) = \boldsymbol{x}^2 - \boldsymbol{x}$ and $\boldsymbol{f}_2(\boldsymbol{x}) = (\boldsymbol{x} - \frac{1}{2})^2 - \frac{1}{4}$, respectively. We see that different formulations lead to different natural interval extensions, which, as a consequence, give different enclosures of $f(x)$. Thus, natural interval extension is formulation dependent and one has to pay attention which of the various equivalent formulations to use.

In Example 2.7, the third formulation $f(x) = (x - \frac{1}{2})^2 - \frac{1}{4}$ gave rise to the tightest enclosure to $f(\boldsymbol{x})$. This is not by chance. The reason is that the variable appears only ones in that expression, in which case the natural interval extension yields the optimal enclosures (Moore, 1966; Skelboe, 1974).

Theorem 2.16. *Suppose that in an expression of $f\colon \mathbb{R}^n \to \mathbb{R}$ each variable $x_1, \ldots, x_n$ appears at most once. The corresponding natural interval extension $\boldsymbol{f}(\boldsymbol{x})$ satisfies for every $\boldsymbol{x} \in \mathbb{IR}^n$: $f(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x})$.*

Proof. Inclusion "$\subseteq$" holds by Theorems 2.14 and 2.15. Inclusion "$\supseteq$" follows from the assumption and optimality of interval arithmetic. Let $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{IR}$ and $\circ$ an arithmetic operation. Now, for each $c \in \boldsymbol{a} \circ \boldsymbol{b}$ there are $a \in \boldsymbol{a}$ and $b \in \boldsymbol{b}$ such that $c = a \circ b$. By induction, we have that for each $y \in \boldsymbol{f}(\boldsymbol{x})$ there is $x \in \boldsymbol{x}$ such that $y = f(x)$. We used the fact that each variable $x_1, \ldots, x_n$ appears at most once in the expression for $f(x)$, so there is no conflict in the selection of $x \in \boldsymbol{x}$. □

Apart from the natural interval extension, there are other inclusion isotonic interval extensions; some of them are discussed in Chapter 6. Basically, there exists many interval extensions: If $\boldsymbol{f}(\boldsymbol{x})$ is an interval extension of a function $f\colon \mathbb{R} \to \mathbb{R}$, then $\boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{x} - \boldsymbol{x}$ is another interval extension of f, and one can iterate further. Examples of interval functions that are interval extensions of some basic ones, but not inclusion isotonic, are, e.g., $\|\boldsymbol{A}\|$, $\mathrm{mag}(\boldsymbol{a})$, etc.

Notes and further reading. It is easy to see that inclusion isotonicity is preserved under function composition. Note that the extended interval arithmetic is also inclusion isotonic (Kearfott, 1996a; Ratz, 1996).

2.5 Interval hull of sets

In general, it is a very hard problem to compute the interval hull of a set $\mathcal{S} \subseteq \mathbb{R}^n$. One way to compute $\square\mathcal{S}$ is to reduce the problem to $2n$ optimization problems

$$\inf(\square\Sigma)_i = \min\{x_i;\, x \in \mathcal{S}\}, \quad i = 1, \ldots, n$$
$$\sup(\square\Sigma)_i = \max\{x_i;\, x \in \mathcal{S}\}, \quad i = 1, \ldots, n,$$

which calculate the minimum and maximum of ith coordinate x_i, respectively. Even though these optimization problems are difficult to solve in general, they become tractable as long as the set $\mathcal{S}$ has favourable properties, such as convexity. This is particularly the case if $\mathcal{S}$ is a convex polyhedron.

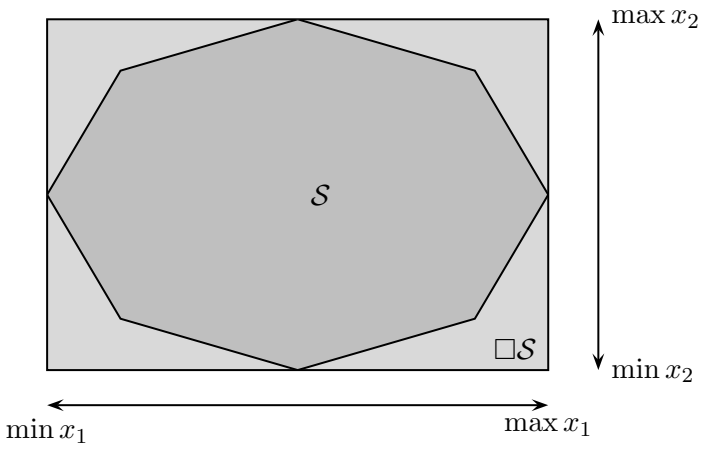


Figure 2.2: The interval hull $\square\mathcal{S}$ of a convex polyhedron $\mathcal{S}$.

Convex polyhedra. If the set $\mathcal{S}$ is described by a system of linear inequalities $Ax \leq b$, then $\mathcal{S}$ forms a convex polyhedron (see page 16). Calculation of the interval hull $\square\mathcal{S}$ then reduces to solving $2n$ linear programs

$$\inf(\square\Sigma)_i = \min\{x_i;\ Ax \leq b\}, \quad i = 1, \ldots, n$$
$$\sup(\square\Sigma)_i = \max\{x_i;\ Ax \leq b\}, \quad i = 1, \ldots, n.$$

Figure 2.2 illustrates this situation.

Even though linear programs are polynomially solvable, it takes some time of computation to solve $2n$ of them. Since all the linear programs share the same feasible set, there is an open room for an efficient computation of them. This was the aim of Achterberg's heuristic introduced in Baharev et al. (2009); it tries to reduce the computational effort by a suitable order of solving the linear programs.

Zonotopes. Zonotopes are special convex polyhedra that are defined as images of interval vectors under a linear map. More formally, let $A \in \mathbb{R}^{m\times n}$ and $\boldsymbol{x} \in \mathbb{IR}^n$. Then the corresponding *zonotope* is defined as

$$\mathcal{Z} \coloneqq \{Ax;\ x \in \boldsymbol{x}\}.$$

Thus, we have in hand neither a system of linear inequalities nor a list or vertices; zonotopes are characterized by another mean. Indeed, it may be computationally consuming to determine vertices of $\mathcal{Z}$ or its defining inequalities simply because there can be exponentially many such inequalities.

Zonotopes possess many favourable and interesting properties. They are centrally symmetric about point Ax^c, and each face of $\mathcal{Z}$ is again a zonotope. The interval hull of a zonotope is easily calculated by interval arithmetic,

$$\square\mathcal{Z} = A\boldsymbol{x}.$$

Example 2.17. Let

$$A = \begin{pmatrix} 1 & 3 & -6 & -10 \\ -3 & 5 & -2 & 10 \end{pmatrix}, \quad \boldsymbol{x} = [-1, 1]^4.$$

Figure 2.3 depicts the corresponding zonotope and its interval hull. □

2.6 Historical notes

Interval approximations date back to Archimedes, who enclosed the irrational number π by the interval $[3\frac{10}{71}, 3\frac{1}{7}]$. During the time, several approaches to deal with interval (or more general) values appeared (Kantorovich, 1962; Warmus, 1956; Young, 1931). The foundations of modern interval computation were set up by Moore (1962, 1966); Sunaga (1958). They come up with the idea of computing rigorous bounds

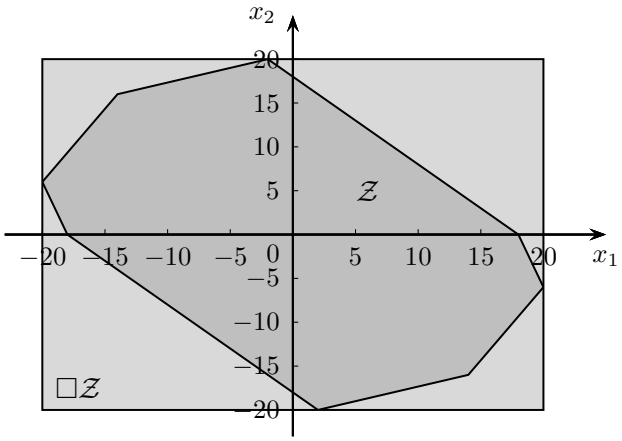


Figure 2.3: (Example 2.3) The interval hull $\square\mathcal{Z}$ of zonotope $\mathcal{Z}$.

by using interval enclosures for numerical problems such as solving nonlinear equations (Section 8.1.1), numerical integration (Section 7.3), or initial value problems in ordinary differential equations (Section 7.4).

In particular, Moore's dissertation in 1962 and book (Moore, 1966) achieved high interest and initiated a large research in interval computation. In his honor, there is the *Moore Prize* (`https://interval.louisiana.edu/Moore_prize.html`) for the best dissertation or paper in applications of interval analysis. So far, there have been the following winners:

- *Warwick Tucker (2002)* has proved, using interval techniques, that the renowned Lorenz equations do in fact possess a strange attractor. This solved the 14th Smale's problem (Smale, 1998).
- *Thomas C. Hales (2004)* received Moore Prize for his solution of the Kepler conjecture about the densest arrangement of spheres in space.
- *Kyoko Makino and Martin Berz (2008)* received Moore Prize for their work on suppression of the wrapping effect by Taylor model-based verified integrators.
- *Luc Jaulin (2012)* received the prize for his paper "A nonlinear set-membership approach for the localization and map building of an underwater robot using interval constraint propagation" published in IEEE Transactions on Robotics in 2009 (Vol. 25, No. 1, pp. 88–98).
- *Kenta Kobayashi (2014)* was awarded the prize for his paper "Computer-assisted uniqueness proof for Stokes' wave of extreme form" published in Nankai Series in Pure, Applied Mathematics and Theoretical Physics in 2013 (Vol. 10, pp. 54–67).
- *Balazs Banhelyi, Tibor Csendes, Tibor Krisztin and Arnold Neumaier (2016)* were awarded the prize for their paper "Global attractivity of the zero solution for Wright's equation" published in SIAM Journal on Applied Dynamical Systems in 2014 (Vol. 13, No. 1, pp. 537–563).
- *Jordi-Lluís Figueras, Àlex Haro and Alejandro Luque (2018)* were awarded the prize for their paper "Rigorous computer-assisted application of KAM theory: A modern approach" published in Foundations of Computational Mathematics in 2017 (Vol. 17, No. 5, pp. 1123–1193).
- *Marko Lange and Siegfried M. Rump (2021)* were awarded the prize for their paper "Verified inclusions for a nearest matrix of specified rank deficiency via a generalization of Wedin's $\sin(\theta)$ theorem" published in BIT Numerical Mathematics in 2021 (Vol. 61, pp. 361–380).

First systematic treatment of interval matrices is due to Apostolatos and Kulisch (1968).

Interval community

Interval community shares its knowledge and experiences by organising conferences, publishing book and journal papers, and posting information on the Internet, among others.

- **Mailing list**

  The Interval community has its interval mailing list `[Reliable Computing]`; see

  `http://www.cs.utep.edu/interval-comp/ftp.html`

  The mailing list serves for conference announcements and for discussion on various interval computation problems.

- **Web pages**

  Web pages

  `http://www.cs.utep.edu/interval-comp/`

  administrated by Vladik Kreinovich contain many useful links related to interval computation, including information on interval software, conferences, publications, applications, and researchers.

- **Journals**

  The primal journal of the interval computations community is the *Reliable Computing*. I is "an open electronic journal devoted to mathematical computations with guaranteed accuracy, bounding of ranges, mathematical proofs based on floating point arithmetic, and other theory and applications of interval arithmetic and directed rounding". It started in 1991 under the name *Interval Computations*, and then changed to Reliable Computing in 1995 being published by Kluwer Academic Publishers. Since 2010, it became a free electronic journal.

  Nevertheless, interval community researchers publish in a lot of other prestigious journals on computing, numerical analysis, linear algebra, fuzzy theory and optimization.

- **Conferences**

  There are two international conferences devoted directly to interval computation. First,

  SCAN, International Symposium on Scientific Computing, Computer Arithmetics and Verified Numerics[3)]

  is organized (mostly) biannually since 1989. It was initiated by the University of Karlsruhe, Germany, and supported by GAMM (International Association of Applied Mathematics and Mechanics) and IMACS (International Association for Mathematics and Computers in Simulation).

  Second,

  SWIM, Summer Workshop on Interval Methods,
  (until 2015 Small Workshop on Interval Methods),

  established by Luc Jaulin and Nacim Ramdani, is held annually since 2008.

  Moreover, interval computation sessions are organized at many other international conferences; see the regularly updated list of forthcoming conferences

  `http://www.cs.utep.edu/interval-comp/conf.html`

- **Intervals Day** (March 14)

  Mathematicians are celebrating this day as Pi Day since 03/14 reminds of $\pi = 3.14\ldots$ Since Archimedes was among the first who estimated π from below and from above, the interval community appropriates this day also as Intervals Day.

[3)] Until 2014 known as *the GAMM – IMACS International Symposium on Scientific Computing, Computer Arithmetic, and Verified Numerical Computation.*

Recognition

- L.N. Trefethen, IV.21 Numerical Analysis, in Gowers (2008), the section about the future (p. 614):

  > *When these are combined with techniques of interval arithmetic, there is even the prospect of accuracy guaranteed with respect to rounding as well as discretization error.*

2.7 Software

Despite the fast development of hardware, software, programming languages and related areas, and being aware of getting out-of-date soon, we mention current implementations of interval computation below.

INTLAB (Rump, 1999a) is a MATLAB toolbox capable of performing interval calculations. It implements interval arithmetic and elementary functions of interval arguments and provides a flexible programming environment.

VERSOFT (Rohn, 2019b) is a verification software written in INTLAB; implementation of linear algebraic functions for real / complex or interval data. It includes linear system solving, eigenvalues, matrix decompositions, basic optimization problems, among others.

Both INTLAB and VERSOFT are free of charge for noncommercial use.

- *Interval* package for Octave (by O. Heimlich),
  free package of verified interval functions
  `https://wiki.octave.org/Interval_package`

There are several interval libraries for C++ in GNU Linux, for instance:

- *C-XSC* (Hofschuster and Krämer, 2004),
  The C-XSC (eXtended Scientific Computing) library for reliable computing provides useful data types, such as real and complex intervals and matrices, and functionalities, such as overloading, controlled rounding or dot products in arbitrary / exact precision.
  `http://www.xsc.de/`, `http://www2.math.uni-wuppertal.de/~xsc/index_en.html`

- *Boost Interval*,
  `http://www.boost.org/doc/libs/1_53_0/libs/numeric/interval/doc/interval.htm`

- *FILIB++* (Lerch et al., 2001),
  evaluation of basic functions, enables containment sets,
  `http://www2.math.uni-wuppertal.de/wrswt/software/filib.html`

- *GAOL* (by F. Goualard),
  `http://sourceforge.net/projects/gaol/`

- *MPFI (by N. Revol, F. Rouillier)*,
  multi precision, portable `http://perso.ens-lyon.fr/nathalie.revol/software.html`

- *PROFIL / BIAS* (by O. Knüppel et al.),
  efficient, portable, `http://www.ti3.tu-harburg.de/Software/PROFILEnglisch.html`

Concerning other applications and languages,

- *PASCAL-XSC* (Klatte et al., 1991),
  This language provides interval (real and complex) arithmetic, controlled rounding, exact scalar product, and other tools for numerical computation. Moreover, there have been implemented some application modules solving basic numerical problems (linear and nonlinear systems, eigenvalues or optimization).
  `http://www.xsc.de/`, `http://www2.math.uni-wuppertal.de/~xsc/index_en.html`

- *Fortran interval packages*,
  including *INTLIB* (Kearfott et al., 1994), *INTERVAL_ARITHMETIC* (Kearfott, 1996b), *FORTRAN-XSC* (Walter, 1993).

- (Hyvönen and De Pascale, 1996), interval arithmetic embedded in MS Excel,

Software for particular interval problems is listed in .... (CSP, GO) TODO

IEEE standard for interval arithmetic

With the extension of interval packages for many programming languages, there emerged the need of standardization of interval arithmetic. After years of work (2008–2015) of the IEEE Interval Standard Working Group – P1788

`http://grouper.ieee.org/groups/1788/`

there was released *IEEE Std 1788-2015 – IEEE Standard for Interval Arithmetic*; see

`http://standards.ieee.org/findstds/standard/1788-2015.html`

This standard specifies basic interval arithmetic operations, and it provides a layer between the hardware and the programming language levels. A simplified version of this standard was published as *IEEE Std 1788.1-2017 - IEEE Standard for Interval Arithmetic (Simplified)*; see

`https://standards.ieee.org/findstds/standard/1788.1-2017.html`

Citing from the web page: "It includes those operations and features of the latter that in the editors' view are most commonly used in practice. Compared to IEEE Std 1788-2015, this standard aims to be minimalistic, yet to cover much of the functionality needed for interval computations. As such, it is more accessible and will be much easier to implement, and thus will speed up the production of implementations."

For more information about computer arithmetic, its implementation, rounding effects and verified computations see Kulisch (2013); Rump (2010).

Chapter 3

Systems of interval linear equations

Systems of linear equations is a basic problem in linear algebra, and it is also essential for many other disciplines. Not surprisingly, problems related to interval systems of linear equations are fundamental in interval computation.

3.1 Solution set

In this section, we introduce the solution set. We present its topological and other properties, and we discuss the related computational complexity issues.

The solution set. Let $\boldsymbol{A} \in \mathbb{IR}^{m\times n}$ and $\boldsymbol{b} \in \mathbb{IR}^m$. Then the system of interval linear equations is a family of linear systems

$$Ax = b, \quad A \in \boldsymbol{A},\ b \in \boldsymbol{b}.$$

We denote this family in short as

$$\boldsymbol{A}x = \boldsymbol{b},$$

but the aim is *not* to find an (interval) vector x that satisfies these equations. A *solution* is defined as a solution to a system $Ax = b$ for some $A \in \boldsymbol{A}$ and $b \in \boldsymbol{b}$. Naturally, *the solution set* is defined as the set of all solutions and denoted

$$\Sigma \coloneqq \{x \in \mathbb{R}^n;\ \exists A \in \boldsymbol{A}\, \exists b \in \boldsymbol{b} : Ax = b\}.$$

This definition is also called *the united solution set* since we join together all solutions over all real instances of the interval system. Extensions using another quantification are addressed in Section 3.10.

The solution set is characterized by diverse ways. First, we mention the characterization by Beeck.

Theorem 3.1 (Beeck, 1972, 1974)**.** *We have*

(1) $x \in \Sigma$ *if and only if* $\boldsymbol{A}x \cap \boldsymbol{b} \neq \emptyset$,

(2) $x \in \Sigma$ *if and only if* $0 \in \boldsymbol{A}x - \boldsymbol{b}$.

Proof.

(1) When $x \in \Sigma$, then $Ax = b$ for some $A \in \boldsymbol{A}$ and $b \in \boldsymbol{b}$, and thus $Ax = b \in \boldsymbol{A}x \cap \boldsymbol{b}$. On the other hand, when $b \in \boldsymbol{A}x \cap \boldsymbol{b}$, then clearly $b \in \boldsymbol{b}$. By Proposition 2.11, there is $A \in \boldsymbol{A}$ such that $Ax = b$.

(2) This follows from the previous item and Proposition 2.9(4). □

The famous and very useful characterization of Σ comes from Oettli and Prager (1964); see also Fiedler et al. (2006).

Theorem 3.2 (Oettli and Prager, 1964)**.** *The solution set* Σ *is described by the inequality system*

$$|A^c x - b^c| \leq A^\Delta |x| + b^\Delta. \tag{3.1}$$

x is a solution of $\boldsymbol{A}x = \boldsymbol{b}$
$\emptyset \neq \boldsymbol{A}x \cap \boldsymbol{b}$
$0 \in \boldsymbol{A}x - \boldsymbol{b}$
$\lvert A^c x - b^c\rvert \leq A^\Delta \lvert x\rvert + b^\Delta$
$\lvert \operatorname{mid}(\boldsymbol{A}x - \boldsymbol{b})\rvert \leq \operatorname{rad}(\boldsymbol{A}x - \boldsymbol{b})$
$\operatorname{mig}(\boldsymbol{A}x - b^c) \leq b^\Delta$

Table 3.1: Equivalent characterizations of the (weak) solutions.

Proof. By Proposition 3.1, $x \in \Sigma$ if and only if $0 \in \boldsymbol{A}x - \boldsymbol{b}$. By Proposition 2.9(3), it is equivalent with $|\operatorname{mid}(\boldsymbol{A}x - \boldsymbol{b})| \leq \operatorname{rad}(\boldsymbol{A}x - \boldsymbol{b})$, which is simplified to (3.1) by Proposition 2.9(1)–(2). □

As stated in the proof, condition (3.1) can be equivalently stated by using a symmetric formula from Rohn (2014); see Mayer (2017),

$$|\operatorname{mid}(\boldsymbol{A}x - \boldsymbol{b})| \leq \operatorname{rad}(\boldsymbol{A}x - \boldsymbol{b}).$$

Table 3.1 summarizes various ways to characterize the solutions.

When $x \in \Sigma$, it is sometimes useful to know for which realization of intervals the solution is attained, that is, for which $A \in \boldsymbol{A}$ and $b \in \boldsymbol{b}$ we have $Ax = b$. Below, we adopt the result from Fiedler et al. (2006), and in the proof, there is explained the method for finding the corresponding $A \in \boldsymbol{A}$ and $b \in \boldsymbol{b}$.

Proposition 3.3. *Every solution $x \in \Sigma$ is a solution of the special system $A_{yz}x = b_y$ for some $y \in [-1,1]^m$ and $z \in \{\pm 1\}^n$.*

Proof. Let $x \in \mathbb{R}^n$ be a solution to (3.1). Define the vector $y \in [-1,1]^m$ componentwise as

$$y_i = \begin{cases} \frac{(A^c x - b^c)_i}{(A^\Delta |x| + b^\Delta)_i} & \text{if } (A^\Delta |x| + b^\Delta)_i > 0, \\ 1 & \text{otherwise.} \end{cases}$$

Now, we have $(A^c x - b^c)_i = y_i (A^\Delta |x| + b^\Delta)_i$, or,

$$A^c x - b^c = \operatorname{diag}(y)(A^\Delta |x| + b^\Delta).$$

Define $z \coloneqq \operatorname{sgn}(x)$, then $|x| = \operatorname{diag}(z)x$ and we can write

$$A^c x - b^c = \operatorname{diag}(y) A^\Delta \operatorname{diag}(z) x + \operatorname{diag}(y) b^\Delta,$$

or

$$(A^c - \operatorname{diag}(y) A^\Delta \operatorname{diag}(z))x = b^c + \operatorname{diag}(y) b^\Delta.$$

Thus $A_{yz}x = b_y$ and so $x \in \Sigma$. □

The following theorem gives another normal form realization by means of the endpoints of the interval coefficients; for each equation, only one of the endpoints could possibly be in the interior of its interval domain. For a survey of normal forms see (Rohn, 2012b, Sec. 4.2.3).

Proposition 3.4 (Rohn, 1985)**.** *Let $x^* \in \Sigma$. Then there are $A \in \boldsymbol{A}$ and $b \in \boldsymbol{b}$ such that $Ax^* = b$, and for each $i \in \{1, \ldots, m\}$ we have $a_{ij} \in \{\underline{a}_{ij}, \overline{a}_{ij}\}$ and $b_i \in \{\underline{b}_i, \overline{b}_i\}$ for all but at most one entry.*

Proof. Let $i \in \{1, \ldots, m\}$ and denote the ith equation of the system $\boldsymbol{A}x = \boldsymbol{b}$ by $\boldsymbol{a}^T x = \boldsymbol{\beta}$. By the Oettli–Prager theorem

$$(a^c)^T x^* - \beta^c \leq (a^\Delta)^T |x^*| + \beta^\Delta,$$
$$-(a^c)^T x^* + \beta^c \leq (a^\Delta)^T |x^*| + \beta^\Delta.$$

Denote $s \coloneqq \operatorname{sgn}(x^*)$. Then $|x^*| = \operatorname{diag}(s)x^*$ and the above inequalities draw

$$(a^c - \operatorname{diag}(s)a^\Delta)^T x^* \le \overline{\beta},$$
$$(a^c + \operatorname{diag}(s)a^\Delta)^T x^* \ge \underline{\beta}.$$

Denote $r \coloneqq s$ and $a_r \coloneqq a^c - \operatorname{diag}(r)a^\Delta$. Put $j \coloneqq 0$. While $a_r^T x^* < \overline{\beta}$ do $j \coloneqq j+1$ and $r_j \coloneqq -r_j$. Now, there are two cases. If we have $a_r^T x^* \ge \overline{\beta}$, then obviously $a_r^T x^* = \beta$ for some $r_j \in [-1,1]$. Thus, we found a realization in which only the jth coefficient is possibly from the interior of the interval $\boldsymbol{a}_j$. In the second case, $r = -s$ and

$$\underline{\beta} \le (a^c + \operatorname{diag}(s)a^\Delta)^T x^* = a_r^T x^* < \overline{\beta}.$$

Hence $(a^c + \operatorname{diag}(s)a^\Delta)^T x^* = \beta$ for some $\beta \in \boldsymbol{\beta}$. So we found a realization in which only the right-hand side entry is possibly from the interior of $\boldsymbol{\beta}$. □

Geometrically, the solution set Σ has a polyhedral shape. It might be nonconvex, but it is provably convex when restricted to any orthant (Beeck, 1973; Oettli, 1965).

Theorem 3.5 (Oettli, 1965)**.** *In each orthant, Σ is either empty or a convex polyhedral set.*

Proof. Let an orthant corresponding to a sign vector $s \in \{\pm 1\}^n$ be given. This orthant is described by $\operatorname{diag}(s)x \ge 0$, and the solution set restricted to this orthant reads

$$|A^c x - b^c| \le A^\Delta |x| + b^\Delta,\ \operatorname{diag}(s)x \ge 0.$$

Since $|x| = \operatorname{diag}(s)x$, we have

$$|A^c x - b^c| \le A^\Delta \operatorname{diag}(s)x + b^\Delta,\ \operatorname{diag}(s)x \ge 0,$$

or

$$A^c x - b^c \le A^\Delta \operatorname{diag}(s)x + b^\Delta,\ -(A^c x - b^c) \le A^\Delta \operatorname{diag}(s)x + b^\Delta,\ \operatorname{diag}(s)x \ge 0.$$

By rearranging the expression, we obtain linear inequalities

$$(A^c - A^\Delta \operatorname{diag}(s))x \le \overline{b},\ (-A^c - A^\Delta \operatorname{diag}(s))x \le -\underline{b},\ \operatorname{diag}(s)x \ge 0. \tag{3.2}$$

or in compact form

$$A_{es}x \le \overline{b},\ -A_{-es}x \le -\underline{b},\ \operatorname{diag}(s)x \ge 0.$$ □

The proof revealed that the solution set can be partitioned as a union of convex polyhedra, the interior of which is pairwise disjoint.

Corollary 3.6. *We have*

$$\Sigma = \bigcup_{s \in \{\pm 1\}^n} \left\{ x;\ A_{es}x \le \overline{b},\ -A_{-es}x \le -\underline{b},\ \operatorname{diag}(s)x \ge 0 \right\}. \tag{3.3}$$

In particular, if we are interested in the nonnegative solutions, then the solution set is a convex polyhedron and checking solvability is polynomial by linear programming.

Corollary 3.7. *The solution set to the interval system $\boldsymbol{A}x = \boldsymbol{b}$, $x \ge 0$ is described by*

$$\underline{A}x \le \overline{b},\ \overline{A}x \ge \underline{b},\ x \ge 0. \tag{3.4}$$

Proof. Put $s \coloneqq e$ and by (3.2) we have that the solution set in the orthant $x \ge 0$ draws

$$(A^c - A^\Delta)x \le \overline{b},\ (-A^c - A^\Delta)x \le -\underline{b},\ x \ge 0, \tag{3.5}$$

which is equivalent to (3.4). □

Example 3.8. Consider the interval system of linear equations (Barth and Nuding, 1974)

$$\begin{pmatrix} [2,4] & [-2,1] \\ [-1,2] & [2,4] \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} [-2,2] \\ [-2,2] \end{pmatrix}. \tag{3.6}$$

Its solution set is illustrated in Figure 3.1. □

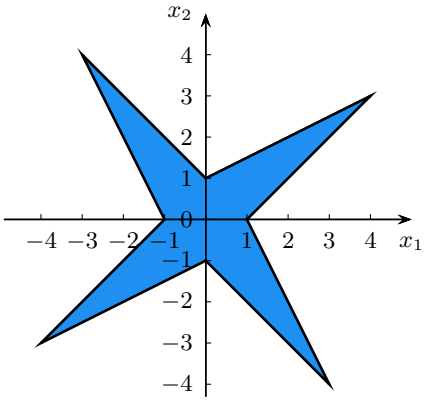


Figure 3.1: (Example 3.8) The solution set to (3.6).

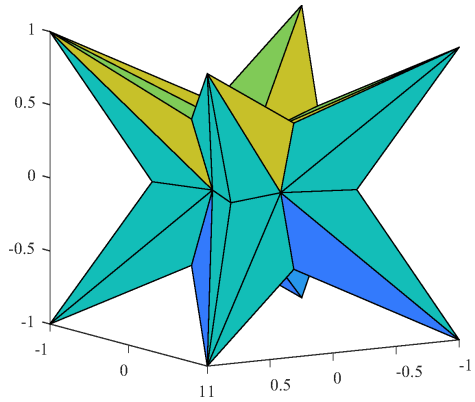


Figure 3.2: (Example 3.9) The solution set to (3.7).

Example 3.9. The solution set for this three-dimensional system

$$\begin{pmatrix} [3,5] & [-1,1] & [-1,1] \\ [-1,1] & [3,5] & [-1,1] \\ [-1,1] & [-1,1] & [3,5] \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} [-1,1] \\ [-1,1] \\ [-1,1] \end{pmatrix} \tag{3.7}$$

is illustrated in Figure 3.2. □

Remark 3.10. The solution set can also be decomposed as follows

$$\Sigma = \bigcup_{s\in\{\pm 1\}^n} \{x;\ A_{es}x \le \overline{b},\ -A_{-es}x \le -\underline{b}\} \tag{3.8}$$

which is an alternative to (3.3). Inclusion "$\subseteq$" is easy to see in view of (3.3). Inclusion "$\supseteq$" can be shown as in the proof of Theorem 3.5: The system

$$A_{es}x \le \overline{b},\quad -A_{-es}x \le -\underline{b}$$

is equivalent to

$$A^c x - b^c \le A^\Delta \operatorname{diag}(s)x + b^\Delta,\ -(A^c x - b^c) \le A^\Delta \operatorname{diag}(s)x + b^\Delta,$$

from which

$$|A^c x - b^c| \le A^\Delta \operatorname{diag}(s)x + b^\Delta \le A^\Delta |x| + b^\Delta.$$

Decomposition (3.8) does not provide a pairwise disjoint interior splitting of the solution set. On the other hand, the particular convex polyhedra are larger.

Example 3.11. Consider the interval linear system from Example 3.8. Figure 3.3 illustrates overlapping decomposition (3.8). □

The problem of finding a solution $x \in \Sigma$ is intractable. The first proof is by Lakeev and Noskov (1994); we follow the proof by Fiedler et al. (2006).

Theorem 3.12 (Lakeev and Noskov, 1994)**.** *Checking $\Sigma \neq \emptyset$ is an NP-hard problem, even in the class of problems with $m = n$.*

Proof. Let the system (1.3) from Theorem 1.35 be given and consider the interval system

$$[A, A]x = [-e, e],\ [-e, e]^T x = [1, 1].$$

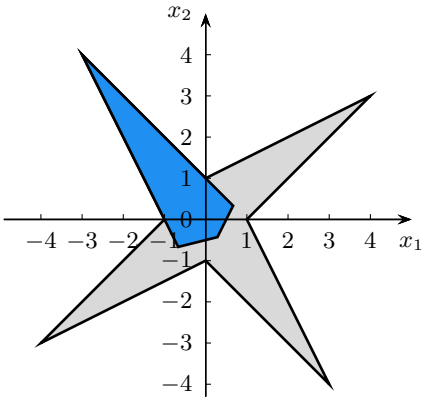
According to the Oettli–Prager Theorem 3.2, the solution set is described by

$$|Ax - 0| \le 0|x| + e,\ |0x - 1| \le e^T|x| + 0,$$

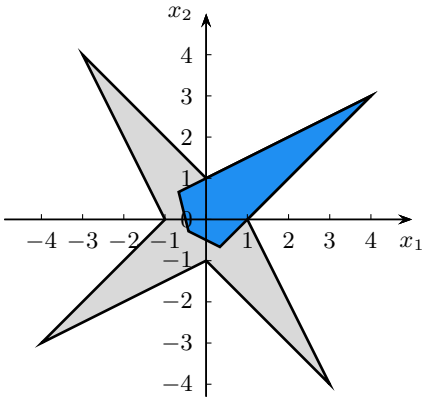
which is equivalent to (1.3). To obtain a square system, we impose an idle variable y resulting in an equivalent square interval system

$$[A, A]x + 0y = [-e, e],\ [-e, e]^T x + 0y = [1, 1].$$

□



(a) The part of the solution set corresponding to $s = (-1, 1)^T$ is in blue.



(b) The part of the solution set corresponding to $s = (1, 1)^T$ is in blue.

Figure 3.3: (Example 3.11): Overlapping decomposition (3.8) of the solution set.

Remark 3.13 (Polynomial cases)**.** On the other hand, there are classes of interval equations solvable in polynomial time. First, problems with nonnegative variables as mentioned in Theorem 3.7.

Second, problems with real A and intervals in the right-hand side vector $\boldsymbol{b}$ only. In this case, the solution set Σ is described by linear inequalities $\underline{b} \leq Ax \leq \overline{b}$. Geometrically, it represents an image of the box $\boldsymbol{b}$ under the mapping $b \mapsto A^{-1}b$, that is, a zonotope (see page 33). In this case, the interval hull of the solution set is simply calculated by interval arithmetic, that is, $\square\Sigma = A^{-1}\boldsymbol{b}$.

Third, suppose that the intervals are situated $\boldsymbol{b}$ and in at most k columns of $\boldsymbol{A}$. Then in (3.3), we do not need to enumerate all 2^n sign vectors, but only 2^k, where the variable signs correspond to interval columns of $\boldsymbol{A}$. For the sake of simplicity, suppose that the last k columns of $\boldsymbol{A}$ are interval. Then (3.3) takes the form

$$\Sigma = \cup_{s\in\{1\}^{n-k}\times\{\pm1\}^k} \left\{x;\ A_{es}x \leq \overline{b},\ -A_{-es}x \leq -\underline{b},\ \operatorname{diag}(s)x \geq 0\right\}.$$

Other complexity results related to the solution set are discussed at the end of Section 3.6.2.

Remark 3.14. There are other interesting topological properties of the solution set. For instance, if $\boldsymbol{A}$ is regular and $\boldsymbol{b}$ is real and nonzero, then the solution set Σ cannot simultaneously intersect two opposite orthants (Rohn, 1984, 2012b).

There are some open problems related topology of the solution set, too. Rohn (2012b, Thm. 144) shows that when $\boldsymbol{A}$ is regular, every two points in Σ can be connected by a piecewise linear curve having at most n segments. Can this value decrease to 2? Or even stronger version: Exists $x^* \in \Sigma$ such that for each $x \in \Sigma$ the whole segment joining x^* and x lies in Σ?

Notes. Visualization of the solution set is possible by using diverse programs: MATLAB package for visualization in 2D and 3D is available in (Sharaya, 2014b), and an online service for computations webComputing in Popova (2006a).

Chen et al. (2010) applied interval linear equations in static code analysis, by introducing a new abstract domain. Its benefit is that, based on the properties of interval linear equations, it can capture certain noncovex properties.

3.2 Methods for the square case – preliminaries

In this section, we restrict ourselves to the most common square case of interval systems with $m = n$. Sections 3.3.1 and 3.5.1 are devoted to direct methods and Section 3.4 is devoted to iterative methods for enclosing the solution set. Methods to compute the exact interval hull $\square\Sigma$ are studied in Section 3.5.4. Special cases that are easily solved are dealt with in Section 3.5.5.

The structure od the solution set Σ is rather complicated, so one is more interested in determining its interval hull $\square\Sigma$. By Theorem 3.12, computation $\square\Sigma$ is *NP*-hard (if we know $\square\Sigma$, then we can easily

decide whether $\Sigma \neq \emptyset$), and this is also true for a family of systems with regular matrices satisfying $\rho(|(A^c)^{-1}|A^\Delta) = 0$; see (Fiedler et al., 2006, Thm. 2.38) or (Rohn, 2012b, Thm. 167) and compare Section 3.6.2. That is why one is usually satisfied with a tight enclosure. Nevertheless, even calculation of a sharp enclosure may be hard. As shown in Kreinovich and Lakeyev (1996) (cf. Kreinovich et al. (1998)), computing an enclosure with a prescribed relative or absolute accuracy is still *NP*-hard.

Problem statement. The main problem of Sections 3.2–3.5 is to find $\square\Sigma$ or its tight enclosure.

In the rest of Section 3.2, we will discuss several classes of interval matrices and two auxiliary transformations (preconditioning and the residual form) that transform interval systems to a (usually) more convenient form.

Regularity and special matrices. As in the real case, solvability of interval systems and the structure of Σ depends on whether the constraint matrices are nonsingular. This leads to regularity of interval matrices (regularity is dealt with in more detail in Section 3.6).

Definition 3.15. An interval matrix $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$ is called *regular* if each $A \in \boldsymbol{A}$ is nonsingular. Otherwise, $\boldsymbol{A}$ is *irregular*.

Two important classes of regular interval matrices are M-matrices and its generalization H-matrices.

Definition 3.16. An interval matrix $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$ is called *M-matrix* if each $A \in \boldsymbol{A}$ is an M-matrix.

Theorem 3.17 (Barth and Nuding, 1974)**.** *An interval matrix $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$ is an M-matrix if and only if $\underline{A}$ is an M-matrix and $\overline{A}_{ij} \leq 0$ for all $i \neq j$.*

Proof. One direction is obvious. For the converse one we employ Theorem 1.18(2). Since $\underline{A}$ is an M-matrix, there is $v > 0$ such that $\underline{A}v > 0$. Now, for any $A \in \boldsymbol{A}$ we have $Av \geq \underline{A}v > 0$. □

In other words, we can say that $\boldsymbol{A}$ is an M-matrix if and only if both the lower and the upper bound matrices $\underline{A}$ and $\overline{A}$ are M-matrices.

Definition 3.18. An interval matrix $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$ is called *H-matrix* if each $A \in \boldsymbol{A}$ is an H-matrix.

Theorem 3.19 (Neumaier, 1984)**.** *An interval matrix $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$ is an H-matrix if and only if $\langle \boldsymbol{A}\rangle$ is an M-matrix, where*

$$\begin{aligned}\langle \boldsymbol{A}\rangle_{ii} &= \operatorname{mig}(\boldsymbol{a}_{ii}),\\ \langle \boldsymbol{A}\rangle_{ij} &= -\operatorname{mag}(\boldsymbol{a}_{ii}), \quad i \neq j.\end{aligned}$$

Proof. "Only if". Since $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$ is an H-matrix, $\langle A\rangle$ is an M-matrix for each $A \in \boldsymbol{A}$, so even for those $A^* \in \boldsymbol{A}$ satisfying $\langle A^*\rangle = \langle \boldsymbol{A}\rangle$.

"If". Let $v > 0$ be such that $\langle \boldsymbol{A}\rangle v > 0$. Then for each $A \in \boldsymbol{A}$ we have $\langle A\rangle v \geq \langle \boldsymbol{A}\rangle v > 0$. □

The proof also showed that $\boldsymbol{A}$ is an H-matrix if and only if the H-matrix property holds for the selection $A^* \in \boldsymbol{A}$ defined as

$$\begin{aligned}a^*_{ii} &= \arg\min\{|a_{ii}|;\ a_{ii} \in \boldsymbol{a}_{ii}\},\\ a^*_{ij} &= \arg\max\{|a_{ij}|;\ a_{ij} \in \boldsymbol{a}_{ij}\}, \quad i \neq j.\end{aligned}$$

3.2.1 Preconditioning

Usually, a so-called *preconditioning* is performed in order to obtain tight enclosures; the origin of this approach dates back to Hansen (1965); Hansen and Smith (1967). Let $C \in \mathbb{R}^{n\times n}$, then the preconditioning the interval system $\boldsymbol{A}x = \boldsymbol{b}$ means that we multiply both sides by C to obtain a new interval system

$$(C\boldsymbol{A})x = C\boldsymbol{b},$$

where the interval matrix $C\boldsymbol{A}$ and the interval vector $C\boldsymbol{b}$ are evaluated by interval arithmetic. Due to the virtue of interval arithmetic, the solution set to the new system contains the old one. It may seem that this also leads to overestimating the enclosures of the solution set calculated by various methods. However, most methods work well with preconditioning.

After preconditioning, the constraint matrix reads $C\boldsymbol{A} = [CA^c - |C|A^\Delta, CA^c + |C|A^\Delta]$, that is, its midpoint is CA^c and its radius is $|C|A^\Delta$. Usually, one preconditions by $C = (A^c)^{-1}$ or its numerical approximation. The reason that preconditioning helps is that the center of $C\boldsymbol{A}$ is the identity matrix, and for such cases the interval hull of the solution set can be exactly determined (Section 3.5.1). Moreover, the inverse midpoint preconditioner behaves best in some measures (Neumaier, 1984, 1990), which, however, does not mean that no other choice can be more useful in some cases.

Preconditioning is successfully performed for all the methods presented here, including the interval Gaussian elimination, the interval Gauss–Seidel method and the Krawczyk method.

One can hardly compute $C = (A^c)^{-1}$ exactly by using floating point arithmetic. However, some results later on (Theorem 3.43 and many others) rely on having $\operatorname{mid}(C\boldsymbol{A}) = I_n$. This can be achieved by relaxing $C\boldsymbol{A}$ and enclosing it as $C\boldsymbol{A} \subseteq [I_n - \Delta, I_n + \Delta]$, where $\Delta \coloneqq \operatorname{mag}(I_n - C\boldsymbol{A})$.

Proposition 3.20. *The matrix $[I_n - \Delta, I_n + \Delta]$ is the smallest matrix (with respect to inclusion) enclosing $C\boldsymbol{A}$ and having its midpoint to be I_n.*

Proof. Let $\boldsymbol{M}$ be such that $C\boldsymbol{A} \subseteq \boldsymbol{M}$ and $M^c = I_n$. Then $\operatorname{mag}(I_n - C\boldsymbol{A}) \leq \operatorname{mag}(I_n - \boldsymbol{M}) = M^\Delta$, so $M^\Delta \coloneqq \Delta$ is the best choice. □

A preconditioning that is optimal in some sense was proposed by Kearfott (1990); Kearfott et al. (1991); Kearfott (1996a) for the interval Gauss–Seidel method. The preconditioning is done separately for each variable by solving an appropriate linear program. Another preconditioners, which are based on the Gauss–Jordan elimination, are discussed in Montanher et al. (2017).

Example 3.21. Reconsider Example 3.8 by (Barth and Nuding, 1974) again. Preconditioning the interval system by $(A^c)^{-1}$, we obtain an interval system

$$\begin{pmatrix} [0.5945, 1.4055] & [-0.5406, 0.5406] \\ [-0.5406, 0.5406] & [0.5945, 1.4055] \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} [-0.7568, 0.7568] \\ [-0.7568, 0.7568] \end{pmatrix}.$$

In this example, the solution set blowed up a lot. Figure 3.4 illustrates the original solution set (in blue), the solution set of the preconditioned system (in gray) and its interval hull in light gray. One can see that the preconditioning caused high overestimation.

Nevertheless, in a typical situation, the overestimation is mild. Consider, for example, the system

$$\begin{pmatrix} [4, 5] & [-1, 1] \\ [1, 2] & -[3, 5] \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} [-1, 4] \\ -[3, 4] \end{pmatrix}. \tag{3.9}$$

Its solution set and the preconditioned one are drawn in Figure 3.5. □

3.2.2 Residual form

The residual form approach to solving $\boldsymbol{A}x = \boldsymbol{b}$ employs a vector $x^* \in \mathbb{R}^n$, and by using the substitution $y \equiv x - x^*$, it transforms the problem to solving

$$\boldsymbol{A}y = \boldsymbol{b} - \boldsymbol{A}x^*,$$

where the right-hand side is evaluated by interval arithmetic. This is again an interval linear system of equations, and can be solved by any solver. If $\boldsymbol{y} \in \mathbb{IR}^n$ is an enclosure to the solution set of the transformed system, then $x^* + \boldsymbol{y}$ encloses the original one.

The vector x^* is usually taken as a solution to $A^c x = b^c$, and therefore the midpoints of both interval vectors $\boldsymbol{b} - \boldsymbol{A}x^*$ and $\boldsymbol{y}$ are near to zero.

The transformation to the residual form of interval equations usually does not tighten the resulting enclosures and the convergence speed, but it has some good properties. This form is convenient, in particular when solving systems with dependencies between interval coefficients; see Section 3.9.

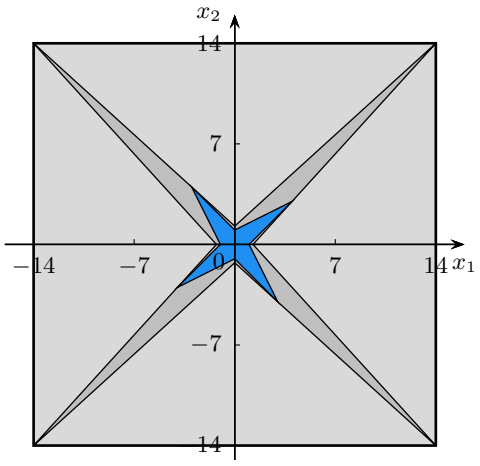


Figure 3.4: The solution set and the preconditioned one for (3.6).

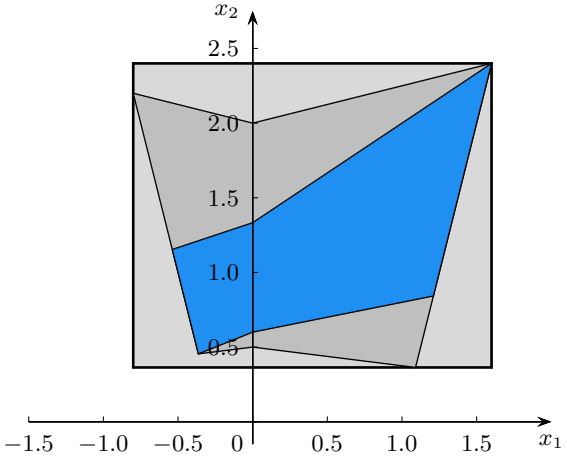


Figure 3.5: The solution set and the preconditioned one for (3.9).

3.3 Direct methods for the square case

In this section, we briefly present the interval version of the well-known Gaussian elimination. Other direct methods based on the limit versions of iterative methods will be discussed in Section 3.4.4*.

3.3.1 Interval Gaussian elimination

The interval Gaussian elimination works in the same fashion as Gaussian elimination, but with interval arithmetic instead of the real (or floating point) one. For more details see, e.g., Alefeld and Herzberger (1983); Neumaier (1990); Mayer (2017), from which the results of this section are derived.

Notice that in each step, the pivot must be an interval not containing the zero. If we cannot find an appropriate pivot, then the method fails.

The elements below the pivot can be set directly to zero as in the classical Gaussian elimination and need not be superfluously evaluated by interval arithmetic. The reason is that the transformation of the interval matrix

$$\begin{pmatrix} \boldsymbol{a}_{11} & \dots \\ \boldsymbol{a}_{21} & \dots \\ \vdots & \end{pmatrix}$$

changes the $(2,1)$th entry to $\boldsymbol{a}_{21} - \frac{\boldsymbol{a}_{21}}{\boldsymbol{a}_{11}}\boldsymbol{a}_{11}$. However, a straightforward evaluation of this expression is not convenient. As we observed in Example 2.7, among others, it is better to use another formula for the same function. Since $a_{21} - \frac{a_{21}}{a_{11}}a_{11} = 0$, we simply evaluate the zero function on $\boldsymbol{a}_{11} \times \boldsymbol{a}_{21}$, which is again zero.

The algorithm (interval Gaussian elimination). Now we state formally the algorithm. First, we transform the extended interval matrix $\boldsymbol{C} \coloneqq (\boldsymbol{A} \mid \boldsymbol{b})$ into the row echelon form:

1: **for** $k = 1, \dots, n$ **do**
2: suppose that $0 \notin \boldsymbol{c}_{kk}$ (otherwise switch row k with some row below it)
3: **for** $i = k+1, \dots, n$ **do**
4: $\boldsymbol{c}_{ik} \coloneqq 0$
5: **for** $j = k+1, \dots, n$ **do**
6: $\boldsymbol{c}_{ij} \coloneqq \boldsymbol{c}_{ij} - \frac{\boldsymbol{c}_{ik}}{\boldsymbol{c}_{kk}}\boldsymbol{c}_{kj}$
7: **end for**
8: **end for**
9: **end for**

If step 2 cannot be performed, then the algorithm cannot compute a bounded enclosure of Σ, and $\boldsymbol{A}$ is possibly irregular. Otherwise, the algorithm succeeds, and we apply the backward substitution to compute an enclosure $\boldsymbol{x} \supseteq \Sigma$:

1: **for** $k = n, \dots, 1$ **do**
2: $\quad \boldsymbol{x}_k \coloneqq \boldsymbol{b}_k - \frac{1}{\boldsymbol{a}_{kk}} \sum_{j=k+1}^{n} \boldsymbol{a}_{kj} \boldsymbol{x}_j$
3: **end for**
4: **return** "$\boldsymbol{x}$"

Example 3.22. Reconsider Example 3.8. The interval Gaussian elimination proceeds as follows

$$\begin{pmatrix} \boldsymbol{A} & \boldsymbol{b} \end{pmatrix} \sim \begin{pmatrix} [2,4] & [-2,1] & [-2,2] \\ [-1,2] & [2,4] & [-2,2] \end{pmatrix} \sim \begin{pmatrix} [2,4] & [-2,1] & [-2,2] \\ 0 & [1,6] & [-4,4] \end{pmatrix}.$$

Notice that after pivoting operations, the values below the pivots vanish, as in the classical Gaussian elimination. By back substitution, we compute

$$\begin{aligned} \boldsymbol{x}_2 &= [-4,4], \\ \boldsymbol{x}_1 &= \big([-2,2] - [-2,1] \cdot [-4,4]\big) \,/\, [2,4] = [-5,5]. \end{aligned}$$

The resulting enclosure of the solution set is $\boldsymbol{x} = ([-5,5],\, [-4,4])^T$. □

The interval Gaussian elimination can fail even if $\boldsymbol{A}$ is regular (one pre-explanation is that checking regularity of $\boldsymbol{A}$ is hard, but the interval Gaussian elimination is a polynomial method). This is not the case for $n = 1$ or $n = 2$, but one can to find a three-dimensional counterexample; see Mayer (2017); Reichmann (1979). We present another counterexample here.

Example 3.23. Consider the interval matrix

$$\boldsymbol{A} = \begin{pmatrix} [1,2] & 2 & 3 \\ 4 & 5 & 6 \\ -1 & 1 & 0 \end{pmatrix}.$$

It is not hard to verify that $\boldsymbol{A}$ is regular. Nevertheless, the transformation to the row echelon form yields

$$\begin{pmatrix} [1,2] & 2 & 3 \\ 4 & 5 & 6 \\ -1 & 1 & 0 \end{pmatrix} \sim \begin{pmatrix} [1,2] & 2 & 3 \\ 0 & [-3,1] & [-6,0] \\ 0 & [2,3] & [1.5,3] \end{pmatrix} \sim \begin{pmatrix} [1,2] & 2 & 3 \\ 0 & [2,3] & [1.5,3] \\ 0 & [-3,1] & [-6,0] \end{pmatrix}$$
$$\sim \begin{pmatrix} [1,2] & 2 & 3 \\ 0 & [2,3] & [1.5,3] \\ 0 & 0 & [-7.5,4.5] \end{pmatrix}.$$

The third pivot is an interval containing the zero, so the interval Gaussian elimination cannot be performed regardless the values of the right-hand side interval vector $\boldsymbol{b}$. □

Due to this property, it is natural to seek for a class of regular matrices for which the interval Gaussian elimination does not fail. Such classes are M-matrices or H-matrices, among some others; we prove this result for M-matrices first, and then we generalize it for H-matrices.

Theorem 3.24 (Alefeld, 1977)**.** *Let $\boldsymbol{A}$ be an M-matrix. Then for each $k = 1, \dots, n-1$, after the kth loop of the interval Gaussian elimination, the submatrix formed by the last $n-k$ rows and columns is an M-matrix.*

Proof. We show it for $k = 1$ only; the others follow by induction. Let $v > 0$ such that $\underline{A}v > 0$ and let $\boldsymbol{A}'$ be the matrix after the first loop. First, observe that $\boldsymbol{a}'_{ij} = \boldsymbol{a}_{ij} - \boldsymbol{a}_{1j}\boldsymbol{a}_{i1}/\boldsymbol{a}_{11} \le 0$ for $i \ne j$. Now, for $i = 2, \dots, n$, using $\sum_{j=1}^{n} \underline{a}_{1j} v_j \ge 0$ and $\sum_{j=1}^{n} \underline{a}_{ij} v_j > 0$ we get

$$\begin{aligned} \sum_{j=2}^{n} \underline{a}'_{ij} v_j &= \sum_{j=2}^{n} \left(\underline{a}_{ij} - \underline{a}_{1j} \frac{\underline{a}_{i1}}{\underline{a}_{11}} \right) v_j = \sum_{j=2}^{n} \underline{a}_{ij} v_j - \frac{\underline{a}_{i1}}{\underline{a}_{11}} \sum_{j=2}^{n} \underline{a}_{1j} v_j \\ &\ge \sum_{j=2}^{n} \underline{a}_{ij} v_j + \frac{\underline{a}_{i1}}{\underline{a}_{11}} \underline{a}_{11} v_1 = \sum_{j=1}^{n} \underline{a}_{ij} v_j > 0, \end{aligned}$$

which means that $\boldsymbol{A}'$ without the first row and column is an M-matrix. □

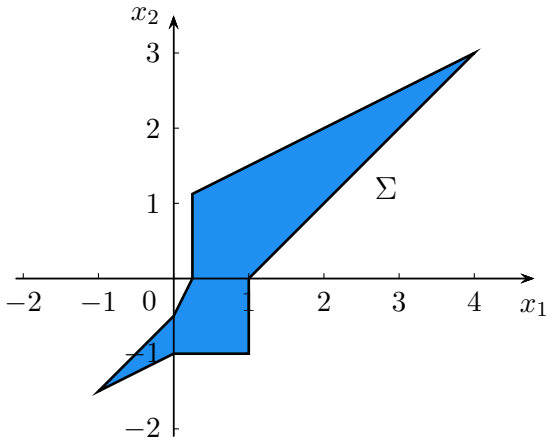


Figure 3.6: (Example 3.26) The solution set Σ.

Corollary 3.25 (Alefeld, 1977)**.** *Let $\boldsymbol{A}$ be an M-matrix. Then the interval Gaussian elimination can be carried out without any pivoting.*

Proof. From the assumption, there is $v > 0$ such that $\underline{A}v > 0$. In particular, $\underline{a}_{11}v_1 > -\sum_{i=2}^{n} \underline{a}_{1i}v_i \geq 0$. Thus, $\boldsymbol{a}_{11}$ is positive, and one loop of the interval Gaussian elimination can be performed. Since the active submatrix is an M-matrix by Theorem 3.24, this argument applies repeatedly in each loop. □

Even though the interval Gaussian elimination does not fail for M-matrices, it need not produce the interval hull of the solution set.

Example 3.26. Consider the interval linear system of equations $\boldsymbol{A}x = \boldsymbol{b}$ from (Mayer, 2017, Ex. 5.5.19), where

$$\boldsymbol{A} = \begin{pmatrix} [2,4] & [-2,0] \\ [-1,0] & [2,4] \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} [1,2] \\ [-2,2] \end{pmatrix}.$$

The solution set Σ is depicted in Figure 3.6. Its interval hull is

$$\square\Sigma = ([-1,4],\, [-1.5,3])^T.$$

Even though $\boldsymbol{A}$ is an M-matrix, the interval Gaussian elimination yields an overestimated enclosure

$$([-1.5,4],\, [-2,3])^T. \qquad \square$$

Theorem 3.27 (Alefeld, 1977)**.** *Let $\boldsymbol{A}$ be an H-matrix. Then for each $k = 1, \ldots, n-1$, after the kth loop of the interval Gaussian elimination, the submatrix formed by the last $n-k$ rows and columns is an H-matrix.*

Corollary 3.28 (Alefeld, 1977)**.** *Let $\boldsymbol{A}$ be an H-matrix. Then the interval Gaussian elimination can be carried out without any pivoting.*

Proof. By Theorem 3.19, $\langle \boldsymbol{A} \rangle$ is an H-matrix, so $0 \notin \boldsymbol{a}_{11}$, and the first loop of the interval Gaussian elimination can be performed. Then we proceed by induction. □

Gaussian elimination closely relates to LU decomposition. For H-matrices, we have the following interval version of LU decomposition.

Theorem 3.29 (Alefeld, 1977)**.** *Let $\boldsymbol{A}$ be an H-matrix. Then there are lower and upper triangular interval matrices $\boldsymbol{L}, \boldsymbol{U} \in \mathbb{IR}^{n \times n}$ such that the diagonal of $\boldsymbol{L}$ consists of ones, and $\boldsymbol{A} \subseteq \boldsymbol{LU}$.*

Proof. For an H-matrix $A \in \boldsymbol{A}$, the LU decomposition can be done as follows: Transform the matrix $(A \mid I_n)$ to the row echelon form $(U \mid L')$, where U is upper triangular and L' is lower triangular with ones on the diagonal. Hence $L'A = U$, from which we obtain the LU decomposition as $A = LU$, where $L = L'^{-1}$.

We proceed similarly for the interval matrix $\boldsymbol{A}$. By Corollary 3.28, we can transform the interval matrix $(\boldsymbol{A} \mid I_n)$ to the row echelon form $(\boldsymbol{U} \mid \boldsymbol{L}')$. Since $\boldsymbol{L}'$ is lower triangular with ones on the diagonal, we can enclose the inverses of all its realizations by transforming $(\boldsymbol{L}' \mid I_n)$ to the reduced row echelon form $(I_n \mid \boldsymbol{L})$, which always exists. By the inclusion property of interval arithmetic, $\boldsymbol{A} \subseteq \boldsymbol{LU}$. □

To summarize, provided $\boldsymbol{A}$ is an H-matrix or a similar special matrix (such as diagonally dominant), the interval Gaussian works well. Moreover, when $\boldsymbol{A}$ is an M-matrix, and one of $0 \in \boldsymbol{b}$, $\underline{b} \geq 0$, or $\overline{b} \leq 0$ holds true, then the interval Gaussian elimination yields $\square\Sigma$; see Barth and Nuding (1974); Beeck (1974); Mayer (2017). This result is worth comparing with Theorem 3.57, where an analogous result for inverse nonnegative interval matrices is stated.

Preconditioning. In general, preconditioning is commonly recommended. Without a preconditioning, the interval Gaussian elimination may yield very poor results; for an exponential overestimation see Neumaier (1990). For systems preconditioned by $(A^c)^{-1}$ the method works well in the sense that we get a finite enclosure provided the constraint matrix stays regular (since the interval matrix is an H-matrix then). Moreover, it is superior to the interval Gauss–Seidel method (Section 3.4.1) in this case. Under some assumptions, it also yields the tightest enclosure of the preconditioned system (Mayer and Rohn, 1998). On the other hand, we can find examples for which the direct Gaussian elimination works, but the midpoint inverse preconditioning fails.

Notes and further reading. Attempts to make the interval Gaussian elimination more efficient involve, e.g., pivot tightening (Garloff, 2009). Mayer (2017) discusses several classes of interval matrices (i.a. Hessenberg matrices) and conditions under which interval Gaussian elimination can be applied. He also characterizes a class of interval matrices (including, e.g., tridiagonal matrices) for which interval Gaussian elimination is feasible if and only if it is feasible for each realization.

3.4 Iterative methods for the square case

Several methods for finding an enclosure to Σ, such as the interval Jacobi, interval Gauss–Seidel and Krawczyk method, are based on iterative steps sequentially improving an initial enclosure $\boldsymbol{x}^0$. The iteration methods that we will present do not require regularity of the matrix, which is convenient when searching for the solutions in a prescribed box only; cf. Section 8.1. The methods presented in this section mostly perform better if preconditioning is used.

Initial enclosure. The following formula gives a cheap initial enclosure; see Moore et al. (2009). Again, it performs well if the interval linear system $\boldsymbol{A}x = \boldsymbol{b}$ is preconditioned by the midpoint inverse.

Theorem 3.30. *If* $\|I_n - \boldsymbol{A}\|_\infty < 1$*, then*

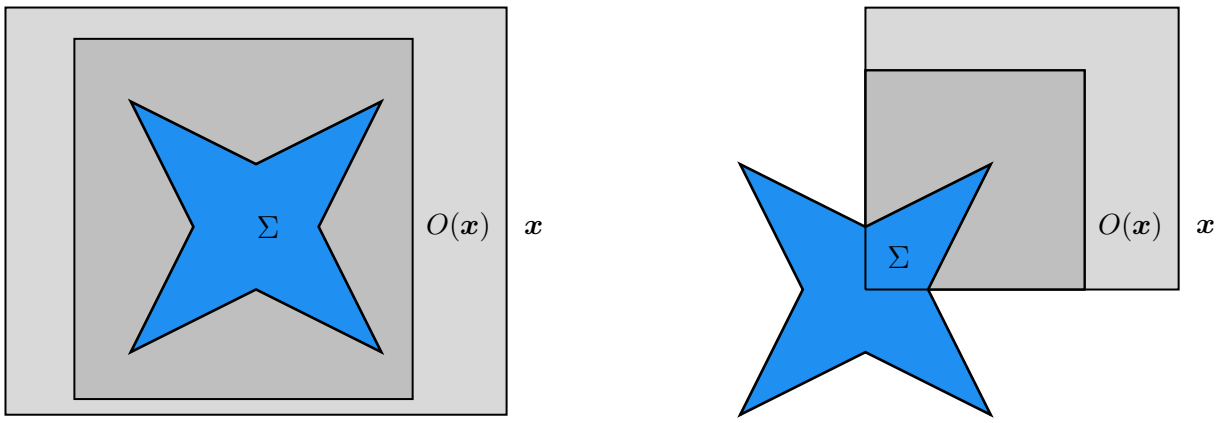
$$\boldsymbol{x}^0 := \frac{\|\boldsymbol{b}\|_\infty}{1 - \|I_n - \boldsymbol{A}\|_\infty}[-1, 1]^n \tag{3.10}$$

encloses Σ*.*

Proof. Let $x \in \Sigma$, so $Ax = b$ for some $A \in \boldsymbol{A}$ and $b \in \boldsymbol{b}$. Thus $x = Ax + (I_n - A)x = b + (I_n - A)x$, whence for any matrix norm induced by a vector norm, one has $\|x\| = \|b + (I_n - A)x\| \leq \|b\| + \|I_n - A\| \cdot \|x\|$. Hence by factoring $\|x\|$

$$\|x\| \leq \frac{\|b\|}{1 - \|I_n - A\|} \leq \frac{\|\boldsymbol{b}\|}{1 - \|I_n - \boldsymbol{A}\|}.$$

Choosing the maximum norm we obtain (3.10). □



(a) The initial vector $\boldsymbol{x}$ encloses the whole solution set Σ.

(b) The initial vector $\boldsymbol{x}$ encloses a part of the solution set Σ.

Figure 3.7: An interval operator $O(\boldsymbol{x})$ applied on an interval vector $\boldsymbol{x}$.

Interval operators. After having an initial enclosure to Σ, the iterations can usually be written in a generic way as follows

1: $\boldsymbol{x}^k \coloneqq O(\boldsymbol{x}^{k-1})$;
2: $\boldsymbol{x}^k \coloneqq \boldsymbol{x}^k \cap \boldsymbol{x}^{k-1}$

Herein, $O\colon \mathbb{IR}^n \to \mathbb{IR}^n$ is an interval operator. This operator must not miss any solution. That is, if $x \in \boldsymbol{x} \cap \Sigma$, then $x \in O(\boldsymbol{x})$. This property is illustrated by Figure 3.7. Subfigure 3.7a shows the situation when the initial vector $\boldsymbol{x}$ encloses the whole solution set Σ, whereas in Subfigure 3.7b $\boldsymbol{x}$ encloses only a part of Σ. The later corresponds to a situation when only a part of Σ is of interest (e.g., nonnegative solutions).

Notice that step 2 ensures that the sequence $\boldsymbol{x}^0 \supseteq \boldsymbol{x}^1 \supseteq \ldots$ is nested and so converges to some interval vector. Under some assumptions, discussed in the following section, step 2 can be omitted and the iterations still converge to an enclosure of Σ.

Interval operators based on splittings. Most of the operator used are based on a splitting of the constraint matrix. Consider first the real case and split $A = M - N$, where M is nonsingular. Then the system $Ax = b$ reads $Mx = b + Nx$, from which $x = M^{-1}(b + Nx)$. This gives rise to an interval operator

$$O(\boldsymbol{x}) = \boldsymbol{M}^{-1}(\boldsymbol{b} + \boldsymbol{N}\boldsymbol{x}). \tag{3.11}$$

If $\boldsymbol{M}$ is easily invertible (such as a diagonal matrix), then we can use this explicit form. Otherwise, we compute the value of $O(\boldsymbol{x})$ by solving the interval linear system

$$\boldsymbol{M} \cdot O(\boldsymbol{x}) = \boldsymbol{b} + \boldsymbol{N}\boldsymbol{x}. \tag{3.12}$$

The following sections will be devoted to particular types of splittings and the corresponding interval operators, in particular Jacobi, Gauss–Seidel and Krawczyk iterations. More information on iterations based on splittings is provided by Mayer (2017).

3.4.1 The interval Jacobi and Gauss–Seidel method

The Jacobi and Gauss–Seidel methods are classical iterative methods to solve a real system of linear equations $Ax = b$. We will adapt them for solving the interval systems.

The interval Jacobi operator. The ith equation of the real system $Ax = b$ reads $\sum_{j=1}^{n} a_{ij}x_j = b_i$, from which

$$x_i = \tfrac{1}{a_{ii}} \left(b_i - \textstyle\sum_{j=1}^{i \neq i} a_{ij}x_j \right).$$

This gives rise to the classical Jacobi method. Starting with some initial approximation $x^0 \in \mathbb{R}^n$, the kth iteration for the Jacobi method is

1: **for** $i = 1, \dots, n$ **do**
2: $\quad x_i^k := \frac{1}{a_{ii}} \left(b_i - \sum_{j \neq i}^{i-1} a_{ij} x_j^{k-1} \right)$,
3: **end for**

For the interval system $\boldsymbol{A}x = \boldsymbol{b}$, we have to calculate an initial enclosure $\boldsymbol{x}^0 \supseteq \Sigma$ first. Next, the kth interval Jacobi iteration draws as follows.

1: **for** $i = 1, \dots, n$ **do**
2: $\quad \boldsymbol{x}_i^k := \frac{1}{\boldsymbol{a}_{ii}} \left(\boldsymbol{b}_i - \sum_{j \neq i} \boldsymbol{a}_{ij} \boldsymbol{x}_j^{k-1} \right)$;
3: $\quad \boldsymbol{x}_i^k := \boldsymbol{x}_i^k \cap \boldsymbol{x}_i^{k-1}$.
4: **end for**

In other words, it is the iterative method from Section 3.4 using the interval Jacobi operator

$$J(\boldsymbol{x})_i := \frac{1}{\boldsymbol{a}_{ii}} \left(\boldsymbol{b}_i - \sum_{j \neq i} \boldsymbol{a}_{ij} \boldsymbol{x}_j \right).$$

Naturally, we have to assume that $0 \notin \boldsymbol{a}_{ii}$ for all $i \in \{1, \dots, n\}$, unless we employ some type of an extended interval arithmetic (see Remark 3.33). If $A^c = I_n$, which is obtained by a suitable preconditioning, then the condition $\exists i : 0 \in \boldsymbol{a}_{ii}$ implies irregularity of $\boldsymbol{A}$ and thus Σ could be unbounded.

Denote by $\boldsymbol{D}$ the diagonal matrix with entries equal to the diagonal of $\boldsymbol{A}$, and by $\boldsymbol{A'}$ denote $\boldsymbol{A}$ with the diagonal set to zero. Then the interval Jacobi operator has a compact form

$$J(\boldsymbol{x}) := \boldsymbol{D}^{-1}(\boldsymbol{b} - \boldsymbol{A'}\boldsymbol{x}),$$

where $\boldsymbol{D}^{-1} = \operatorname{diag}(\boldsymbol{d}_{11}^{-1}, \dots, \boldsymbol{d}_{nn}^{-1})$. Thus, the interval Jacobi operator is an operator based on splitting (3.11), where $\boldsymbol{M}$ is the diagonal of $\boldsymbol{A}$, and $\boldsymbol{N}$ the rest of $\boldsymbol{A}$.

The important property of the interval Jacobi operator (and any similar interval operator) is that it misses out no solution:

Theorem 3.31. *For each $\boldsymbol{x} \in \mathbb{IR}^n$ we have $J(\boldsymbol{x}) \supseteq \boldsymbol{x} \cap \Sigma$.*

Proof. For each $x \in \boldsymbol{x} \cap \Sigma$ we have $x = D^{-1}(b - A'x) \subseteq \boldsymbol{D}^{-1}(\boldsymbol{b} - \boldsymbol{A'}\boldsymbol{x}) = J(\boldsymbol{x})$. □

The interval Gauss–Seidel operator. The classical Gauss–Seidel method for solving the real system $Ax = b$ is based on expressing the ith equation as

$$x_i = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j - \sum_{j=i+1}^{n} a_{ij} x_j \right).$$

This leads us to the iterative process, where the kth iteration reads

1: **for** $i = 1, \dots, n$ **do**
2: $\quad x_i^k := \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^k - \sum_{j=i+1}^{n} a_{ij} x_j^{k-1} \right)$.
3: **end for**

The interval Gauss–Seidel iteration method for solving $\boldsymbol{A}x = \boldsymbol{b}$ then simply utilizes interval arithmetic

1: **for** $i = 1, \dots, n$ **do**
2: $\quad \boldsymbol{x}_i^k := \frac{1}{\boldsymbol{a}_{ii}} \left(\boldsymbol{b}_i - \sum_{j=1}^{i-1} \boldsymbol{a}_{ij} \boldsymbol{x}_j^k - \sum_{j=i+1}^{n} \boldsymbol{a}_{ij} \boldsymbol{x}_j^{k-1} \right)$;
3: $\quad \boldsymbol{x}_i^k := \boldsymbol{x}_i^k \cap \boldsymbol{x}_i^{k-1}$.
4: **end for**

Thus, the corresponding interval Gauss–Seidel operator reads

$$GS(\boldsymbol{x})_i := \frac{1}{\boldsymbol{a}_{ii}} \left(\boldsymbol{b}_i - \sum_{j=1}^{i-1} \boldsymbol{a}_{ij} GS(\boldsymbol{x})_j - \sum_{j=i+1}^{n} \boldsymbol{a}_{ij} \boldsymbol{x}_j \right).$$

The interval Gauss–Seidel operator is also an operator based on splitting (3.12), where $\boldsymbol{M}$ is the lower triangular part of $\boldsymbol{A}$, $\boldsymbol{N}$ the strictly upper triangular part of $-\boldsymbol{A}$, and the resulting value of $GS(\boldsymbol{x})$ is computed by forward substitution.

Also the interval Gauss–Seidel operator omits no solution included in $\boldsymbol{x}$:

Theorem 3.32. *For each $\boldsymbol{x} \in \mathbb{IR}^n$ we have $GS(\boldsymbol{x}) \supseteq \boldsymbol{x} \cap \Sigma$.*

The interval Gauss–Seidel method is similar to the interval Jacobi one, but the former outperforms the latter since the interval Gauss–Seidel iteration employs the updated values $\boldsymbol{x}_1^k, \ldots, \boldsymbol{x}_{i-1}^k$. Nevertheless, the interval Jacobi iterations may come into play when performing parallel computation.

Both methods converge to the same limit. Later in Theorem 3.43 we will present an explicit formula for the limit.

When $\boldsymbol{A}$ is an M-matrix, then the interval Gauss–Seidel method converges to the optimal $\square\Sigma$; see Barth and Nuding (1974); Neumaier (1990). The interval Gauss–Seidel method is superior to any interval operator based on splitting. A more detailed analysis is given in Mayer (2017); Neumaier (1990).

The interval Gauss–Seidel method was introduced by Ris (1972). From that time, various improvements appeared. Efficient implementation of the interval Gauss–Seidel method is due to Hansen and Sengupta (1981); cf. Hansen and Walster (2004). It takes into account the order of equations that are processed and enables division by zero by using a generalized interval arithmetics, too.

Remark 3.33 (Extended interval arithmetic)**.** The interval Gauss–Seidel method is suitable for employing the extended interval arithmetic (Section 2.3), which is implemented as follows (Neumaier, 1990). Consider an auxiliary one-variable interval equation $\boldsymbol{a}x = \boldsymbol{b}$ with an initial domain $\boldsymbol{x} \in \mathbb{IR}$. Then the restricted solution set

$$\Sigma(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{x}) \coloneqq \{x \in \boldsymbol{x};\ ax = b \text{ for some } a \in \boldsymbol{a}, b \in \boldsymbol{b}\},$$

which is utilized in each step of the interval Gauss–Seidel method when dividing by a diagonal entry, reads

$$\Sigma(\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{x}) = \begin{cases} \frac{\boldsymbol{b}}{\boldsymbol{a}} \cap \boldsymbol{x} & \text{if } 0 \notin \boldsymbol{a}, \\ \boldsymbol{x} & \text{if } 0 \in \boldsymbol{a}, 0 \in \boldsymbol{b}, \\ \square\,(\boldsymbol{x} \cap ([-\infty, \underline{b}/\underline{a}] \cup [\underline{b}/\overline{a}, \infty])) & \text{if } 0 \in \boldsymbol{a}, 0 < \underline{b}, \\ \square\,\big(\boldsymbol{x} \cap \big([-\infty, \overline{b}/\overline{a}] \cup [\overline{b}/\underline{a}, \infty]\big)\big) & \text{if } 0 \in \boldsymbol{a}, 0 > \underline{b}. \end{cases}$$

For concreteness, consider the interval equation

$$[-2, 1]x = [6, 7], \quad x \in \boldsymbol{x} = [-5, 5].$$

By using the extended interval arithmetic, we can shrink the initial domain $\boldsymbol{x}$ to $\Sigma([-2, 1], [6, 7], [-5, 5]) = [-5, -3]$.

3.4.2 Krawczyk method

The Krawczyk method (Krawczyk, 1969) is an interval operator based on splitting (3.11), where $\boldsymbol{M} = I_n$ and $\boldsymbol{N} = I_n - \boldsymbol{A}$. The Krawczyk operator thus reads

$$K(\boldsymbol{x}) \coloneqq \boldsymbol{b} + (I_n - \boldsymbol{A})\boldsymbol{x}.$$

Reliability of the Krawczyk operator is easy to see.

Proposition 3.34. *We have $x \in \boldsymbol{x} \cap \Sigma \subseteq K(\boldsymbol{x})$.*

Proof. Let $x \in \boldsymbol{x} \cap \Sigma$, so $Ax = b$ for some $A \in \boldsymbol{A}$ and $b \in \boldsymbol{b}$. Therefore $x = b + (I_n - A)x \in \boldsymbol{b} + (I_n - \boldsymbol{A})\boldsymbol{x} = K(\boldsymbol{x})$. □

Analysing the Krawczyk method, we obtain a simple bound for the overestimation of the enclosure after one iteration. Moreover, the overestimation is easily calculated by utilizing intermediate calculations.

Proposition 3.35 (Rump, 1990)**.** *Let $\boldsymbol{x} \supseteq \Sigma$ be given. Then the result of one Krawczyk iteration overestimates $\square\Sigma$ by no more than $2\,\mathrm{mag}(I_n - \boldsymbol{A})x^{\Delta}$ from below and from above.*

Proof. Denote

$$\alpha \coloneqq \sup(\boldsymbol{b} + (I_n - \boldsymbol{A})x^c), \quad \beta \coloneqq \operatorname{mag}(I_n - \boldsymbol{A})x^\Delta.$$

Any $x \in \Sigma$ can be written as $x = b + (I_n - A)x = b + (I_n - A)x^c + (I_n - A)(x - x^c)$. Thus,

$$\sup(\square\Sigma) \geq \sup(\boldsymbol{b} + (I_n - \boldsymbol{A})x^c) + \min_{A \in \boldsymbol{A}} (I_n - A)(x - x^c) \geq \alpha - \beta.$$

Similarly,

$$\sup(\square\Sigma) \leq \sup(K(\boldsymbol{x})) = \sup(\boldsymbol{b} + (I_n - \boldsymbol{A})\boldsymbol{x}) \leq \sup(\boldsymbol{b} + (I_n - \boldsymbol{A})x^c + (I_n - \boldsymbol{A})(\boldsymbol{x} - x^c)) \leq \alpha + \beta.$$

Now,

$$\alpha - \beta \leq \sup(\square\Sigma) \leq \sup(K(\boldsymbol{x})) \leq \alpha + \beta,$$

so the overestimation is at most 2β. □

Now, we show that the Krawczyk method yields not as tight enclosures as the interval Jacobi or Gauss–Seidel method. Notice, however, that the Krawczyk method is still worthy of consideration for some purposes, e.g., in parametric interval systems (Section 3.9).

Proposition 3.36. *Suppose that no diagonal entry of $\boldsymbol{A}$ contains the zero. Then one step of the interval Jacobi or Gauss–Seidel iteration is superior to one step of the Krawczyk one.*

Proof. It suffices to compare the Jacobi iteration with the Krawczyk one. Let $\boldsymbol{x}$ be an initial enclosure. The ith step of the Jacobi iteration draws

$$\boldsymbol{x}_i^J \coloneqq \tfrac{1}{\boldsymbol{a}_{ii}} \Big(\boldsymbol{b}_i - \textstyle\sum_{j \neq i} \boldsymbol{a}_{ij} \boldsymbol{x}_j \Big) \cap \boldsymbol{x}_i$$

whereas the Krawczyk iteration puts

$$\boldsymbol{x}_i^K \coloneqq \Big(\boldsymbol{b}_i - \textstyle\sum_{j \neq i} \boldsymbol{a}_{ij} \boldsymbol{x}_j - (\boldsymbol{a}_{ii} - 1)\boldsymbol{x}_i \Big) \cap \boldsymbol{x}_i.$$

Denoting $\boldsymbol{a} \coloneqq \boldsymbol{a}_{ii}$ and $\boldsymbol{c} \coloneqq (\boldsymbol{b})_i - \sum_{j \neq i} (\boldsymbol{A})_{ij} \boldsymbol{x}_j$, the above reads

$$\boldsymbol{x}_i^J \coloneqq \tfrac{\boldsymbol{c}}{\boldsymbol{a}} \cap \boldsymbol{x}_i \quad \text{and} \quad \boldsymbol{x}_i^K \coloneqq (\boldsymbol{c} - (\boldsymbol{a} - 1)\boldsymbol{x}_i) \cap \boldsymbol{x}_i.$$

Suppose that $\boldsymbol{a}_i > 0$ and $\overline{c} \geq 0$; the other situations are dealt with similarly. We compare the right endpoints only since the left endpoints are compared analogously. Suppose that $\boldsymbol{x}_i^J \neq \emptyset$, otherwise we are done. This implies that $\underline{x_i} \leq \overline{c}/\underline{a}$. If $\underline{a} \geq 1$, then

$$\sup(\boldsymbol{c} - (\boldsymbol{a} - 1)\boldsymbol{x}_i) \geq \overline{c} - (\underline{a} - 1)\underline{x}_i \geq \overline{c} - (\underline{a} - 1)\frac{\overline{c}}{\underline{a}} = \frac{\overline{c}}{\underline{a}} = \sup\Big(\frac{\boldsymbol{c}}{\boldsymbol{a}}\Big),$$

showing $\sup(\boldsymbol{x}_i^K) \geq \sup(\boldsymbol{x}_i^J)$. If $\underline{a} < 1$, then

$$\sup(\boldsymbol{c} - (\boldsymbol{a} - 1)\boldsymbol{x}_i) \geq \overline{c} - (\underline{a} - 1)\overline{x}_i \geq \sup(\boldsymbol{x}_i^J)\underline{a} - (\underline{a} - 1)\sup(\boldsymbol{x}_i^J) = \sup(\boldsymbol{x}_i^J).$$

In this case, $\sup(\boldsymbol{x}_i^K) \geq \sup(\boldsymbol{x}_i^J)$, too. □

A stronger result that preconditioned interval Gauss–Seidel iteration is always superior to the Krawczyk one can be found in (Neumaier, 1990, Thm. 4.3.5).

The residual Krawczyk method. The residual form transformation presented in Section 3.2.2 substitutes the variables $x := x^* + y$ and considers the system in the form $\boldsymbol{A}y = \boldsymbol{b} - \boldsymbol{A}x^*$. If $\boldsymbol{y}$ is an enclosure to the corresponding solution set, then $\boldsymbol{y} + x^*$ encloses the original one. The Krawczyk operator for the transformed system reads

$$K_r(\boldsymbol{y}) := \boldsymbol{b} - \boldsymbol{A}x^* + (I_n - \boldsymbol{A})\boldsymbol{y}.$$

Theorem 3.37 (Neumaier, 1990)**.** *The residual Krawczyk method produces no tighter enclosures than the Krawczyk method.*

Proof. We estimate the Krawczyk operator from above by the residual one

$$\begin{aligned} K(\boldsymbol{x}) &= K(\boldsymbol{y} + x^*) = \boldsymbol{b} + (I_n - \boldsymbol{A})(\boldsymbol{y} + x^*) \\ &\subseteq \boldsymbol{b} + (I_n - \boldsymbol{A})\boldsymbol{y} + x^* - \boldsymbol{A}x^* = \boldsymbol{b} - \boldsymbol{A}x^* + (I_n - \boldsymbol{A})\boldsymbol{y} = K_r(\boldsymbol{y}). \end{aligned}$$

□

Even though this result does not praise the residual version, the situation is not so bad. As long as $C = (A^c)^{-1}$, both methods converge to the same fixed point (see Corollary3.45 and around), even though the speed of the residual one is smaller. The residual form will be particularly convenient when solving parametric interval systems (Section 3.9).

3.4.3 ε-inflation

The previous iteration methods started with an enclosure of Σ, and iteratively tightened it to make it as small as possible. In contrast, ε-inflation goes the opposite direction. It starts with a small interval vector and blows it up until a rigorous enclosure of the solution set is achieved.

Let $\mathcal{O}_{\boldsymbol{A},\boldsymbol{b}}(\boldsymbol{x})$ be an interval operator for Σ satisfying three natural properties:

1. for every $A \in \boldsymbol{A}$ and $b \in \boldsymbol{b}$ and for every $\boldsymbol{x}$ we have $\mathcal{O}_{A,b}(\boldsymbol{x}) \subseteq \mathcal{O}_{\boldsymbol{A},\boldsymbol{b}}(\boldsymbol{x})$, that is, it satisfies some restricted inclusion isotonicity;
2. $\mathcal{O}_{A,b}\colon \mathbb{R}^n \to \mathbb{R}^n$ is continuous for each $A \in \boldsymbol{A}$ and $b \in \boldsymbol{b}$ and when restricted to $\mathbb{R}^n$;
3. $\mathcal{O}_{A,b}(x) = x$ for each $A \in \boldsymbol{A}$, $b \in \boldsymbol{b}$ and $x \in \mathbb{R}^n$ such that $Ax = b$.

An example of such an interval operator is, e.g., the Krawczyk one.

Theorem 3.38. *If $\mathcal{O}_{\boldsymbol{A},\boldsymbol{b}}(\boldsymbol{x}) \subseteq \boldsymbol{x}$, then for each $A \in \boldsymbol{A}$ and $b \in \boldsymbol{b}$, there is a fixed point of $\mathcal{O}_{A,b}$ in $\boldsymbol{x}$, which is a solution of $Ax = b$.*

Proof. It follows from Brouwer's fixed point Theorem 1.26 as $\mathcal{O}_{A,b}(\boldsymbol{x}) \subseteq \boldsymbol{x}$. □

However, as illustrated by the example below, the condition $\mathcal{O}_{\boldsymbol{A},\boldsymbol{b}}(\boldsymbol{x}) \subseteq \boldsymbol{x}$ is not sufficient to prove regularity of $\boldsymbol{A}$ and the inclusion $\Sigma \subseteq \boldsymbol{x}$.

Example 3.39. Let

$$\boldsymbol{A} = \begin{pmatrix} [0,2] & 0 \\ 0 & [0,2] \end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{x} = \begin{pmatrix} [-1,1] \\ [-1,1] \end{pmatrix}.$$

For the Krawczyk operator, we have $K(\boldsymbol{x}) = \boldsymbol{x}$, but $\boldsymbol{A}$ is singular and $\Sigma = \mathbb{R}^2$ is unbounded. □

We derive a stronger result for the Krawczyk operator by adopting the procedure from Rump (2010).

Lemma 3.40. *Let $\boldsymbol{x}, \boldsymbol{v} \in \mathbb{IR}^n$ and $\boldsymbol{V} \in \mathbb{IR}^{n\times n}$. If*

$$\boldsymbol{v} + \boldsymbol{V}\boldsymbol{x} \subseteq \operatorname{int} \boldsymbol{x},$$

then $\rho(\boldsymbol{V}) < 1$.

Proof. Let $V \in \boldsymbol{V}$. Since

$$\boldsymbol{v} + V\boldsymbol{x} = \boldsymbol{v} + Vx^c + [-|V|x^\Delta, |V|x^\Delta] \subseteq \operatorname{int} \boldsymbol{x},$$

we have $|V|x^\Delta < x^\Delta$. Since $x^\Delta > 0$, by the Perron–Frobenius theory (Theorems 1.8 and 1.9), $\rho(V) \leq \rho(|V|) < 1$. □

Applying it to the Krawczyk operator, we get the following.

Theorem 3.41. *Let $\boldsymbol{x} \in \mathbb{IR}^n$. If*

$$\boldsymbol{b} + (I - \boldsymbol{A})\boldsymbol{x} \subseteq \operatorname{int} \boldsymbol{x},$$

then $\boldsymbol{A}$ is regular and $\Sigma \subseteq \boldsymbol{x}$.

Proof. By Lemma 3.40, we have $\rho(I - \boldsymbol{A}) < 1$, which implies regularity of $\boldsymbol{A}$. Now, use Theorem 3.38 to prove $\Sigma \subseteq \boldsymbol{x}$. □

Implementation. The ε-inflation method is usually implemented for the preconditioned residual Krawczyk form. We precondition by the numerically computed inverse of A^c, and we apply the residual form with x^* being a numerically computed solution of $A^c x = b^c$. After this transformation, the iterations go as follows. Set $\boldsymbol{x} := \boldsymbol{b}$ and repeat performing the inflation $\boldsymbol{y} := [0.9, 1.1]\boldsymbol{x} + 10^{-20}[-1, 1]^n$ and updating $\boldsymbol{x} := \boldsymbol{b} + (I - \boldsymbol{A})\boldsymbol{y}$ until $\boldsymbol{x} \subseteq \operatorname{int} \boldsymbol{y}$. If we succeed, $\Sigma \subseteq x^* + \boldsymbol{x}$.

Notes. The ε-inflation method was introduced by Alefeld and Apostolatos (1968); Caprani and Madsen (1978) and popularized (and named) by Rump (1980, 1983, 1992). It was implemented in INTLAB up to version 6 in the first stage of the function `verifylss` by using the Krawczyk operator.

3.4.4* Interval operators in a general framework

As we saw in the previous sections, the interval operator often takes the form of

$$O(\boldsymbol{x}) := \boldsymbol{U}(\boldsymbol{v} + \boldsymbol{V}\boldsymbol{x}). \tag{3.13}$$

for some interval matrices $\boldsymbol{U}, \boldsymbol{V} \in \mathbb{IR}^{n \times n}$.

Based on the iterations, one can derive a direct method to enclose Σ. Notice, however, that it requires another verified inverse (or a system of linear equations) calculation.

Theorem 3.42. *If $\rho(\operatorname{mag}(\boldsymbol{U}) \operatorname{mag}(\boldsymbol{V})) < 1$, then*

$$\Sigma \subseteq \boldsymbol{U}(\boldsymbol{v} + [-1, 1]\big((I_n - \operatorname{mag}(\boldsymbol{V}) \operatorname{mag}(\boldsymbol{U}))^{-1} - I_n\big) \operatorname{mag}(\boldsymbol{v})).$$

Moreover, provided $V^c = 0$ and $\boldsymbol{U}$ is diagonal, then the enclosure is the limit of the iterations.

Proof. The idea is to find $\boldsymbol{x}$ such that such that $\boldsymbol{x} \supseteq \boldsymbol{U}(\boldsymbol{v}+\boldsymbol{V}\boldsymbol{x})$. This inclusion implies that for each $v \in \boldsymbol{v}$, $U \in \boldsymbol{U}$ and $V \in \boldsymbol{V}$ we have $\boldsymbol{x} \supseteq Uv + UV\boldsymbol{x}$ and $\rho(UV) \leq \rho(\operatorname{mag}(\boldsymbol{U}) \operatorname{mag}(\boldsymbol{V})) < 1$. By Theorem 1.27, the operator $x \mapsto Uv + UVx$ has a unique fixed point, and it is in $\boldsymbol{x}$. Since the fixed points of the mapping correspond to the solutions of the system of linear equations, $\Sigma \subseteq \boldsymbol{x}$.

To find such an interval vector, we overestimate the right-hand side and solve the algebraic equation

$$\boldsymbol{x} = \boldsymbol{U}(\boldsymbol{v} + [-1, 1] \operatorname{mag}(\boldsymbol{V}) \operatorname{mag}(\boldsymbol{x})).$$

Note that if $V^c = 0$ and $\boldsymbol{U}$ is diagonal, then the right-hand sides are the same and no overestimation happens. The magnitude of $\boldsymbol{x}$ draws

$$\operatorname{mag}(\boldsymbol{x}) \leq \operatorname{mag}(\boldsymbol{U})(\operatorname{mag}(\boldsymbol{v}) + \operatorname{mag}(\boldsymbol{V}) \operatorname{mag}(\boldsymbol{x})),$$

and it holds as equation as long as $\boldsymbol{U}$ is diagonal. Factoring $\operatorname{mag}(\boldsymbol{x})$, we get

$$\operatorname{mag}(\boldsymbol{x}) \leq (I_n - \operatorname{mag}(\boldsymbol{U}) \operatorname{mag}(\boldsymbol{V}))^{-1} \operatorname{mag}(\boldsymbol{U}) \operatorname{mag}(\boldsymbol{v}). \tag{3.14}$$

Since $\rho(\operatorname{mag}(\boldsymbol{U}) \operatorname{mag}(\boldsymbol{V})) < 1$ and $\operatorname{mag}(\boldsymbol{U}), \operatorname{mag}(\boldsymbol{V}) \geq 0$, Theorem 1.10 implies $(I_n - \operatorname{mag}(\boldsymbol{U}) \operatorname{mag}(\boldsymbol{V}))^{-1} = \sum_{j=0}^{\infty} (\operatorname{mag}(\boldsymbol{U}) \operatorname{mag}(\boldsymbol{V}))^j \geq 0$, and therefore the above inequality is valid and $\operatorname{mag}(\boldsymbol{x}) \geq 0$. Substituting for $\operatorname{mag}(\boldsymbol{x})$ its upper bound and using Neumann series again, we obtain the enclosure

$$\begin{aligned} \boldsymbol{x}^* &= \boldsymbol{U}(\boldsymbol{v} + [-1, 1] \operatorname{mag}(\boldsymbol{V})(I_n - \operatorname{mag}(\boldsymbol{U}) \operatorname{mag}(\boldsymbol{V}))^{-1} \operatorname{mag}(\boldsymbol{U}) \operatorname{mag}(\boldsymbol{v})) \\ &= \boldsymbol{U}(\boldsymbol{v} + [-1, 1]((I_n - \operatorname{mag}(\boldsymbol{V}) \operatorname{mag}(\boldsymbol{U}))^{-1} - I_n) \operatorname{mag}(\boldsymbol{v})). \end{aligned} \tag{3.15}$$

If $V^c = 0$ and $\boldsymbol{U}$ is diagonal, then the above enclosure satisfies $\boldsymbol{x}^* = \boldsymbol{U}(\boldsymbol{v} + \boldsymbol{V}\boldsymbol{x}^*)$. We claim that the iterations $\boldsymbol{x} \mapsto \boldsymbol{U}(\boldsymbol{v} + \boldsymbol{V}\boldsymbol{x})$ converge to this unique interval vector for any initial $\boldsymbol{x} \in \mathbb{IR}^n$. The magnitude of $\boldsymbol{x}$ changes as follows

$$\operatorname{mag}(\boldsymbol{x}) \mapsto \operatorname{mag}(\boldsymbol{U})(\operatorname{mag}(\boldsymbol{v}) + \operatorname{mag}(\boldsymbol{V}) \operatorname{mag}(\boldsymbol{x})).$$

Since $\rho(\operatorname{mag}(\boldsymbol{U}) \operatorname{mag}(\boldsymbol{V})) < 1$, we have by Theorem 1.27 that the magnitudes $\operatorname{mag}(\boldsymbol{x})$ converge to the right-hand side of (3.14), which is $\operatorname{mag}(\boldsymbol{x}^*)$. Now, for any $\boldsymbol{x} \in \mathbb{IR}^n$ with $\operatorname{mag}(\boldsymbol{x}) = \operatorname{mag}(\boldsymbol{x}^*)$ we have

$$O(\boldsymbol{x}) \ = \boldsymbol{U}(\boldsymbol{v} + \boldsymbol{V}\boldsymbol{x}) = \boldsymbol{U}(\boldsymbol{v} + \boldsymbol{V}\boldsymbol{x}^*) = \boldsymbol{x}^*,$$

which completes the proof. □

Interval Jacobi and Gauss–Seidel methods. By a slight modification of Theorem 3.42, we obtain a limit of the interval Jacobi (and Gauss–Seidel) iterations. Notice that an equivalent expression is given in (Neumaier, 1990, Thm. 4.4.10).

Theorem 3.43. *If $\rho(A^\Delta) < 1$ and $A^c = I_n$, then*

$$\Sigma \subseteq \boldsymbol{D}^{-1}(\boldsymbol{b} + [-1, 1] \operatorname{mag}(\boldsymbol{A}')\underline{A}^{-1} \operatorname{mag}(\boldsymbol{b})),$$

and the enclosure is the limit of the Jacobi iterations.

Proof. We call Theorem 3.42 with $\boldsymbol{U} \coloneqq \boldsymbol{D}^{-1}$, $\boldsymbol{V} \coloneqq -\boldsymbol{A}'$, and $\boldsymbol{v} \coloneqq \boldsymbol{b}$. Using the form (3.15) and the fact that $\operatorname{mag}(\boldsymbol{D}^{-1}) = \underline{D}^{-1}$, we get

$$\Sigma \subseteq \underline{D}^{-1}(\boldsymbol{b} + [-1, 1] \operatorname{mag}(\boldsymbol{A}')(I_n - \underline{D}^{-1} \operatorname{mag}(\boldsymbol{A}'))^{-1}\underline{D}^{-1} \operatorname{mag}(\boldsymbol{b})).$$

Since

$$(I_n - \underline{D}^{-1} \operatorname{mag}(\boldsymbol{A}'))^{-1}\underline{D}^{-1} = (\underline{D}(I_n - \underline{D}^{-1} \operatorname{mag}(\boldsymbol{A}')))^{-1} = (\underline{D} - \operatorname{mag}(\boldsymbol{A}'))^{-1} = \underline{A}^{-1},$$

we have the final form. Notice that the assumption of Theorem 3.42 is satisfied since from $\rho(A^\Delta) < 1$ there is $x > 0$ such that $(A'^\Delta + D^\Delta)x = A^\Delta x < x$, whence $A'^\Delta x < (I_n - D^\Delta)x$. Now,

$$\operatorname{mag}(\boldsymbol{D}^{-1}) \operatorname{mag}(\boldsymbol{A}')x = (I_n - D^\Delta)^{-1}A'^\Delta x < (I_n - D^\Delta)^{-1}(I_n - D^\Delta)x = x.$$

Therefore $\rho(\operatorname{mag}(\boldsymbol{D}^{-1}) \operatorname{mag}(\boldsymbol{A}')) < 1$. □

Krawczyk method. By Theorem 3.42, the direct enclosure to Σ derived from the Krawczyk iterations is as follows; we simply substitute $\boldsymbol{U} :- I_n$, $\boldsymbol{V} :- I_n - \boldsymbol{A}$ and $\boldsymbol{v} \coloneqq \boldsymbol{b}$.

Corollary 3.44. *If $\rho(\operatorname{mag}(I_n - \boldsymbol{A})) < 1$, then*

$$\Sigma \subseteq \boldsymbol{b} + [-1, 1]\big((I_n - \operatorname{mag}(I_n - \boldsymbol{A}))^{-1} - I_n\big) \operatorname{mag}(\boldsymbol{b}).$$

If $A^c = I_n$, then the enclosure is the limit of the iterations.

The residual Krawczyk method. If we apply the residual form of an interval system of linear equations, then the enclosure has the form of $\boldsymbol{x} = x^* + \boldsymbol{y}$, where $\boldsymbol{y}$ encloses the solution set of $\boldsymbol{A}y = \boldsymbol{b} - \boldsymbol{A}x^*$ and $x^* \in \mathbb{R}^n$ is fixed. Applying the above result for the Krawczyk method to this form, we get the following result.

Corollary 3.45. *If $\rho(\operatorname{mag}(I_n - \boldsymbol{A})) < 1$, then*

$$\Sigma \subseteq x^* + \boldsymbol{b} - \boldsymbol{A}x^* + [-1, 1]\big((I_n - \operatorname{mag}(I_n - \boldsymbol{A}))^{-1} - I_n\big) \operatorname{mag}(\boldsymbol{b} - \boldsymbol{A}x^*). \tag{3.16}$$

If $A^c = I_n$, then the enclosure is the limit of the iterations.

It is an easy exercise to show by comparing the centers and radii that this enclosure equals to that from the Krawczyk method (Corollary 3.44) provided $A^c = I_n$ and $x^* = b^c$. Thus, both have the same fixed point. Surprisingly, the generalization of the well-known Bauer–Skeel bounds for the solution of a perturbed linear system yields the same enclosure; see Rohn (2010).

Remark 3.46. Besides (3.16), sometimes a slightly simpler enclosing formula

$$\Sigma \subseteq x^* + [-1,1](I_n - \operatorname{mag}(I_n - \boldsymbol{A}))^{-1} \operatorname{mag}(\boldsymbol{b} - \boldsymbol{A}x^*) \tag{3.17}$$

is used; see Section 3.9. If $A^c = I_n$ and $x^* = b^c$, then both enclosures are the same, but in general (3.16) is provably tighter.

Notes and further reading. More on interval operators and their convergence can be found in Alefeld and Herzberger (1983); Neumaier (1990). For example, in the case with $\boldsymbol{U} = 0$, the iterations converge to a unique fixed point $\boldsymbol{x}^*$ for every starting point $\boldsymbol{x}^0$ if and only if $\rho(\operatorname{mag}(\boldsymbol{V})) < 1$. In which case, $\boldsymbol{x}^*$ contains the whole solution set.

Iterations of general interval operators and their convergence are discussed by Mayer (2017).

3.5 The square case – hull methods and special cases

This section is devoted to the methods that compute $\square\Sigma$. Due to its intractability, the methods have exponential complexity in the worst case. That is why we also consider special kinds of interval matrices, for which the interval hull of the solution set can be calculated more easily.

3.5.1 Hansen–Bliek–Rohn method

Herein, we describe a direct method for calculating the tightest enclosure to Σ, that is, $\square\Sigma$. We make use of the assumption $A^c = I_n$. This property may be achieved by preconditioning the constraint interval matrix by the midpoint inverse, and by a slight inflation such that $A^c = I_n$ exactly; cf. Section 3.2.1. The method was developed independently by Hansen (1992) and Bliek (1992) and further improved by a rigorous proof by Rohn (1993a); see also Ning and Kearfott (1997); Neumaier (1999).

Theorem 3.47. *Let $A^c = I_n$ and $\rho(A^\Delta) < 1$. Denote*

$$\begin{aligned} M &\coloneqq \underline{A}^{-1}, \\ x^* &\coloneqq M \operatorname{mag}(\boldsymbol{b}). \end{aligned}$$

For every $i = 1, \ldots, n$, we have

$$\sup(\square\Sigma)_i \le \max\left\{x_i^* + M_{ii}(b_i^c - |b_i^c|),\ \tfrac{1}{2M_{ii}-1}\left(x_i^* + M_{ii}(b_i^c - |b_i^c|)\right)\right\}, \tag{3.18a}$$

$$\inf(\square\Sigma)_i \ge \min\left\{-x_i^* + M_{ii}(b_i^c + |b_i^c|),\ \tfrac{1}{2M_{ii}-1}\left(-x_i^* + M_{ii}(b_i^c + |b_i^c|)\right)\right\}. \tag{3.18b}$$

Proof. By the Oettli–Prager Theorem 3.2,

$$|x - b^c| \le A^\Delta |x| + b^\Delta.$$

from which we can derive two inequalities

$$x - b^c \le A^\Delta |x| + b^\Delta,$$

and

$$|x| - |b^c| \le A^\Delta |x| + b^\Delta.$$

Taking the ith inequality from the former system and the other inequalities from the latter system we have

$$|x| + (x_i - |x|_i)e_i - |b^c| - (b_i^c - |b_i^c|)e_i \leq A^\Delta |x| + b^\Delta,$$

or

$$(I_n - A^\Delta)|x| + (x_i - |x|_i)e_i \leq |b^c| + (b_i^c - |b_i^c|)e_i + b^\Delta.$$

By Theorem 1.10 we have

$$M = \left(I_n - A^\Delta\right)^{-1} = \sum_{k=0}^{\infty} \left(A^\Delta\right)^k \geq I_n \geq 0.$$

Thus, multiplying the previous inequalities by M gives

$$|x| + (x_i - |x|_i)Me_i \leq x^* + (b_i^c - |b_i^c|)Me_i.$$

The ith inequality draws

$$|x|_i + (x_i - |x|_i)M_{ii} \leq x_i^* + (b_i^c - |b_i^c|)M_{ii}.$$

We distinguish two cases. If $x_i \geq 0$, then

$$x_i \leq x_i^* + (b_i^c - |b_i^c|)M_{ii}.$$

Provided $x_i < 0$ we get

$$(2M_{ii} - 1)x_i \leq x_i^* + (b_i^c - |b_i^c|)M_{ii}.$$

Since $M \geq I_n$, we have $2M_{ii} > 1$ and so

$$x_i \leq \frac{1}{(2M_{ii} - 1)} \left(x_i^* + (b_i^c - |b_i^c|)M_{ii}\right).$$

Putting together, we obtain (3.18a).

The lower bound (3.18b) can be proven by reducing to the previous case by transformation $\boldsymbol{A}(-x) = -\boldsymbol{b}$. In this way, b^c changes its sign and for any $x \in \Sigma$ we have

$$-x_i \leq \max\left\{x_i^* + M_{ii}(-b_i^c - |b_i^c|), \frac{1}{2M_{ii} - 1}\left(x_i^* + M_{ii}(-b_i^c - |b_i^c|)\right)\right\}.$$

from which (3.18b) follows. □

We have even a stronger result – the Hansen–Bliek–Rohn method produces $\square\Sigma$ exactly, up to the numerical accuracy.

Theorem 3.48. *The Hansen–Bliek–Rohn method yields $\square\Sigma$.*

Proof. Let $i \in \{1, \ldots, n\}$. We have $M = \underline{A}^{-1}$, $x_i^* = M \operatorname{mag}(\boldsymbol{b})$. We will find two realizations of $\boldsymbol{A}$ and $\boldsymbol{b}$ for which the ith entries of their solutions are equal to x_i^* and $\frac{1}{2M_{ii}-1}x_i^*$. This will prove the upper bound; the lower bound can be proven accordingly.

For any $y, z \in \{\pm 1\}^n$ consider the system $A_{yz}x = b_y$, i.e.,

$$(I_n - \operatorname{diag}(y)A^\Delta \operatorname{diag}(z))x = b^c + \operatorname{diag}(y)b^\Delta.$$

Put $z' := z + (1 - z_i)e_i$, and substitute $x' := \operatorname{diag}(z')x$. Now, the system reads

$$\left(I_n - \operatorname{diag}(y)A^\Delta \operatorname{diag}(z)\right) \operatorname{diag}(z')x' = b^c + \operatorname{diag}(y)b^\Delta.$$

Note that the solutions of both systems have the same ith entry. Multiplying by $\operatorname{diag}(z')$ we obtain

$$\left(I_n - \operatorname{diag}(z')\operatorname{diag}(y)A^\Delta \operatorname{diag}(z)\operatorname{diag}(z')\right)x' = \operatorname{diag}(z')b^c + \operatorname{diag}(z')\operatorname{diag}(y)b^\Delta.$$

Now, we associate $y \coloneqq z'$ and $z_j = z'_j \coloneqq \operatorname{sgn}(b^c_j)$, $j \neq i$. Thus, the only free parameter is z_i and the system takes the form of

$$\left(I_n - A^\Delta \operatorname{diag}(z)\operatorname{diag}(z')\right)x' = \operatorname{diag}(z')b^c + b^\Delta,$$

or

$$\left(I_n - A^\Delta \operatorname{diag}(z)\operatorname{diag}(z')\right)x' = \sum_{j\neq i}(|b^c_j| + b^\Delta_j)e_j + \overline{b}_i e_i.$$

We distinguish two cases. First, if $z_i \coloneqq 1$, then $z = z'$ and the system draws

$$\left(I_n - A^\Delta\right)x' = \sum_{j\neq i}(|b^c_j| + b^\Delta_j)e_j + \overline{b}_i e_i,$$

or

$$\underline{A}x' = \sum_{j\neq i}(|b^c_j| + b^\Delta_j)e_j + \overline{b}_i e_j,$$

Its solution is

$$x' = M\left(\sum_{j\neq i}(|b^c_j| + b^\Delta_j)e_j + \overline{b}_i e_i\right) = \sum_{j\neq i}(|b^c_j| + b^\Delta_j)M_{*j} + \overline{b}_i M_{*i}$$

with the ith entry $x'_i = \sum_{j\neq i} M_{ij}(|b^c_j| + b^\Delta_j) + M_{ii}\overline{b}_i = x^*_i$.

Next, put $z_i \coloneqq -1 = -z'_i$. Hence the system can be written as

$$\left(\underline{A} + 2A^\Delta_{*i}\right)x' = \sum_{j\neq i}(|b^c_j| + b^\Delta_j)e_j + \overline{b}_i e_i,$$

To determine its solution, we utilize the Sherman–Morrison formula (Theorem 1.2) to calculate the inverse of $(\underline{A} + 2A^\Delta_{*i})$. Put $a \coloneqq 2A^\Delta_{*i} = 2(e_i - \underline{A}_{*i}) = 2(e_i - M^{-1}_{*i})$ and derive

$$\begin{aligned}\left(\underline{A} + 2A^\Delta_{*i}\right)^{-1} &= \left(\underline{A} + ae_i^T\right)^{-1} = M - \frac{1}{1 + e_i^T Ma}Mae_i^T M \\ &= M - \frac{2}{1 + 2M_{i*}(e_i - M^{-1}_{*i})}M(e_i - M^{-1}_{*i})M_{i*} = M - \frac{2}{2M_{ii} - 1}(M_{*i} - e_i)M_{i*}.\end{aligned}$$

The ith row of the inverse is

$$\left(\underline{A} + 2A^\Delta_{*i}\right)^{-1}_{i*} = M_{i*} - \frac{2}{2M_{ii} - 1}(M_i - 1)M_{i*} = \frac{1}{2M_{ii} - 1}M_{i*}.$$

Hence the solution of the system has the ith entry $x'_i = \frac{1}{2M_{ii}-1}x^*_i$. □

It follows from the above theorems and their proofs that $b^c_i \geq 0$ implies $\sup(\square\Sigma)_i = x^*_i$ and $\inf(\square\Sigma)_i \geq -x^*_i$, whence $\operatorname{mag}(\square\Sigma)_i = x^*_i$. Similarly, if $b^c_i \leq 0$ then $\sup(\square\Sigma)_i \leq x^*_i$ and $\inf(\square\Sigma)_i = -x^*_i$, whence $\operatorname{mag}(\square\Sigma)_i = x^*_i$. In any case,

$$\operatorname{mag}(\square\Sigma) = x^* = \underline{A}^{-1}\operatorname{mag}(\boldsymbol{b}).$$

3.5.2 Comparison of methods

Suppose that $A^c = I_n$, and let $\boldsymbol{A}$ be regular. Denote by $\boldsymbol{x}^K$, $\boldsymbol{x}^{GS}$, $\boldsymbol{x}^G$, $\boldsymbol{x}^{HBR}$ the limiting enclosures of the Krawczyk and Gauss–Seidel method, the enclosure by the interval Gaussian elimination, and the Hansen–Bliek–Rohn one, respectively.

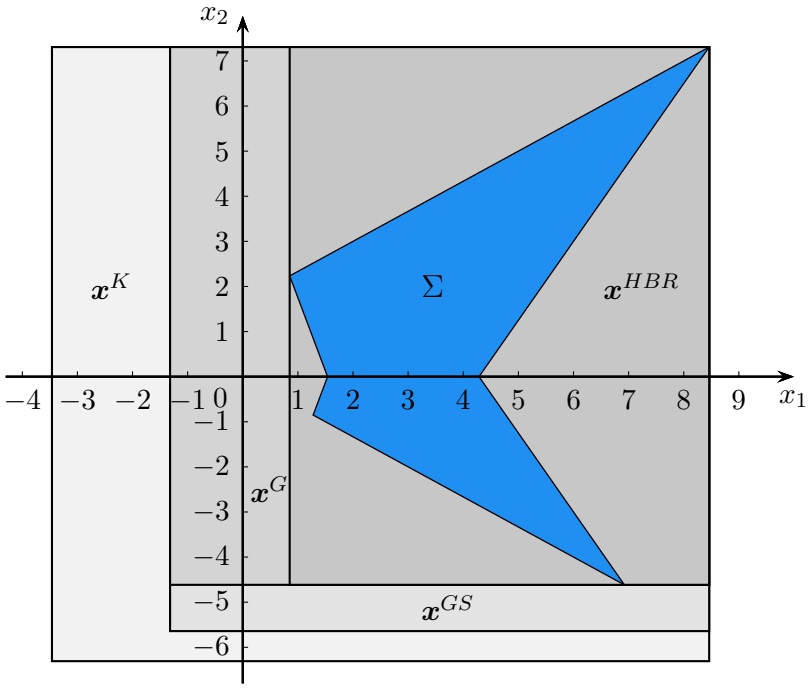


Figure 3.8: (Example 3.50) Comparison of methods for enclosing the solution set Σ.

Theorem 3.49. *If $A^c = I_n$, then*

$$\operatorname{mag}(\square\Sigma) = \operatorname{mag}(\boldsymbol{x}^{HBR}) = \operatorname{mag}(\boldsymbol{x}^G) = \operatorname{mag}(\boldsymbol{x}^{GS}) = \operatorname{mag}(\boldsymbol{x}^K) = \underline{A}^{-1} \operatorname{mag}(\boldsymbol{b}),$$
$$\square\Sigma = \boldsymbol{x}^{HBR} \subseteq \boldsymbol{x}^G \subseteq \boldsymbol{x}^{GS} \subseteq \boldsymbol{x}^K.$$

Proof. It follows from the results developed in the previous section. The comparison with the Gaussian elimination comes from (Neumaier, 1990, Thm. 4.5.11). □

Example 3.50. Consider the interval linear system of equations $\boldsymbol{A}x = \boldsymbol{b}$, where

$$\boldsymbol{A} = \begin{pmatrix} [0.7, 1.3] & [-0.4, 0.4] \\ [-0.4, 0.4] & [0.6, 1.4] \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} [2, 3] \\ [0, 1] \end{pmatrix}.$$

Notice that $A^c = I_n$, so we do not apply preconditioning. Enclosures calculated by the particular methods are:

$$\boldsymbol{x}^{HBR} = \begin{pmatrix} [0.8510, 8.4616] \\ [-4.6154, 7.3077] \end{pmatrix}, \qquad \boldsymbol{x}^G = \begin{pmatrix} [-1.3187, 8.4616] \\ [-4.6154, 7.3077] \end{pmatrix},$$
$$\boldsymbol{x}^{GS} = \begin{pmatrix} [-1.3187, 8.4616] \\ [-5.6411, 7.3077] \end{pmatrix}, \qquad \boldsymbol{x}^K = \begin{pmatrix} [-3.4616, 8.4616] \\ [-6.3077, 7.3077] \end{pmatrix}.$$

Figure 3.8 illustrates the solution set Σ and the enclosures nested in the way as stated in Theorem 3.49. Notice also, in accordance with the theorem, that in each entry the interval enclosures have at least one bound tight (in this case the upper bound). □

The theorem says that, under the assumption, all the basic enclosure methods are exact in at least one endpoint of each coordinate, and their tightness differs only in the second endpoint. This property motivated the magnitude method Hladík (2014b) based on a special operator starting with the enclosure $[-\underline{A}^{-1} \operatorname{mag}(\boldsymbol{b}), \underline{A}^{-1} \operatorname{mag}(\boldsymbol{b})]$ and trying to approach the Hansen–Bliek–Rohn bounds. It yields a tight enclosure in a very low time.

3.5.3 Tightening methods

By a *tightening method*, we mean a method that starts with an initial enclosure $\boldsymbol{x}$ of the solution set Σ computed by a standard method (discussed in the previous sections) and tries to make it even more tighter. Such tightening methods are expected to run in polynomial time, but the price for tightening usually lies in higher computational cost than for the standard methods.

Monotonicity checking

Monotonicity checking (Hansen, 1969a) enables to fix some interval coefficients at either the lower or upper bounds based on monotonicity of the solution w.r.t. the coefficient. Provided the kth entry of the solution x_k is nondecreasing on the domain, that is,

$$\frac{\partial x_k}{\partial a_{ij}} \geq 0, \tag{3.19}$$

then $\overline{\Sigma}_k$ is attained for the matrix $A \in \boldsymbol{A}$ with $a_{ij} = \overline{a}_{ij}$, and $\underline{\Sigma}_k$ is attained for $a_{ij} = \underline{a}_{ij}$. This idea gives rise to a method that improves the lower and upper bounds of the particular entries of the solution set. Let $k \in \{1, \ldots, n\}$ and let us consider the upper bound case; the lower bounds are tightened accordingly. We fix at upper endpoints $\boldsymbol{a}_{ij} \coloneqq \overline{a}_{ij}$ all i, j for which x_k is nondecreasing with respect to x_k, and we fix at lower endpoints $\boldsymbol{a}_{ij} \coloneqq \underline{a}_{ij}$ all i, j for which x_k is nonincreasing with respect to x_k. Similarly for $\boldsymbol{b}$. Thus, we have a subset of $\boldsymbol{A}$ and an enclosure method yields a hopefully sharper upper bound on $\overline{\Sigma}_k$.

Checking for monotonicity (3.19) is not easy in general, so we propose a simple sufficient condition. Recall the formulae from Theorem 1.25

$$\frac{\partial x_k}{\partial a_{ij}} = -(A^{-1})_{ki} x_j, \quad \frac{\partial x_k}{\partial b_i} = (A^{-1})_{ki}.$$

Let $\boldsymbol{C} \in \mathbb{IR}^{n \times n}$ be an enclosure of the set of inverses $\{A^{-1};\ A \in \boldsymbol{A}\}$; compare Section 5.4.4. Now, if $-\boldsymbol{c}_{ki}\boldsymbol{x}_j \geq 0$, then (3.19) holds. Analogously, $\boldsymbol{c}_{ki} \geq 0$ implies monotonicity of x_k with respect to b_i.

Notice that similar ideas are used also for tightening enclosures of other functions; compare Chapter 6.

3.5.4 Hull methods

The previous methods are designed to efficiently calculate enclosures of the solution set. The calculated enclosures are usually tight enough, but not optimal in general. Hull methods is a name for methods that determine the interval hull of the solution set (exactly up to the numerical calculations). In view of the *NP*-hardness, they are exponential in the worst case. However, they are not necessarily exponential in every case.

Jansson's algorithm

The first naive algorithm is based on Theorem 3.5. In each orthant, the solution set forms a convex polyhedral set. Thus, we can go through all orthants, determine bounds of the particular convex polyhedral sets (if they are non-empty) and output the maximal bounds. Since there are 2^n orthants, this leads to a very inefficient algorithm.

The naive approach can easily be improved by computing an enclosure $\boldsymbol{x}^0$ to Σ by any of the mentioned methods. Now, it is sufficient to inspect orthants that have nonempty intersection with $\boldsymbol{x}^0$. Hence we check the orthants corresponding to sign vectors $s \in \{\pm 1\}^n$ satisfying $s_i = 1$ if $\underline{x}^0_i \geq 0$, $s_i = -1$ if $\overline{x}^0_i \leq 0$, and $s_i = \pm 1$ otherwise. This approach is much faster in average, but may fail in the case we cannot determine an initial enclosure. The initial enclosure may intersect all orthants, too.

Another hull method was proposed by Jansson (1997). It is based on the topological properties of the solution set Σ.

Theorem 3.51 (Jansson, 1997)**.** *When $\Sigma \neq \emptyset$, then exactly one of the following alternatives holds true:*

(1) $\boldsymbol{A}$ is regular, in which case Σ is bounded and connected.

(2) $\boldsymbol{A}$ is irregular, in which case each topologically connected component of Σ is unbounded.

Proof. (1) Due to regularity, $Ax = b$ has a unique solution $x_{A,b}$, which is a continuous function with respect to $A \in \boldsymbol{A}$ and $b \in \boldsymbol{b}$. Therefore the range is compact and connected.

(2) Let $x^0 \in \Sigma$ be arbitrary, so there are $A^0 \in \boldsymbol{A}$, and $b^0 \in \boldsymbol{b}$ such that $A^0x^0 = b^0$. Suppose to the contrary that the component $\mathcal{K}$ containing x^0 is bounded. Hence A^0 must be nonsingular.

Since $\boldsymbol{A}$ is irregular, there is a singular $A^1 \in \boldsymbol{A}$. Define $A(\lambda) \coloneqq (1-\lambda)A^0 + \lambda A^1 \in \boldsymbol{A}$ with $\lambda \in [0,1]$, and

$$\lambda^* \coloneqq \inf\{\lambda \in [0,1];\ A(\lambda) \text{ is singular}\}.$$

We have $\lambda^* > 0$ since $A(0) = A^0$ is nonsingular, so any sufficiently close matrix is nonsingular, too. The solution $x(\lambda)$ to the system $A(\lambda)x = b^0$ depends continuously on $\lambda \in [0, \lambda^*)$, so $x(\lambda) \in \mathcal{K}$ for any $\lambda \in [0, \lambda^*)$. Take a sequence $\lambda_i \in [0, \lambda^*)$, $i = 1, \ldots,$ converging to λ^*, and select a subsequence such that $x(\lambda_i)$ converges to some $x^* \in \mathcal{K}$. Now,

$$A(\lambda^*)x^* = \lim_{i\to\infty} A(\lambda_i) \lim_{i\to\infty} x(\lambda_i) = \lim_{i\to\infty} A(\lambda_i)x(\lambda_i) = b^0.$$

Since $A(\lambda^*)$ is singular, the solution set to $A(\lambda^*)x^* = b^0$ is unbounded and therefore also $\mathcal{K}$. □

Based on the above topological characteristics of Σ, the hull method works as follows: Find any solution $x^c \in \Sigma$, for instance, by solving the midpoint system $A^cx = b^c$. Next, check whether or not the component $\mathcal{K}$ containing x^c is bounded.

More precisely, let $s = \operatorname{sgn}(x^c)$. Determine the interval hull, denoted by $\boldsymbol{x}^s$, of the part of Σ that lies in the orthant determined by the sign vector s. Provided $\boldsymbol{x}^s$ is unbounded, we stop; otherwise, for each $i \in \{1, \ldots, n\}$ do the following. If $\underline{x}^s_i = 0$, then we have to inspect the neighboring ortant that corresponds to the sign vector $(s_1, \ldots, s_{i-1}, -1, s_{i+1}, \ldots, s_n)$. Similarly, if $\overline{x}^s_i = 0$, then we have to inspect the neighboring ortant corresponding to the sign vector $(s_1, \ldots, s_{i-1}, 1, s_{i+1}, \ldots, s_n)$. In this way, we sequentially go through the orthants until we either decide on unboundedness or inspect all orthants in which $\mathcal{K}$ is lying.

Rohn's algorithm

Another approach to hull computation is due to Rohn (1989a, 2006a, 2012a); in a slightly different fashion described also in (Neumaier, 1990, Chap. 6).

Theorem 3.52 (Rohn, 1989a)**.** *Let $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$ be regular and $\boldsymbol{b} \in \mathbb{IR}^n$. For each $y \in \{\pm 1\}^n$, the system $A^cx - \operatorname{diag}(y)A^\Delta|x| = b_y$ has a unique solution solution $x_y \in \Sigma$ and*

$$\operatorname{conv}\Sigma = \operatorname{conv}\{x_y;\ y \in \{\pm 1\}^n\}.$$

Proof. See Rohn (1989a, 2006a). □

The points x_y, $y \in \{\pm 1\}^n$, are thus important points of the solution set. Geometrically, they are the extremal vertices of Σ – the spires of the star-looking solution sets; see Figures 3.1, 3.2, 3.4 and 3.5. We will consider these points also in Section 4.4.1, where they arise from another perspective.

These points relate to another result by Rohn (2012b): For each $y \in \{\pm 1\}^n$, the whole segment with endpoints x_y and x_{-y} lies in Σ.

For solving the absolute value equation $A^cx - \operatorname{diag}(y)A^\Delta|x| = b_y$, Rohn proposed *The sign accord algorithm* (Rohn, 1989a, 2012a). The basic idea is the following. If we know the sign vector $z = \operatorname{sgn}(x)$ of the solution, we can easily solve the equation. Using $|x| = \operatorname{diag}(z)x$, it reduces to the linear equations $A_{yz}x = b_y$. Since typically we do not know the signs in advance, we will solve the equation iteratively. We can start, e.g., with $z \coloneqq \operatorname{sgn}((A^c)^{-1}b^c)$, and while z does not meet the signs of the solution to $A_{yz}x = b_y$, we switch the violated z_i (i.e., $z_i \neq \operatorname{sgn}(x_i)$) with the smallest index i, and solve $A_{yz}x = b_y$. The pseudocode of the method is displayed in Algorithm 3.1.

Rohn (1989a) proved that the algorithm is finite for each $y \in \{\pm 1\}^n$ and each starting $z \in \{\pm 1\}^n$ as long as $\boldsymbol{A}$ is regular. Nevertheless, if we switch the signs of z in other way, there is less known; cf. open problems in Chapter 11.

Algorithm 3.1 Sign accord

Solve $A^c x = b^c$, and put $z \coloneqq \operatorname{sgn}(x)$,
solve $A_{yz} x = b_y$,
while $\exists i : z_i x_i < 0$ **do**
  put $k \coloneqq \min\{i;\ z_i x_i < 0\}$,
  put $z_k \coloneqq -z_k$,
  solve $A_{yz} x = b_y$,
end while
return x

Using the sign accord algorithm directly to determine $\square\Sigma$ is not suitable since we have to solve 2^n linear equations. Utilizing the signs of an initial enclosure of Σ is not so easy here. However, by using a dual approach, Rohn (2012a) proposed a not a priori exponential algorithm, which was implemented in VERSOFT.

Theorem 3.52 directly implies the following result stating that the extremal points of the solution set Σ are attained for the endpoints of $\boldsymbol{A}$ and $\boldsymbol{b}$. This reduces the problem to 2^{n^2+n} instances. A self-contained proof of the statement can also be found in Mayer (2017).

Corollary 3.53 (Beeck, 1972; Nickel, 1977; Harfiel, 1980)**.** *Let $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ be regular and $\boldsymbol{b} \in \mathbb{IR}^n$. Then*

$$\operatorname{conv}\Sigma = \operatorname{conv}\{A^{-1}b;\ |A - A^c| = A^\Delta,\ |b - b^c| = b^\Delta\}.$$

3.5.5 Easy case – inverse nonnegative matrices

Even though computing $\square\Sigma$ is *NP*-hard in general, there are some polynomially computable subclasses. One such subclass ($A^c = I_n$) was discussed in Section 3.5.1. Another easy-to-solve subclass is that with inverse nonnegative matrices, and, as a particular case, M-matrices. M-matrices naturally appear in diverse situations, e.g., in economics Jerrell (1996, 1997) and others; see Neumaier (1990) and references therein.

If $\boldsymbol{A}$ is an M-matrix, then $A^{-1} \geq 0$ for every $A \in \boldsymbol{A}$. This inverse nonnegativity holds also for other matrices, and, surprisingly, can easily be checked by handling two real matrices only (Kuttler, 1971). For other characterization, see Neumaier (1990); Rohn (1987, 2012b).

Theorem 3.54 (Kuttler, 1971)**.** *An interval matrix $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$ is inverse nonnegative if and only if $\underline{A}^{-1} \geq 0$ and $\overline{A}^{-1} \geq 0$. In which case, $\square\{A^{-1};\ A \in \boldsymbol{A}\} = [\overline{A}^{-1}, \underline{A}^{-1}]$.*

Proof. Let $A \in \boldsymbol{A}$. Then $\underline{A} \leq A \leq \overline{A}$ implies

$$\overline{A}^{-1} A \leq I_n \leq \underline{A}^{-1} A. \tag{3.20}$$

Denote $B \coloneqq \overline{A}^{-1} A$; we want to show that it is an M-matrix. Indeed, $B_{ij} \leq 0$ for $i \neq j$ from (3.20), and for $v \coloneqq \underline{A}^{-1} e > 0$ we have $Bv = \overline{A}^{-1} A v \geq \overline{A}^{-1} \underline{A} v = \overline{A}^{-1} e > 0$. Hence B is nonsingular and from $B = \overline{A}^{-1} A$ the matrix A is nonsingular, too. Thus, we can write $A^{-1} = B^{-1} \overline{A}^{-1} \geq 0$.

Eventually, multiplying (3.20) from right by A^{-1} we get $\overline{A}^{-1} \leq A^{-1} \leq \underline{A}^{-1}$. □

Example 3.55. Notice that $\{A^{-1};\ A \in \boldsymbol{A}\} = [\overline{A}^{-1}, \underline{A}^{-1}]$ does not hold in general even for M-matrices, as the following example illustrates:

$$\begin{pmatrix} 1 & a \\ -1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{1}{1+a} & -1 + \frac{1}{1+a} \\ \frac{1}{1+a} & \frac{1}{1+a} \end{pmatrix}, \quad a \in [-\tfrac{1}{2}, 0].$$

The interval hull of the inverse matrices is $\begin{pmatrix} [1,2] & [0,1] \\ [1,2] & [1,2] \end{pmatrix}$ but, e.g., $\left(\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right)$ is never attained. This example also shows that $[\overline{A}^{-1}, \underline{A}^{-1}]$ may contain a singular matrix. □

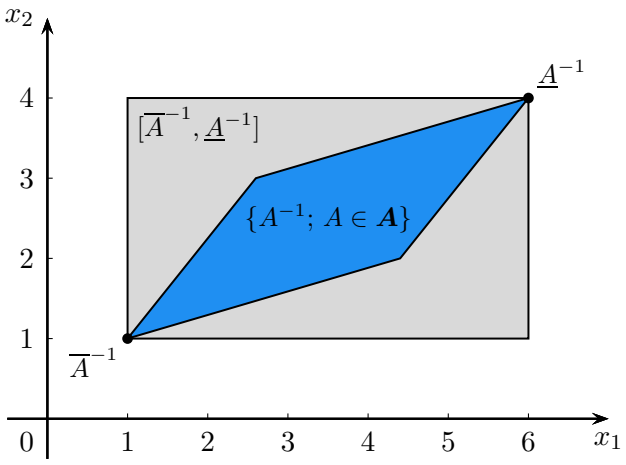


Figure 3.9: (Remark 3.56) Illustration of the set $\{A^{-1};\ A \in \boldsymbol{A}\}$ for inverse nonnegative interval matrices.

Remark 3.56. Figure 3.9 schematically illustrates the set of inverses $\{A^{-1};\ A \in \boldsymbol{A}\}$. The point $\overline{A}^{-1}$ is the entrywise minimum and $\underline{A}^{-1}$ is the maximum. Even though $\{A^{-1};\ A \in \boldsymbol{A}\}$ is not a box in general, it is a convex polyhedron. This is easy to see in view of Corollary 3.7 since the ith column of A^{-1} is the solution of the system $Ax = e_i$ and this solution lives in the nonnegative orthant for every $A \in \boldsymbol{A}$.

Inverse nonnegativity applies in solving linear equations in the following way.

Theorem 3.57. *Let $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$ be inverse nonnegative. Then*

(1) $\square\Sigma = [\overline{A}^{-1}\underline{b}, \underline{A}^{-1}\overline{b}]$ *when* $\underline{b} \geq 0$,

(2) $\square\Sigma = [\underline{A}^{-1}\underline{b}, \overline{A}^{-1}\overline{b}]$ *when* $\overline{b} \leq 0$,

(3) $\square\Sigma = [\underline{A}^{-1}\underline{b}, \underline{A}^{-1}\overline{b}]$ *when* $0 \in \boldsymbol{b}$.

Proof.

(1) Let $A \in \boldsymbol{A}$ and $b \in \boldsymbol{b}$. Since $\overline{b} \geq b \geq \underline{b} \geq 0$ and $\underline{A}^{-1} \geq A^{-1} \geq \overline{A}^{-1} \geq 0$, we get $\overline{A}^{-1}\underline{b} \leq A^{-1}b \leq \underline{A}^{-1}\overline{b}$.

(2) Analogous to (1).

(3) Let $A \in \boldsymbol{A}$ and $b \in \boldsymbol{b}$. Using $\overline{b} \geq 0 \geq \underline{b}$ and $\underline{A}^{-1} \geq A^{-1} \geq \overline{A}^{-1} \geq 0$, we get $\underline{A}^{-1}\underline{b} \leq A^{-1}\underline{b} \leq A^{-1}b \leq A^{-1}\overline{b} \leq \underline{A}^{-1}\overline{b}$. □

For a general $\boldsymbol{b} \in \mathbb{IR}^n$, the interval hull $\square\Sigma$ can be efficiently computed, too. The proof of the following result by Neumaier (1990) describes an iterative process that produces $\inf(\square\Sigma)$ in at most $n+1$ iterations; the case of $\sup(\square\Sigma)$ is analogous.

Theorem 3.58. *Let $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$ be inverse nonnegative. Then $\square\Sigma = [(A_{-e,s})^{-1}\underline{b}, (A_{e,s'})^{-1}\overline{b}]$, where $s = \operatorname{sgn}(\inf(\square\Sigma))$ and $s' = \operatorname{sgn}(\sup(\square\Sigma))$.*

Proof. Denote $A(x) \coloneqq A_{-e,s}$, where $s = \operatorname{sgn}(x)$. For every $A \in \boldsymbol{A}$, we then have

$$A(x)x \geq Ax. \tag{3.21}$$

First, we show that there is $x^* \in \mathbb{R}^n$ such that $A(x^*)x^* = \underline{b}$. Consider the iterations, producing vectors $x^0, x^1, \ldots$, where we start with $x^0 = \overline{A}^{-1}\underline{b}$ and then iterate $x^{i+1} = A(x^i)^{-1}\underline{b}$. In view of (3.21), we have $A(x^0)x^0 \geq \overline{A}x^0 = \underline{b}$, and for $i \geq 1$ we have $A(x^i)x^i \geq A(x^{i-1})x^i = \underline{b}$. Due to inverse nonnegativity, $x^i \geq A(x^i)^{-1}\underline{b} = x^{i+1}$. Hence $x^0 \geq x^1 \geq \ldots$ is a nonincreasing sequence. Since $A(x)$ changes only when some entry of x changes its sign, we have after at most $n+1$ iterations that $A(x^i)x^i = \underline{b}$ for some i.

Now, we show that if $x^* \in \mathbb{R}^n$ satisfies $A(x^*)x^* = \underline{b}$, then $x^* = \inf(\square\Sigma)$. Let $A \in \boldsymbol{A}$ and $b \in \boldsymbol{b}$ be arbitrary and denote $x \coloneqq A^{-1}b$. Then $Ax = b \geq \underline{b} = A(x^*)x^* \geq Ax^*$, whence $x \geq x^*$ by inverse nonnegativity. Therefore $x^* = \inf(\square\Sigma)$. □

Rohn (1989a) gives an alternative characterization of Σ when $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$ is inverse nonnegative:

$$\square\Sigma = [x_{-e}, x_e],$$

where x_{-e} and x_e are solutions of the absolute value equation from Theorem 3.52. Despite the elegance of this formulation, efficiency of its calculation is not clear.

Inverse sign pattern matrices

The above results can be easily generalized to the so-called inverse sign pattern matrices (Rohn, 1989a, 2012b). An interval matrix $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$ has the *inverse sign pattern* (y, z), where $y, z \in \{\pm 1\}^n$, if $\operatorname{diag}(y)A^{-1}\operatorname{diag}(z) \geq 0$ for all $A \in \boldsymbol{A}$.

In this case, $\operatorname{diag}(z)\boldsymbol{A}\operatorname{diag}(y)$ is inverse nonnegative, its left endpoint matrix is $\operatorname{diag}(z)A^c\operatorname{diag}(y) - A^\Delta$ and the right endpoint matrix is $\operatorname{diag}(z)A^c\operatorname{diag}(y) + A^\Delta$, so we have the following consequence of Kuttler's Theorem 3.54.

Corollary 3.59. *An interval matrix $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$ has the inverse sign pattern (y, z) if and only if $(\operatorname{diag}(z)A^c\operatorname{diag}(y) - A^\Delta)^{-1} \geq 0$ and $(\operatorname{diag}(z)A^c\operatorname{diag}(y) + A^\Delta)^{-1} \geq 0$.*

Analogously we proceed for an interval linear system $\boldsymbol{A}x = \boldsymbol{b}$ with $\boldsymbol{A}$ having the inverse sign pattern (y, z). We just employ the substitution $x' := \operatorname{diag}(y)x$ and solve the interval system $(\operatorname{diag}(z)\boldsymbol{A}\operatorname{diag}(y))x' = \operatorname{diag}(z)\boldsymbol{b}$ with the inverse nonnegative matrix $\operatorname{diag}(z)\boldsymbol{A}\operatorname{diag}(y)$.

Inverse stable matrices

Even more general class of interval matrices are those matrices whose inverse matrix entries keep their signs. Formally, $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$ is *inverse stable* if no A^{-1}, $A \in \boldsymbol{A}$, contains a zero entry. This class of matrices was studied, e.g., in Rohn (1993c, 2012a,b).

Notes

There are much more results and topics in interval linear systems that we did not mentioned. For instance, block interval methods (Garloff, 1990), sparse systems (Neumaier, 1990, references on p. 168).

3.6 Regularity of interval matrices

Recall that an interval matrix $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$ is *regular* if every matrix $A \in \boldsymbol{A}$ is nonsingular; otherwise, $\boldsymbol{A}$ is called *irregular* (Definition 3.15). Checking regularity of interval matrices is still a computationally hard problem; *NP*-hardness was proved by Poljak and Rohn (1988, 1993).

Theorem 3.60. *Checking regularity of an interval matrix is co-NP-hard on the class of interval matrices $\boldsymbol{A}$ with A^c non-negative positive definite rational and $A^\Delta = ee^T$.*

Proof. By Theorem 1.35, checking solvability of $|Ax| \leq e$, $e^T|x| \geq 1$ is *NP*-hard on the set of non-negative positive definite rational matrices A. The system is equivalent to

$$\left|\begin{pmatrix} A \\ 0 \end{pmatrix} x - \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right| \leq \begin{pmatrix} ee^T \\ e^T \end{pmatrix} |x| + \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Clearly, if x solves the former system, then it solves the latter. If x solves the latter, then $x \neq 0$ and $x' := \frac{1}{e^T|x|}x$ solves the former. By the Oettli–Prager Theorem 3.2, the latter system describes the solution set of the interval linear system

$$[A - ee^T, A + ee^T]x = 0, \tag{3.22a}$$

$$[-e^T, e^T]x = 1. \tag{3.22b}$$

If $[A - ee^T, A + ee^T]$ is regular, then (3.22a) forces all solutions to be zero, which contradicts (3.22b). If $[A - ee^T, A + ee^T]$ is irregular, then there is a non-zero solution x to (3.22a), which can be normalized by $x' := \frac{1}{e^T|x|}x$ to fulfill (3.22b). Thus, we reduced the *NP*-hard problem of checking solvability of a nonlinear system to checking irregularity of an interval matrix. □

NP-hardness remains true even for other special subclasses of interval matrices. For example, interval matrices with $A^\Delta = I_n$; this is due to *NP*-hardness of P-matrix property and by Rohn (2012c).

Forty necessary and sufficient conditions for regularity of interval matrices were surveyed by Rohn (2009); most of them are explicitly exponential formulae. Below, we state some of the conditions. The first one employs the Oettli–Prager description of interval linear equations. The second condition characterizes regularity by determinants of matrices of type A_{yz}. The third condition states, among others, that if $\boldsymbol{A}$ is not regular, then it contains a singular matrix of a special form $A^c - d\,\mathrm{diag}(y)A^\Delta\,\mathrm{diag}(z)$. The last condition is one of many regularity characterizations by an absolute value equation.

Theorem 3.61. *Necessary and sufficient conditions for regularity of* $\boldsymbol{A}$*:*

(1) The system $|A^c x| \le A^\Delta |x|$ *has the only solution* $x = 0$*.*

(2) $\det(A_{yz})$ *is constantly either positive or negative for each* $y, z \in \{\pm 1\}^n$*.*

(3) Each matrix of the form $A^c - d\,\mathrm{diag}(y)A^\Delta\,\mathrm{diag}(z)$*,* $d \in [0,1]$*,* $y, z \in \{\pm 1\}^n$ *is nonsingular.*

(4) For each $y \in \{\pm 1\}^n$*, the system* $A^c x - \mathrm{diag}(y)A^\Delta |x| = y$ *has a solution.*

Proof. We prove only the first three conditions. The other proofs and formulations are found in Rohn (2009) or in the references therein.

(1) It follows from the Oettli–Prager Theorem 3.2 applied on the interval system $\boldsymbol{A}x = 0$. The interval matrix $\boldsymbol{A}$ is regular if and only if the interval system $\boldsymbol{A}x = 0$ has only the trivial solution.

(2) One implication is trivial due to continuity of the determinant. We show the second implication by negation. Let $\boldsymbol{A}$ be irregular, so that there is a nonzero solution x to $|A^c x| \le A^\Delta |x|$. By the proof of the Oettli–Prager Theorem 3.2 we have $Ax = 0$ for a matrix A having the form of A_{uz} with $u \in [-1,1]^n$ and $z \in \{\pm 1\}^n$. Thus, $\det(A_{uz}) = 0$ and $\det(A_{uz})$ is a linear function in u_i for each $i \in \{1, \ldots, n\}$.

We claim that there exist $y, y' \in \{\pm 1\}^n$ such that $\det(A_{y,t}) \ge 0 \ge \det(A_{y',t})$. First, we set $u_1 := y_1 \in \{\pm 1\}$ such that $\det(A_{u,t})$ does not decrease. Subsequently, we fix $u_2, \ldots, u_n$ to the values $y_2, \ldots, y_n \in \{\pm 1\}$, which we can do due to linearity of u_is, and we arrive at $\det(A_{u,t}) \ge 0$. Analogously we get $\det(A_{u',t}) \le 0$.

Moreover, we can find $y, y' \in \{\pm 1\}^n$ that differ in only one entry. Simply construct a sequence $y^1, \ldots, y^m \in \{\pm 1\}^n$ such that $y^1 = y$, $y^m = y'$ and each pair of neighbors y^k, y^{k+1} differs in exactly one entry. Since $\det(A_{y^1,t}) \ge 0 \ge \det(A_{y^m,t})$, we must get $\det(A_{y^k,t}) \ge 0 \ge \det(A_{y^{k+1},t})$ for some $k \in \{1, \ldots, m-1\}$.

(3) One implication is trivial, and the second is proved by negation. Suppose that $\boldsymbol{A}$ is irregular. If $\det(A^c) = 0$ or $\det(A_{yz}) = 0$ for some $y, z \in \{\pm 1\}^n$, then we are done. Otherwise, by the previous result, there are $y, y', z \in \{\pm 1\}^n$ such that $\det(A_{yz})\det(A_{y'z}) < 0$. Without loss of generality suppose that $\det(A^c)\det(A_{yz}) < 0$. The function $\det(A^c - \alpha\,\mathrm{diag}(y)A^\Delta\,\mathrm{diag}(z))$ is continuous in α, so for some $\alpha \in [0,1]$ there is $\det(A^c - \alpha\,\mathrm{diag}(y)A^\Delta\,\mathrm{diag}(z)) = 0$. □

Kreinovich (2005) showed that the finite reduction in Theorem 3.61(2) is optimal in the sense that regularity of $\boldsymbol{A}$ cannot be checked in general by inspecting a subset of those matrices.

Remark 3.62. If $\boldsymbol{A}$ is irregular, then it contains a singular matrix A of special form. By Rohn (1993b), there are $d \in [0,1]$ and $y, z \in \{\pm 1\}^n$ such that $A := A^c - d\,\mathrm{diag}(y)A^\Delta\,\mathrm{diag}(z) \in \boldsymbol{A}$ is singular. Another special form singular matrix lies on an edge of $\boldsymbol{A}$ (Rohn, 1989a), i.e., there is (k, ℓ) such that $a_{k\ell} \in \boldsymbol{a}_{k\ell}$ and $a_{ij} \in \{\underline{a}_{ij}, \overline{a}_{ij}\}$ for $(i,j) \ne (k,\ell)$.

Another form will be presented in Corollary 5.32. Various special forms are recorded in Rohn (2009, 2012a).

3.6.1 Jansson & Rohn algorithm

Even though checking regularity of an interval matrix is *co-NP*-hard, there is an algorithm by Jansson and Rohn (1999) that is not a priori exponential. It was implemented by Hladík et al. (2011a) for computing real eigenvalues of interval matrices, and it turned out to be more effective than traditional exponential formulae.

The algorithm is based on the topological properties of the solution set of interval linear equations; see Section 3.5.4. Let $b \in \mathbb{R}^n$ be arbitrary. An interval matrix $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$ is regular if and only if the solution set Σ to the interval system $\boldsymbol{A}x = b$ is non-empty and bounded. Boundedness of Σ can be checked by the Jansson algorithm. Recall that this method calculates an initial $x^0 \in \Sigma$ and then check boundedness of the connected component $\mathcal{K}$ of Σ containing x^0. First, we check the orthant containing x^0. Next, we inspect the neighboring orthants having nonempty intersection with $\mathcal{K}$ and so on.

The vector $b \in \mathbb{R}^n$ can be set, e.g., as $b \coloneqq A^c e$. In this case, Σ is not empty since $e \in \Sigma$. More sophisticated heuristic of selection of b is dealt with in Jansson and Rohn (1999). Notice that the choice of b is important for the performance of the method as for various $b \in \mathbb{R}^n$ the component $\mathcal{K}$ intersects various number of orthants and thus influences the computing time.

3.6.2 Sufficient conditions

Since checking regularity of $\boldsymbol{A}$ is intractable, one often utilizes sufficient regularity conditions. Sufficient conditions are surveyed in Rex and Rohn (1998). First, we observe that preconditioning works here, too.

Lemma 3.63. *Let $C \in \mathbb{R}^{n\times n}$ and $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$. If $C\boldsymbol{A}$ is regular, then $\boldsymbol{A}$ is regular.*

Proof. Since $C\boldsymbol{A}$ is regular, every matrix in the form CA, where $A \in \boldsymbol{A}$, is nonsingular. Thus, both matrices C and A are nonsingular. This holds for every $A \in \boldsymbol{A}$, so $\boldsymbol{A}$ is regular. □

Now, we inspect the special case of $A^c = I_n$. Later in Theorem 3.72, we will see that the sufficient condition below is also a necessary condition in this case.

Theorem 3.64 (Beeck, 1975)**.** *Suppose that $A^c = I_n$. If $\rho(A^\Delta) < 1$, then $\boldsymbol{A}$ is regular.*

Proof. For each $A \in \boldsymbol{A}$ we have $|A - A^c| = |A - I_n| \leq A^\Delta$. From Theorem 1.8 it follows $\rho(A - I_n) \leq \rho(A^\Delta) < 1$, so A has no zero eigenvalue and is nonsingular. Hence $\boldsymbol{A}$ is regular. □

Condition $A^c = I_n$ can be achieved by preconditioning with $(A^c)^{-1}$. Indeed, interval matrix $\boldsymbol{M} \coloneqq (A^c)^{-1}\boldsymbol{A}$ satisfies $M^c = I_n$ and $M^\Delta = |(A^c)^{-1}|A^\Delta$. In view of Lemma 3.63, we immediately have the following sufficient condition.

Theorem 3.65 (Beeck, 1975)**.** *If*

$$\rho(|(A^c)^{-1}|A^\Delta) < 1, \tag{3.23}$$

then $\boldsymbol{A}$ is regular.

Condition (3.23) is one of the most frequently used sufficient conditions for regularity. In fact, interval matrices satisfying the Beeck's condition (3.23) are called *strongly regular*.

It may not be obvious at first sight, but strongly regular matrices are closed under subset inclusion: If $\boldsymbol{A}$ is strongly regular and $\boldsymbol{B} \subseteq \boldsymbol{A}$, then $\boldsymbol{B}$ is strongly regular as well (Neumaier, 1990; Rohn, 2012b).

Implementation. We can view Beeck's condition (3.23) as the preconditioning $\boldsymbol{A}$ by $C \coloneqq (A^c)^{-1}$, and then checking regularity of $C\boldsymbol{A}$ by verifying $\rho(\mathrm{rad}(C\boldsymbol{A})) < 1$. As mentioned in Section 3.2.1, we hardly know the midpoint inverse exactly. Hence we do preconditioning by an approximation $C \approx (A^c)^{-1}$, which gives an interval matrix the midpoint of which is only an approximation of I_n. Thus, for a safe implementation, we use $\mathrm{mag}(I_n - C\boldsymbol{A})$ instead of $\mathrm{rad}((A^c)^{-1}\boldsymbol{A})$; see Proposition 3.20. The numerical version of the condition then reads

$$\rho(\mathrm{mag}(I_n - C\boldsymbol{A})) < 1. \tag{3.24}$$

This condition is equivalent to the commonly mentioned form (Rump, 1983; Rex and Rohn, 1998)

$$\rho(|I_n - CA^c| + |C|A^\Delta) < 1$$

since

$$\begin{aligned}\operatorname{mag}(I_n - C\boldsymbol{A}) &= \operatorname{mag}(I_n - CA^c + C[-A^\Delta, A^\Delta]) \\ &= \operatorname{mag}(I_n - CA^c) + \operatorname{mag}(C[-A^\Delta, A^\Delta]) = |I_n - CA^c| + |C|A^\Delta.\end{aligned}$$

The interesting property of (3.24) is that it is not theoretically stronger than the Beeck's condition since the provably optimal preconditioner is $C \coloneqq (A^c)^{-1}$.

Theorem 3.66 (Rex and Rohn, 1995)**.** *If (3.24) holds true for some $C \in \mathbb{R}^{n\times n}$, then A^c is nonsingular and*

$$\rho(|(A^c)^{-1}|A^\Delta) \le \rho(\operatorname{mag}(I_n - C\boldsymbol{A})) = \rho(|I_n - CA^c| + |C|A^\Delta).$$

Proof. By Theorem 1.8, we have

$$\rho(I_n - CA^c) \le \rho(|I_n - CA^c|) \le \rho(|I_n - CA^c| + |C|A^\Delta) < 1. \tag{3.25}$$

Thus, CA^c is nonsingular, whence A^c is nonsingular as well.

Now, define

$$G \coloneqq |I_n - CA^c| + |C|A^\Delta + \varepsilon ee^T, \quad \alpha \coloneqq \rho(G) < 1,$$

where $\varepsilon > 0$ is small enough. Since $G > 0$, by Perron–Frobenius Theorem 1.7 there is $x > 0$ such that $Gx = \alpha x$. Using also that $\alpha < 1$, we derive

$$\alpha|I_n - CA^c|x + |C|A^\Delta x \le |I_n - CA^c|x + |C|A^\Delta x < \alpha x,$$

and from this

$$|C|A^\Delta x < \alpha(I_n - |I_n - CA^c|)x.$$

By Theorem 1.10 and (3.25), $I_n - |I_n - CA^c|$ has a nonnegative inverse, which yields

$$(I_n - |I_n - CA^c|)^{-1}|C|A^\Delta x < \alpha x.$$

Now, from

$$(A^c)^{-1} = (I_n - (I_n - CA^c))^{-1}C = \sum_{i=0}^{\infty}(I_n - CA^c)^i C$$

we derive

$$|(A^c)^{-1}| \le \sum_{i=0}^{\infty}|I_n - CA^c|^i|C| = (I_n - |I_n - CA^c|)^{-1}|C|.$$

Putting all together, we obtain

$$|(A^c)^{-1}|A^\Delta x \le (I_n - |I_n - CA^c|)^{-1}|C|A^\Delta x < \alpha x.$$

By Theorem 1.9, $\rho(|(A^c)^{-1}|A^\Delta) < \alpha$, from which the statement follows due to continuity of the spectral radius (Horn and Johnson, 1985; Meyer, 2000). □

Remarks. Notice that by Theorem 1.19, for any nonnegative A the condition $\rho(A) < 1$ is equivalent to $(I_n - A)^{-1} \ge 0$. The latter form can be more convenient from the computational and verification viewpoints. In this way, we can replace the spectral radius computations from the above theorems by matrix inverse calculations.

The strength of Beeck's sufficient condition was studied by Rump (1997a,b). He showed that every regular $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$ satisfies $\rho(|(A^c)^{-1}|A^\Delta) < (3 + 2\sqrt{2})n$, and for any $n \in \mathbb{N}$ and $\varepsilon > 0$, there is regular $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$ such that $\rho(|(A^c)^{-1}|A^\Delta) > n - \varepsilon$. He also proved that the upper bound can be tightened to $\rho(|(A^c)^{-1}|A^\Delta) < n$ for A^Δ of rank one, and conjectures that this remains true for any matrix; cf. page 189.

Theorem 1.5 gives rise yet to another alternative. Since $\rho(A) \le \|A\|$ for any matrix norm, we can replace the spectral radii by matrix norms in the above theorems. We obtain slightly weaker conditions for checking regularity of interval matrices, but with lower computational effort; some matrix norms, e.g., the maximum one, are very easily calculated.

Relation to interval equations. It is surprising that solving interval equations is computationally hard even for regular matrices. Morevover, it stays hard even a special kind of strongly regular matrices (Rohn and Kreinovich, 1995); we show the proof by Fiedler et al. (2006).

Theorem 3.67 (Rohn and Kreinovich, 1995)**.** *Computing the interval hull of the solution set to $\boldsymbol{A}x = \boldsymbol{b}$ is NP-hard on a class of problems with $\rho(|(A^c)^{-1}|A^\Delta) = 0$.*

Proof. We proceed similarly as in the proof of Theorem 3.12. We know from Theorem 1.35 that checking solvability of

$$-e \leq Az \leq e,\ e^T|z| \geq 1 \tag{3.26}$$

is *NP*-hard on the set of non-negative positive definite matrices. Consider the interval system $\boldsymbol{A}x = \boldsymbol{b}$, where

$$\boldsymbol{A} = \begin{pmatrix} 1 & [-e,e]^T \\ 0 & A \end{pmatrix},\quad \boldsymbol{b} = \begin{pmatrix} 0 \\ [-e,e] \end{pmatrix},\quad x = \begin{pmatrix} y \\ z \end{pmatrix}.$$

It is easy to see that $\rho(|(A^c)^{-1}|A^\Delta) = 0$ since

$$|(A^c)^{-1}|A^\Delta = \begin{pmatrix} 0 & e^T \\ 0 & 0 \end{pmatrix}.$$

Let $\boldsymbol{x}^T = (\boldsymbol{y}, \boldsymbol{z}^T)$ be the interval hull of the solution set of $\boldsymbol{A}x = \boldsymbol{b}$. We will show that checking $\overline{y} \geq 1$ is *NP*-hard. By the Oettli–Prager Theorem 3.2, the value of $\overline{y}$ can be expressed as

$$\overline{y} = \max\{y;\ |y| \leq e^T|z|,\ |Az| \leq e\}.$$

We have that $\overline{y} \geq 1$ if and only if (3.26) is solvable, which completes the proof. □

As a simple consequence of *co-NP*-hardness of checking regularity we have that it is also hard to decide on boundedness of the solution set (Rohn, 1994c).

Corollary 3.68 (Rohn, 1994c)**.** *Deciding whether the solution set of $\boldsymbol{A}x = \boldsymbol{b}$ is bounded is a co-NP-hard problem.*

Proof. By reduction of checking regularity of an interval matrix $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$. Consider the interval system $\boldsymbol{A}x = 0$. If $\boldsymbol{A}$ is regular, then the solution set is $\{0\}$. Otherwise, the solution set contains a line, and so it is unbounded. □

Another consequence is intractability of checking if all solutions of an interval linear system of equations $\boldsymbol{A}x = \boldsymbol{b}$ are nonnegative. It is an open problem if it remains intractable also in the case $\boldsymbol{A}$ is regular.

Corollary 3.69. *Deciding whether $\Sigma \subseteq \{x \geq 0\}$ is a co-NP-hard problem.*

Proof. Consider the interval system $\boldsymbol{A}x = 0$. When $\boldsymbol{A}$ is regular, then $\Sigma = \{0\} \subseteq \{x \geq 0\}$. When $\boldsymbol{A}$ is irregular, then there are $A \in \boldsymbol{A}$ and $x \neq 0$ such that $Ax = 0$. Thus, $\{x, -x\} \subseteq \Sigma$ and Σ cannot liy in the nonnegative orthant only. □

3.6.3 Necessary conditions

Necessary conditions for regularity of an interval matrix $\boldsymbol{A}$ help in checking regularity in the sense that their negations work as sufficient conditions for irregularity; they can cheaply verify that $\boldsymbol{A}$ is irregular.

The following proposition is a direct consequence of Theorem 3.1.

Proposition 3.70. *If $0 \in \boldsymbol{A}x$ for some $x \in \mathbb{R}^n \setminus \{0\}$, then $\boldsymbol{A}$ is irregular.*

Here, one must be careful when implementing the condition $0 \in \boldsymbol{A}x$ on a computer, and choose the right rounding mode. Other possibility is to check the Oettli–Prager inequality $|A^c x| \leq A^\Delta|x|$.

A promising candidate for x is to take a column of $(A^c)^{-1}$. This leads to the following corollary, which is originally from Rohn (1989a).

Corollary 3.71. *If* $(A^\Delta|(A^c)^{-1}|)_{ii} \geq 1$ *for some* $i \in \{1,\dots,n\}$*, then* $\boldsymbol{A}$ *is irregular.*

Proof. We have $\boldsymbol{A}(A^c)^{-1} = [I_n - A^\Delta|(A^c)^{-1}|, I_n + A^\Delta|(A^c)^{-1}|]$, so its ith column includes the zero vector iff $(A^\Delta|(A^c)^{-1}|)_{ii} \geq 1$. □

Another sufficient condition for irregularity is $(|(A^c)^{-1}|A^\Delta)_{ii} \geq 1$ for some $i \in \{1,\dots,n\}$. It is a consequence of the above theorem applied to the interval matrix $\boldsymbol{A}^T$. This form, however, has a nice similarity with the Beeck's sufficient condition for regularity of $\boldsymbol{A}$.

For other necessary conditions see, e.g., Rex and Rohn (1998); Rump (1997b).

3.6.4 Special cases

The following result comes from Rohn (2012b). It draws a class of matrices for which Beeck's sufficient condition is also necessary for regularity. It is particularly the case when $(A^c)^{-1} \geq 0$, or, more specifically, when $A^c = I_n$. In the latter case, $\boldsymbol{A}$ is regular if and only if $\rho(A^\Delta) < 1$.

Theorem 3.72. *Suppose that* $(A^c)^{-1} \geq 0$*. Then* $\boldsymbol{A}$ *is regular if and only if* $\rho(|(A^c)^{-1}|A^\Delta) < 1$*.*

Proof. One direction is obvious. To prove the converse, suppose to the contrary that $\boldsymbol{A}$ is regular, but $\rho^* := \rho(|(A^c)^{-1}|A^\Delta) \geq 1$. By the Perron–Frobenius Theorem 1.7, there is $x \neq o$ such that

$$(A^c)^{-1}A^\Delta x = |(A^c)^{-1}|A^\Delta x = \rho^* x,$$

or,

$$\left(I_n - \frac{1}{\rho^*}(A^c)^{-1}A^\Delta\right)x = o.$$

By premultiplying by A^c, we get

$$\left(A^c - \frac{1}{\rho^*}A^\Delta\right)x = o.$$

Since $(A^c - \frac{1}{\rho^*}A^\Delta) \in \boldsymbol{A}$, the interval matrix $\boldsymbol{A}$ is irregular; a contradiction. □

The statement can be slightly generalized. Regularity of $\boldsymbol{A}$ is equivalent to regularity of $\operatorname{diag}(y)\boldsymbol{A}\operatorname{diag}(z)$ for any $y, z \in \{\pm 1\}^n$. Thus, we have the following result.

Corollary 3.73. *Suppose there are* $y, z \in \{\pm 1\}^n$ *such that* $\operatorname{diag}(z)(A^c)^{-1}\operatorname{diag}(y) \geq 0$*. Then* $\boldsymbol{A}$ *is regular if and only if* $\rho(|(A^c)^{-1}|A^\Delta) < 1$*.*

Another special case is related to H-matrices (Neumaier, 1990, Prop. 4.1.7).

Theorem 3.74. *Let* A^c *be an M-matrix. Then* $\boldsymbol{A}$ *is regular if and only if it is an H-matrix.*

Proof. One direction is obvious as each H-matrix is nonsingular. To prove the converse, suppose to the contrary that there is $A_1 \in \boldsymbol{A}$ not being an H-matrix. Consider the matrix $A(t) := A^c + t(A_1 - A^c) \in \boldsymbol{A}$ depending on $t \in [0,1]$. Define $t_0 := \inf_{t\in[0,1]}$ such that $A(t)$ is not an H-matrix. Obviously $A(t_0)$ is not an H-matrix; otherwise, there is $v > 0$ such that $\langle A(t_0)\rangle > 0$ and this would hold even for a slightly increased t_0. Since $t_0 > 0$, we have that $A(t_0 - \varepsilon)$ is an H-matrix for any sufficiently small $\varepsilon > 0$. Since A^c has a positive diagonal and nonpositive offdiagonal, we have that $\langle A(t_0 - \varepsilon)\rangle$ is an M-matrix that lies in $\boldsymbol{A}$. Hence the real eigenvalues of $\langle A(t_0 - \varepsilon)\rangle$ are positive, and the real parts of the complex eigenvalues are positive as well. Therefore, from continuity of eigenvalues, $\langle A(t_0)\rangle \in \boldsymbol{A}$ must have a zero eigenvalue, which is a contradiction with regularity of $\boldsymbol{A}$. □

Notes and further reading. For tridiagonal interval matrices, Bar-On (2000); Bar-On et al. (1996) presented a method based on LU decomposition, which makes regularity checking effective in this case. Regularity of interval matrices finds an application also in robotics, e.g., in identification of singularities in the workspace; see Remark 5.34.

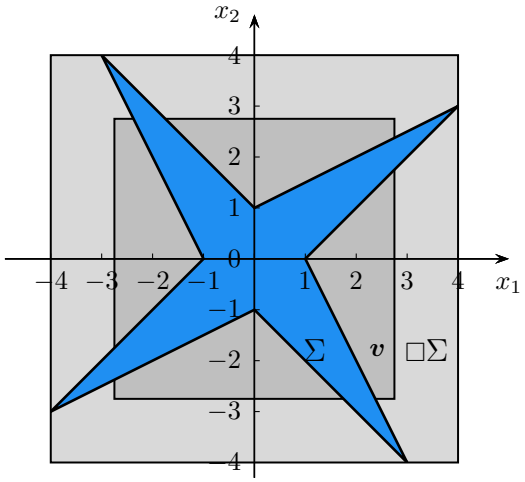


Figure 3.10: An inner estimation $\boldsymbol{v}$ of the solution set Σ in dark gray; the interval hull $\square\Sigma$ in light gray.

3.7 Inner estimation

By *an inner estimation* of the solution set Σ we mean any interval vector $\boldsymbol{v} \in \mathbb{IR}^n$ such that $\boldsymbol{v} \subseteq \square\Sigma$. That is, it is a subset of the interval hull of Σ, but not of the solution set itself; see Figure 3.10.

The advantage of an inner estimation is that it provides us with the quality of an interval enclosure $\boldsymbol{x} \supseteq \Sigma$. For each $i \in \{1, \dots, n\}$, $\overline{x}_i - \overline{v}_i$ and $\underline{v}_i - \underline{x}_i$ are the maximal overestimations of $\boldsymbol{x}$ in the ith coordinate from above and from below, respectively.

We already presented an inner estimation implicitly in Proposition 3.35 on the quality of Krawczyk iterations. Indeed, it yields an inner estimation $\boldsymbol{v}$ in the form

$$\begin{aligned}\underline{v} &= \inf(\boldsymbol{b} + (I_n - \boldsymbol{A})x^c) + \operatorname{mag}(I_n - \boldsymbol{A})x^\Delta,\\ \overline{v} &= \sup(\boldsymbol{b} + (I_n - \boldsymbol{A})x^c) - \operatorname{mag}(I_n - \boldsymbol{A})x^\Delta,\end{aligned}$$

where $\boldsymbol{x} \in \mathbb{IR}^n$ is any enclosure of Σ.

Notes and further reading. Inner estimations were initiated by Neumaier (1987), but we followed more practical formulae by Rump (1990); see also Mayer (2017); Rump (2010) for various adaptations.

3.8 Overdetermined systems

Consider an interval linear system

$$\boldsymbol{A}x = \boldsymbol{b}, \tag{3.27}$$

where $\boldsymbol{A} \in \mathbb{IR}^{m\times n}$, $\boldsymbol{b} \in \mathbb{IR}^m$ and $m > n$.

3.8.1 Least square solutions

(Neumaier, 1990, p. 169)

The least square method is a well-known method to approximate a solution to overdetermined linear equations $Ax = b$. The least square solution minimizes the difference between the right and left hand sides in Euclidean norm, that is

$$\min_{x\in\mathbb{R}^n} \|Ax - b\|_2.$$

It is known that the solution can equivalently be characterized as a solution to the square system

$$A^T Ax = A^T b,$$

or, alternatively, as x-solution to the system

$$\begin{pmatrix} 0 & A^T \\ A & I_m \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ b \end{pmatrix}. \tag{3.28}$$

Moreover, the least square solution has nice statistical properties. It is the best linear unbiased estimator provided the errors have a mean of zero, are uncorrelated, and have equal finite variances. Under these assumptions, it is also an asymptotically consistent estimator, that is, as the number of data points increases, the resulting sequence of estimates converges in probability to the true parameters.

The least square solution set to (3.27) is defined as the set of all least square solutions

$$\Sigma_{LSQ} \coloneqq \{x \in \mathbb{R}^n;\ \exists A \in \boldsymbol{A}\, \exists b \in \boldsymbol{b} : A^T A x = A^T b\}.$$

The solution set is always non-empty, however, checking its boundedness is *co-NP*-hard (Černý et al., 2013) even for real right-hand side systems ($b^\Delta = 0$).

Since Σ_{LSQ} is defined as the solution set of a square interval system of size n with nonlinear dependencies, we can employ any method from Section 3.9 to compute an enclosure to Σ_{LSQ}. By (3.28), we have an equivalent characterization

$$\Sigma_{LSQ} = \{x \in \mathbb{R}^n;\ \exists A \in \boldsymbol{A}\, \exists b \in \boldsymbol{b} \exists y \in \mathbb{R}^m : A^T y = 0,\ Ax + y = b\}.$$

It represents a solution set to a square interval system of size $m+n$ with linear dependencies. Moreover, it belongs to the symmetric case, so approaches from Section 3.9.2 are applicable. This form was discussed, e.g., in Karlík (2008); Neumaier (1986); Rump (1983) and extended to interval data with linear dependencies in Popova (2006c). Despite the larger dimension, this form usually yields much tighter enclosures (Černý et al., 2013). Moreover, when the systems are relaxed, the second form yields provably tighter enclosures.

Proposition 3.75. *Relaxing the dependencies, the solution set to* $\boldsymbol{A}^T\boldsymbol{A}x = \boldsymbol{A}^T\boldsymbol{b}$ *is contained in the solution set to*

$$\begin{pmatrix} 0 & \boldsymbol{A}^T \\ \boldsymbol{A} & I_m \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ \boldsymbol{b} \end{pmatrix}. \tag{3.29}$$

Proof. Any instance of the system (3.29) reads

$$A^T y = 0,\ A'x + y = b$$

for some $A, A' \in \boldsymbol{A}$ and $b \in \boldsymbol{b}$. Now, substituting $y = b - A'x$ into the first equation gives $A^T(b - A'x) = 0$, whence $A^T A'x = A^T b$. Since $A^T A' \in \boldsymbol{A}^T\boldsymbol{A}$ and $A^T b \in \boldsymbol{A}^T\boldsymbol{b}$, we are done. □

Notes. There are other alternative approaches to solve the problem (Markov, 1990). Some of them adapt the QR factorization for interval data (Gay, 1988), for instance, by using interval Householder's method (Bentbib, 2002; Manteuffel, 1981). A monotonicity checking, similar to the one presented in Section 3.5.3, but extended to the least square solutions, was discussed in Deif (1986); Shary and Moradi (2021).

If $\boldsymbol{A} = A$ is point and has full column rank, then $\Sigma_{LSQ} = \{(A^T A)^{-1} A^T b;\ b \in \boldsymbol{b}\}$ is a zonotope (see page 33), and its interval hull is computable directly by interval arithmetic as $\square\Sigma_{LSQ} = (A^T A)^{-1} A^T \boldsymbol{b}$. Inner and outer estimations by ellipsoids, and enumeration of the vertices and facets of the zonotope were thoroughly investigated by Černý et al. (2013).

Applications in electrical networks are discussed in Deif (1986), and in stock market forecasting in Hu and He (2007); He and Hu (2009).

3.8.2 True solutions

Notes and further reading (Section 3.8). Underdetermined interval systems of linear equations also naturally appear in many situations. For example, in verification of optimization problems (Jansson et al., 2007); see Section 10.1.

3.9 Parametric interval systems

When solving interval linear systems of equations $\boldsymbol{A}x = \boldsymbol{b}$, we have so far assumed that the matrix and the right-hand side entries vary inside the corresponding intervals independently. However, this is hardly true for real-life problems. Usually, the system reads $A(p)x = b(p)$, where the entries of $A(p)$ and $b(p)$ depend on parameters $p_1, \dots, p_K$, whose domains are intervals $\boldsymbol{p}_1, \dots, \boldsymbol{p}_K$. Calculating the ranges (or their enclosures) of $\boldsymbol{A} \coloneqq A(\boldsymbol{p})$ and $\boldsymbol{b} \coloneqq b(\boldsymbol{p})$ leads to the relaxed system of standard interval linear equations $\boldsymbol{A}x = \boldsymbol{b}$, but the loss of dependence structure may cause a huge overestimation. Thus, appropriate methods for dependent linear systems are of high importance.

The solution set of a parametric interval system is defined as

$$\Sigma_{\mathrm{p}} = \{x \in \mathbb{R}^n;\ A(p)x = b(p) \text{ for some } p \in \boldsymbol{p}\}.$$

It may have a complicated structure and finding a tight enclosure is a very challenging task.

In general, we may apply preconditioning

$$CA(\boldsymbol{p})x = Cb(\boldsymbol{p}).$$

Again, we may simply relax the system to standard interval linear equations, which is useful provided we are able to somehow simplify and tightly evaluate the entries of $CA(\boldsymbol{p})$ and $Cb(\boldsymbol{p})$.

A promising approach is to employ the residual form (Section 3.2.2). Let $x^* \in \mathbb{R}^n$, for example, the solution of $A(p^c)x = b(p^c)$. Then we find an enclosure to Σ_{p} in the form of $x^* + \boldsymbol{y}$, where $\boldsymbol{y}$ encloses the solution set to

$$CA(\boldsymbol{p})y = C(b(\boldsymbol{p}) - A(\boldsymbol{p})x^*). \tag{3.30}$$

Now, it depends on the special structure of dependencies how effectively we are able to enclose both terms $C(b(\boldsymbol{p}) - A(\boldsymbol{p})x^*)$ and $CA(\boldsymbol{p})$. Since the dependencies may be very complex, no general recipe can easily be given.

Remark 3.76. Applying this approach to the residual Krawczyk method, the direct Krawczyk enclosure (3.17) employed on our system works as follows. Let $R \coloneqq \operatorname{mag}(I_n - CA(\boldsymbol{p}))$. If $\rho(R) < 1$, then

$$\Sigma_{\mathrm{p}} \subseteq x^* + [-1, 1](I_n - R)^{-1} \operatorname{mag}(C(b(\boldsymbol{p}) - A(\boldsymbol{p})x^*)).$$

This type of an enclosure was originally proposed by Skalna (2012).

As it is very hard to handle the parametric systems in the general form, we focus on the particular case of a linear dependence structure.

Notes and further reading. General systems with nonlinear dependences between interval quantities were handled, e.g., by Garloff et al. (2009); Kolev (2004a); Popova (2007).

3.9.1 Linear dependencies

We will consider a linear parametric structure, that is, a system

$$A(p)x = b(p), \tag{3.31}$$

where $A(p) = \sum_{k=1}^{K} A^k p_k$, $b(p) = \sum_{k=1}^{K} b^k p_k$ and $p \in \boldsymbol{p}$ for some given interval vector $\boldsymbol{p} \in \mathbb{IR}^K$, matrices $A^1, \dots, A^K \in \mathbb{R}^{n \times n}$ and vectors $b^1, \dots, b^K \in \mathbb{R}^n$. This linear parametric case covers a wide area of interval systems with dependencies. For instance, it comprises the interval system $\boldsymbol{A}x = \boldsymbol{b}$, where the constraint matrix is supposed to be symmetric, skew-symmetric, circulant, Toeplitz or Hankel.

Example 3.77 (Standard interval matrix)**.** Obviously, linear parametric systems generalize the standard interval systems. For example, an interval matrix $\boldsymbol{A} \in \mathbb{IR}^{2\times 2}$ can be expressed as a linear interval parametric matrix $A(p)$, $p \in \boldsymbol{p} \in \mathbb{IR}^4$ as follows

$$A(p) = \begin{pmatrix} p_1 & p_2 \\ p_3 & p_4 \end{pmatrix} = \sum_{k=1}^{4} p_k A^k = p_1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + p_2 \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + p_3 \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + p_4 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

so that we associate $\boldsymbol{p}_1 = \boldsymbol{a}_{11}$, $\boldsymbol{p}_2 = \boldsymbol{a}_{12}$, $\boldsymbol{p}_3 = \boldsymbol{a}_{21}$ and $\boldsymbol{p}_4 = \boldsymbol{a}_{22}$. □

Example 3.78 (Symmetric interval matrix)**.** Symmetric interval matrices have a linear interval parametric form, too. For example, any 2×2 symmetric interval matrix $\boldsymbol{A}^S$ can be expressed as a linear interval parametric matrix $A(p)$, $p \in \boldsymbol{p} \in \mathbb{IR}^3$ as follows

$$A(p) = \begin{pmatrix} p_1 & p_2 \\ p_2 & p_3 \end{pmatrix} = \sum_{k=1}^{3} p_k A^k = p_1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + p_2 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + p_3 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix},$$

where we associate $\boldsymbol{p}_1 = \boldsymbol{a}_{11}$, $\boldsymbol{p}_2 = \boldsymbol{a}_{12} = \boldsymbol{a}_{21}$ and $\boldsymbol{p}_3 = \boldsymbol{a}_{22}$. □

Example 3.79 (Circulant interval matrix)**.** Any 3×3 circulant interval matrix has the form of a linear interval parametric matrix

$$A(p) = \begin{pmatrix} p_1 & p_2 & p_3 \\ p_3 & p_1 & p_2 \\ p_2 & p_3 & p_1 \end{pmatrix} = \sum_{k=1}^{3} p_k A^k = p_1 \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + p_2 \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} + p_3 \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix},$$

where we associate $\boldsymbol{p}_1 = \boldsymbol{a}_{11}$, $\boldsymbol{p}_2 = \boldsymbol{a}_{12}$ and $\boldsymbol{p}_3 = \boldsymbol{a}_{13}$. □

Characterization of the solution set. It is a tempting problem to determine a simple characterization of Σ_{p}. By using a time consuming Fourier–Motzkin elimination (Alefeld et al., 2003; Popova, 2015) we describe the parametric solution set by a possibly double exponential number of nonlinear inequalities. The description, however, shows that the shape of Σ_{p} is characterized by polynomials. For special classes of parametric interval systems, the explicit description of the solution set was developed by Hladík (2008b); Popova (2009). Shapes of the particular solution sets were first analyzed in Alefeld et al. (1997, 2003). Deeper insight gives Section 5.2 in Mayer (2017).

The following necessary characterization of Σ_{p} is by Popova (2009); see also Hladík (2012d).

Theorem 3.80 (Popova, 2009)**.** *If $x \in \Sigma_{\mathrm{p}}$, then it solves*

$$|A(p^c)x - b(p^c)| \le \sum_{k=1}^{K} p_k^{\Delta} |A^k x - b^k|. \tag{3.32}$$

Proof. Let $x \in \mathbb{R}^n$ be a solution to $A(p)x = b(p)$ for some $p \in \boldsymbol{p}$. Then,

$$\begin{aligned} |A(p^c)x - b(p^c)| &= \left| \sum_{k=1}^{K} p_k^c (A^k x - b^k) \right| = \left| \sum_{k=1}^{K} p_k^c (A^k x - b^k) - \sum_{k=1}^{K} p_k (A^k x - b^k) \right| \\ &= \left| \sum_{k=1}^{K} (p_k^c - p_k)(A^k x - b^k) \right| \le \sum_{k=1}^{K} |p_k^c - p_k| |A^k x - b^k| \le \sum_{k=1}^{K} p_k^{\Delta} |A^k x - b^k|. \end{aligned}$$ □

Popova (2009) showed that it is the complete characterization of Σ_{p} as long as no interval parameter appears in more than one equation. Thus, it can serve as a generalization of the Oettli–Prager Theorem 3.2 and a basis for deriving direct enclosures to the parametric solution set; see Hladík (2012d).

Theorem 3.81 (Popova, 2009)**.** *Suppose that for every $k = 1, \dots, K$, the matrix $(A^k \mid b^k)$ has only one nonzero row. Then $x \in \Sigma_{\mathrm{p}}$ if and only if x solves (3.32).*

Proof. By Theorem 3.80, we have necessity of (3.32). Here, we show sufficiency. Let $i \in \{1, \dots, n\}$ and $\mathcal{K}_i \subseteq \{1, \dots, K\}$ the set of parameters appearing in the ith row. From (3.32), there is $\alpha \in [0, 1]$ such that

$$|A(p^c)_{i*}x - b(p^c)_i| = \sum_{k \in \mathcal{K}_i} \alpha p_k^{\Delta} |A_{i*}^k x - b_i^k|.$$

Hence we can find $q_k \in [-p_k^{\Delta}, p_k^{\Delta}]$, $k \in \mathcal{K}_i$ such that

$$A(p^c)_{i*}x - b(p^c)_i = \sum_{k \in \mathcal{K}_i} q_k (A_{i*}^k x - b_i^k),$$

which after rearrangement reads

$$\sum_{k\in\mathcal{K}_i} A^k_{i*}(p^c_k - q_k)x = \sum_{k\in\mathcal{K}_i} b^k_i(p^c_k - q_k).$$

Therefore, x solves the realization $A(p)x = b(p)$ with $p = p^c - q \in \boldsymbol{p}$, where $q_k := 0$ for $k \notin \cup_i \mathcal{K}_i$. Since the sets $\mathcal{K}_i$, $i \in \{1,\ldots,n\}$, are mutually disjoint, there is no conflict in the selection of p. □

Even though (3.32) is only a necessary condition, it is a basis for the following characterization. From a certain viewpoint, the system (3.33) below is composed of a union of systems (3.32) over all possible preconditionings. An open question here is how to obtain $\square\Sigma_{\mathrm{p}}$ by a reduction to finitely many choices of $y \in \mathbb{R}^n$.

Theorem 3.82 (Hladík, 2012d)**.** *We have that $x \in \Sigma_{\mathrm{p}}$ if and only if it solves*

$$y^T\left(A(p^c)x - b(p^c)\right) \le \sum_{k=1}^{K} p^\Delta_k |y^T(A^k x - b^k)| \tag{3.33}$$

for every $y \in \mathbb{R}^n$.

Proof. Let $x \in \mathbb{R}^n$. Then $x \in \Sigma_{\mathrm{p}}$ if and only if there is a vector $q \in [-1,1]^K$ such that

$$A(p^c)x - b(p^c) = \sum_{k=1}^{K} q_k p^\Delta_k (A^k x - b^k).$$

Denote $d := A(p^c)x - b(p^c)$, and let $D \in \mathbb{R}^{n\times K}$ be a matrix whose kth column is equal to $p^\Delta_k(A^k x - b^k)$, $k = 1,\ldots,K$. Then $x \in \Sigma_{\mathrm{p}}$ if and only if there is a feasible solution of the linear system

$$Dq = d,\ -e \le q \le e,$$

or, in other words, if and only if the linear program

$$\max\ 0^T q\ \text{ subject to }\ Dq = d,\ -e \le q \le e$$

has an optimal solution. Consider the corresponding dual problem

$$\begin{aligned}&\min\ d^T y + e^T(u+v)\\ &\text{subject to }\ D^T y + u - v = 0,\ u,v \ge 0,\end{aligned}$$

which is always feasible. According to the theory of duality in linear programming (Theorem 1.32), existence of an optimal solution to one problem implies the same for the second one and the optimal values are equal.

For an optimal solution of the dual problem and every $i \in \{1,\ldots,K\}$ either $u_i = 0$ or $v_i = 0$; otherwise we can subtract a small positive amount from both u_i and v_i and decrease the optimal value. If $u_i = 0$ then $(u+v)_i = v_i = (D^T y)_i \ge 0$. Similarly, $v_i = 0$ implies $(u+v)_i = u_i = -(D^T y)_i \ge 0$. Hence we can derive $u + v = |D^T y|$, and the dual problem takes the form

$$\min\ d^T y + e^T|D^T y|\ \text{ subject to }\ y \in \mathbb{R}^n.$$

Since the objective function is positive homogeneous, the problem has an optimal solution (equal to zero) if and only if the objective function is non-negative, i.e.

$$d^T y + e^T|D^T y| \ge 0 \quad \forall y \in \mathbb{R}^n.$$

By substituting $y := -y$, we have $y^T d \le |y^T D|e$, and by substituting for D and d, we eventually obtain (3.33). □

Checking $x \in \Sigma_{\mathrm{p}}$. Even though it is hard to describe the solution set in general, checking $x \in \Sigma_{\mathrm{p}}$ for a given $x \in \mathbb{R}^n$ is a polynomial problem carried out by a suitable linear programming solver since the constraints $A(p)x = b(p)$, $p \in \boldsymbol{p}$ are linear. On the other hand, we cannot check it much more efficiently since it is a P-complete problem.

Theorem 3.83 (Hladík, 2020a)**.** *Checking $x^* \in \Sigma_{\mathrm{p}}$ is a P-complete problem (under NC-reduction).*

Proof. Recall that checking solvability of a linear system $Ax = b$, $x \geq 0$ is P-complete, where $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$. We reduce this system to our problem. First, rewrite it to

$$\sum_k (A_{*k}e_k^T)xp_k = b, \quad p \in \boldsymbol{p}, \tag{3.34}$$

where $\boldsymbol{p} = [0, c]^n$ and $c > 0$ is sufficiently large. Notice that $(A_{*k}e_k^T)e = A_{*k}$. Thus, the linear system $Ax = b$, $x \geq 0$ is solvable iff $x := e$ is a solution of (3.34). Since c can be chosen such that it has a polynomial size with respect to the input size (Schrijver, 1998), and the transformation is obviously in NC, the reduction is done. □

This result also shows that we can hardly hope for having a simple Oettli–Prager type characterization of Σ_{p}; it is believed that P-complete problems cannot have a closed form arithmetic characterization of a polynomial size. Notice that an analogous result can be stated for the symmetric case discussed in Section 3.9.2.

Enclosures of solutions.

Naive approach – direct relaxation. Relaxing the dependencies, we get an ordinary interval system of equations $\boldsymbol{A}x = \boldsymbol{b}$, where $\boldsymbol{A}$ and $\boldsymbol{b}$ are obtained by interval evaluation of particular entries as follows

$$\boldsymbol{A} \coloneqq A(\boldsymbol{p}) = \sum_{k=1}^{K} A^k \boldsymbol{p}_k, \quad \boldsymbol{b} \coloneqq b(\boldsymbol{p}) = \sum_{k=1}^{K} b^k \boldsymbol{p}_k.$$

The solution set of $\boldsymbol{A}x = \boldsymbol{b}$ includes Σ_{p}, so any enclosure of the relaxed system also encloses the solution set of the original system (3.31). This approach does not take into account the dependencies at all, and so the overestimation can be high.

Preconditioning. Many interval system solvers use preconditioning, however, it is not effective to precondition the relaxed system $C\boldsymbol{A}x = C\boldsymbol{b}$. Provably tighter relaxation is obtained by preconditioning the original system and after relaxing, which leads to the interval system $\boldsymbol{A'}x = \boldsymbol{b'}$ with

$$\boldsymbol{A'} \coloneqq \sum_{k=1}^{K} \boldsymbol{p}_k(CA^k), \quad \boldsymbol{b'} \coloneqq \sum_{k=1}^{K} \boldsymbol{p}_k(Cb^k).$$

Observation 3.84. *We have $\boldsymbol{A'} \subseteq \boldsymbol{A}$ and $\boldsymbol{b'} \subseteq \boldsymbol{b}$.*

Proof. From sub-distributivity of interval arithmetic, it simply follows

$$\sum_{k=1}^{K} (CA^k)\boldsymbol{p}_k \subseteq C \sum_{k=1}^{K} A^k \boldsymbol{p}_k,$$

and similarly for the right-hand side. □

Often, we use a numerical approximate of the midpoint inverse $C \approx A(p^c)^{-1}$ as a preconditioner. However, this need not be the best choice. In some sense optimal preconditioning related to the interval parametric Gauss–Seidel method was proposed in Hladík (2016b).

Algorithm 3.2 Framework for a linear parametric system of equations (3.31)

1: Choose a preconditioner $C \in \mathbb{R}^{n\times n}$, for instance $C \approx A(p^c)^{-1}$.
2: Choose a center $x^* \in \mathbb{R}^n$, for instance $x^* \approx A(p^c)^{-1}b(p^c)$.
3: Evaluate $\boldsymbol{A}' \coloneqq \sum_{k=1}^K \boldsymbol{p}_k(CA^k)$ and $\boldsymbol{b}' = \sum_{k=1}^K \boldsymbol{p}_k(C(b^k - A^k x^*))$.
4: Find an enclosure $\boldsymbol{y}$ of the solution set of $\boldsymbol{A}'y = \boldsymbol{b}'$ by a favourite solver.
5: **return** $\boldsymbol{x} \coloneqq x^* + \boldsymbol{y}$.

Residual form. As mentioned in the general case, it seems convenient to seek for an enclosure by using the preconditioned residual form (3.30), which takes the form

$$\Big(\sum_{k=1}^K \boldsymbol{p}_k(CA^k)\Big)y = \sum_{k=1}^K \boldsymbol{p}_k(C(b^k - A^k x^*)). \tag{3.35}$$

Due to the sub-distributivity law, the right-hand side interval vector of (3.35), and sometimes the interval matrix, too, are in general tighter than the simple relaxation $\boldsymbol{A} \coloneqq A(\boldsymbol{p})$, $\boldsymbol{b} \coloneqq b(\boldsymbol{p})$ and transformation to

$$(C\boldsymbol{A})y = C(\boldsymbol{b} - \boldsymbol{A}x^*).$$

On the other hand, using A^k and b^k, $k = 1,\dots,K$, explicitly may be time consuming. When K is large, the matrices A_k are often sparse, and one can think of evaluating $C(b^k - A^k x^*)$ without explicitly constructing A^k and b^k. This will be the case in Section 3.9.2 for the symmetric solution set, and similar rearrangements can also be performed for other specific parametric matrices such as Toeplitz or Hankel matrices. The idea of a delicate evaluation of the right-hand side vector is originally due to Jansson (1991). He applied it on a special parametric system by utilizing the residual Krawczyk operator.

The resulting approach. Evaluating the interval matrix and right-hand side, (3.35) becomes a standard interval linear system, which may be solved by any method presented in Section 3.2, depending on the requirements on the running time and tightness of the enclosure. Thus, the general framework for solving parametric interval systems is as described in Algorithm 3.2.

Remark 3.85. Some methods developed for solving the linear parametric system (3.31) are implicitly following the template of Algorithm 3.2. For example, the residual Krawczyk operator applied to the parametric system (3.35) takes the form

$$K_r(\boldsymbol{y}) = \sum_{k=1}^K C(b^k - A^k x^*)\boldsymbol{p}_k + \Big(I_n - \sum_{k=1}^K (CA^k)\boldsymbol{p}_k\Big)\boldsymbol{y}.$$

The direct enclosure formula (3.17) reads as follows. Let $R \coloneqq \operatorname{mag}\big(I_n - \sum_{k=1}^K (CA^k)\boldsymbol{p}_k\big)$. If $\rho(R) < 1$, then

$$\Sigma_{\mathrm{p}} \subseteq x^* + [-1,1](I_n - R)^{-1}\operatorname{mag}\big(\textstyle\sum_{k=1}^K C(b^k - A^k x^*)\boldsymbol{p}_k)\big).$$

In a slightly different fashion, this enclosure was presented in Hladík (2012d); Skalna (2006).

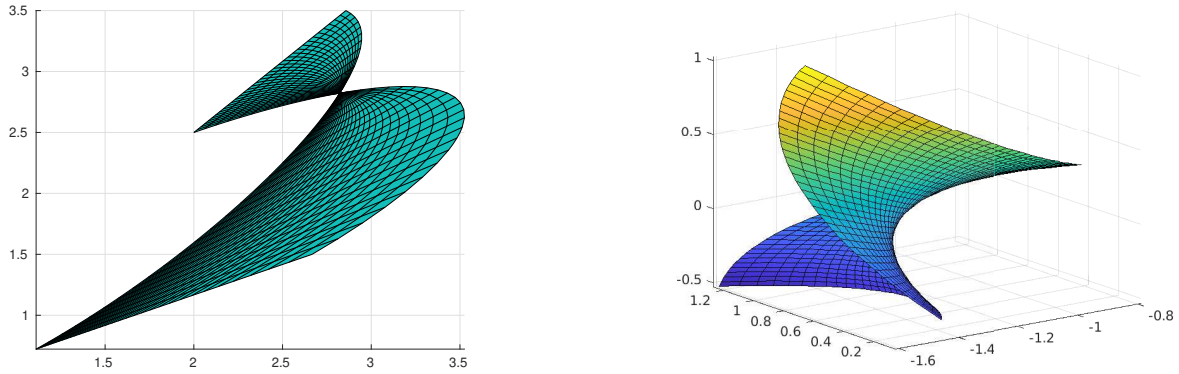
The Hansen–Bliek–Rohn enclosure for the system preconditioned by $A(p^c)^{-1}$ was addressed in Hladík (2012d) and the residual form in addition was discussed in Skalna (2018). A comparison of various methods with and without preconditiong and the residual form was performed in Hladík and Skalna (2019).

Example 3.86. Consider a two-dimensional linear parametric system $A(p)x = b(p)$ with

$$A(p) = \begin{pmatrix} 5+p_1 & -4-p_1+p_2 \\ -5-p_2 & 6 \end{pmatrix}, \quad b(p) = \begin{pmatrix} 4+p_2 \\ 1 \end{pmatrix}, \quad p \in \boldsymbol{p} = \begin{pmatrix} [-2,2] \\ [-2,2] \end{pmatrix}.$$

Figure 3.11a illustrates the area determined by the solution set Σ_{p}. The direct relaxation of the parametric system leads to a standard interval system $\boldsymbol{A}x = \boldsymbol{b}$, the solution set of which is unbounded since $\boldsymbol{A}$ is irregular.

Unfortunately, former preconditioning by $C = A(p^c)^{-1}$ also does not help here, producing again an irregular interval matrix. □



(a) (Example 3.86) Two-dimensional area in plane $\mathbb{R}^2$. (b) (Example 3.87) Two-dimensional surface in space $\mathbb{R}^3$.

Figure 3.11: Illustration of the solution set of a linear interval parametric system.

Example 3.87. Consider now a three-dimensional linear parametric system $A(p)x = b(p)$ with

$$A(p) = \begin{pmatrix} 6-p_2 & 5 & 2-p_1 \\ 2 & -4-p_1 & 4+2p_1+p_2 \\ -5+p_1-p_2 & -3 & 7-p_1 \end{pmatrix}, \quad b(p) = \begin{pmatrix} -5-p_1 \\ -5-p_2 \\ 5 \end{pmatrix}, \quad p \in \boldsymbol{p} = \begin{pmatrix} [-2,1.5] \\ [-2,1.5] \end{pmatrix}.$$

Figure 3.11b depicts the two-dimensional surface determined by the solution set Σ_{p}. Also in this example the direct relaxation of the parametric system produces an unbounded solution set. □

Linear parametric systems appear in diverse problems. We show an example from structure mechanics.

Example 3.88 (Displacements of a truss structure (Skalna, 2006))**.** Consider a 7-bar truss structure as depicted in Figure 3.12. Supposing a downward force at some nodes, the displacements of the nodes are computed by solving the linear system of equations

$$Kd = f, \tag{3.36}$$

where K is the stiffness matrix, and f is the vector of forces at particular nodes. The stiffness matrix has the structure

$$K = \begin{pmatrix} \frac{s_{12}}{2}+s_{13} & -\frac{s_{12}}{2} & -\frac{s_{12}}{2} & -s_{13} & 0 & 0 & 0 \\ -\frac{s_{21}}{2} & \frac{s_{21}+s_{23}}{2}+s_{24} & \frac{s_{21}-s_{23}}{2} & -\frac{s_{23}}{2} & \frac{s_{23}}{2} & -s_{24} & 0 \\ -\frac{s_{21}}{2} & \frac{s_{21}-s_{23}}{2} & \frac{s_{21}+s_{23}}{2} & \frac{s_{23}}{2} & -\frac{s_{23}}{2} & 0 & 0 \\ -s_{31} & -\frac{s_{32}}{2} & \frac{s_{32}}{2} & s_{31}+\frac{s_{32}+s_{34}}{2}+s_{35} & \frac{s_{34}-s_{32}}{2} & -\frac{s_{34}}{2} & -\frac{s_{34}}{2} \\ 0 & \frac{s_{32}}{2} & -\frac{s_{32}}{2} & \frac{s_{34}-s_{32}}{2} & \frac{s_{34}+s_{32}}{2} & -\frac{s_{34}}{2} & -\frac{s_{34}}{2} \\ 0 & -s_{42} & 0 & -\frac{s_{43}}{2} & -\frac{s_{43}}{2} & s_{42}+\frac{s_{43}+s_{45}}{2} & 0 \\ 0 & 0 & 0 & -\frac{s_{43}}{2} & -\frac{s_{43}}{2} & 0 & \frac{s_{43}+s_{45}}{2} \end{pmatrix}.$$

The stiffness of a particular bar (i,j) is computed as $s_{ij} = E\sigma/\ell$, where E is Young's modulus, σ is the cross-sectional area, and ℓ is the length of the bar. Since Young's modulus is usually given only approximately with some error ΔE, the stiffnesses s_{ij} of bars are uncertain and known to lie in some intervals only. Thus, (3.36) has a form of a linear parametric interval system. □

Notes and further reading. The first paper on parametric interval systems with a special structure is by Jansson (1991). The general problem of interval parameter dependent linear systems was first treated in Rump (1994).

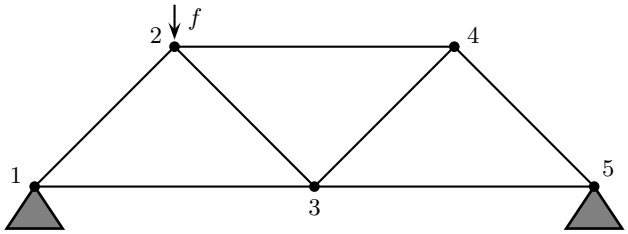


Figure 3.12: (Example 3.88) A truss structure.

Direct methods for computing enclosures to the parametric solution set were studied in Hladík (2012d); Kolev (2006b); Skalna (2006). Iterative methods include parametrized Gauss-Seidel iteration Popova (2001), a fixed-point method by Rump (1994, 2010), an effective method for a special class of parametric systems by Neumaier and Pownuk (2007), and the approach by Kolev (2004b). Monotonicity approach Popova (2004, 2006b); Rohn (2004); Skalna (2008) may substantially improve the accuracy of enclosures since it reduces the domains of some parameters to the endpoints. Relations between the methods and the effects of preconditioning and the residual form were analysed in Hladík and Skalna (2019). To obtain tighter enclosures, affine form concept (see Section 7.1.3) was employed in Skalna and Hladík (2017, 2019) A survey and a detailed description of the state-of-the-art methods are provided in the book by Skalna (2018).

A small selection of a broad area of applications includes those of structural mechanics (Garloff et al., 2009; Smith et al., 2010), mechanical systems (Dessombz et al., 2001), or tolerance analysis in linear circuits (Dreyer, 2007, 2006; Kolev, 1993). Fazzolari and Ferreira (2021) applied parametric interval system solving in testing robust stability of dynamical systems via interval Lyapunov equations; the transformation of Lyapunov equations to standard equations by means of Kronecker product induces the dependencies. Determinants of parametric interval matrices for singularity detection in robotics were applied in Merlet and Donelan (2006) by means of symbolic preconditioning. Regularity of parametric interval matrices was also used in Fazzolari and Ferreira (2021) to check existence of a controller in linear time-invariant control systems.

A Mathematica package for solving parametric interval systems was introduced in Popova (2004), a C-XCS implementation in Popova and Krämer (2007); Zimmer et al. (2012), and an interactive service for computations webComputing in Popova (2006a).

3.9.2 The symmetric solution set

We focus on the symmetric case now. Symmetric matrices naturally appear in many situations, so being able to handle the symmetric case is important. The corresponding solution set is

$$\begin{aligned}\Sigma_{sym} &:= \{x \in \mathbb{R}^n;\ Ax = b \text{ for some } A \in \boldsymbol{A}^S\} \\ &= \{x \in \mathbb{R}^n;\ Ax = b \text{ for some symmetric } A \in \boldsymbol{A}\}.\end{aligned}$$

This interval system obeys the linear parametric structure mentioned above since the symmetry can be simply modelled by using $\frac{1}{2}n(n-1)$ interval parameters. However, using the explicit description of the parametric matrix in the form $\sum_{k=1}^{K} A^k \boldsymbol{p}_k$ would be too much space (and hence also time) consuming. Therefore it is better to handle the symmetry implicitly.

This symmetric case has been studied since 1985; see Rohn (2004). Hladík (2008b) characterized Σ_{sym} by an explicit system of exponentially many nonlinear inequalities. The high number of them is justified by P-completeness of testing whether a given point $x \in \mathbb{R}^n$ belongs to Σ_{sym} – this means that an adaptation of Theorem 3.83 is valid also for the symmetric solution set; see Hladík (2020a). Improved and extended expositions of the symmetric case are provided in a survey paper by Mayer (2012) and the book by Mayer (2017).

Example 3.89. The symmetric interval system in two dimensional space draws $\boldsymbol{A}^S x = \boldsymbol{b}$,

$$\boldsymbol{A}^S = \boldsymbol{a}_{11}\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix} + \boldsymbol{a}_{12}\begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} + \boldsymbol{a}_{22}\begin{pmatrix}0 & 0\\ 0 & 1\end{pmatrix}.$$

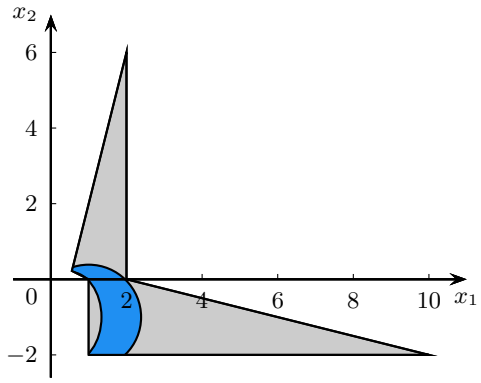


Figure 3.13: (Example 3.89) The solution set arbitrarily larger than the symmetric solution set, $a = 4$.

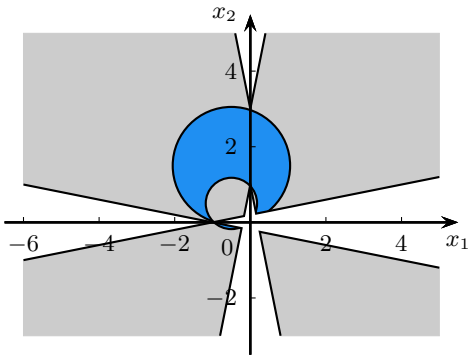


Figure 3.14: (Example 3.89) The solution set is unbounded, but the symmetric solution set is bounded.

Herein, interval entries $\boldsymbol{a}_{11}, \boldsymbol{a}_{12}, \boldsymbol{a}_{22}$ play the role of interval parameters $\boldsymbol{p}_1, \boldsymbol{p}_2, \boldsymbol{p}_3$. By Hladík (2008b) (cf. Mayer (2017)), the symmetric solution set is described by

$$\begin{aligned} |a_{11}^c x_1 + a_{12}^c x_2 - b_1^c| &\le a_{11}^\Delta |x_1| + a_{12}^\Delta |x_2| + b_1^\Delta \\ |a_{21}^c x_1 + a_{22}^c x_2 - b_2^c| &\le a_{21}^\Delta |x_1| + a_{22}^\Delta |x_2| + b_2^\Delta, \\ |-a_{11}^c x_1^2 + a_{22}^c x_2^2 + b_1^c x_1 - b_2^c x_2| &\ge a_{11}^\Delta x_1^2 + a_{22}^\Delta x_2^2 + b_1^\Delta |x_1| + b_2^\Delta |x_2|, \end{aligned}$$

where the first two inequalities are the Oettli–Prager inequalities (3.1). Figure 3.13 illustrates the solution set (light grey color) and symmetric solution set (grey color) for the system

$$\boldsymbol{A} = \begin{pmatrix} [1,2] & [0,a] \\ [0,a] & -1 \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} 2 \\ 2 \end{pmatrix},$$

in which the solution set can be arbitrarily larger than the symmetric one, depending on the real parameter $a > 0$. Figure 3.14 illustrates the case where the solution set is unbounded whereas the symmetric solution set is bounded. The corresponding data are

$$\boldsymbol{A} = \begin{pmatrix} -1 & [-5,5] \\ [-5,5] & 1 \end{pmatrix}, \quad \boldsymbol{b} = \begin{pmatrix} 1 \\ [1,3] \end{pmatrix}. \qquad \square$$

Regularity. Even for checking regularity of parametric interval matrices, we have to develop special techniques since simple relaxation leads to overestimation. For example, let

$$\boldsymbol{A} \coloneqq \begin{pmatrix} 1 & [-1,1] \\ [-1,1] & -1 \end{pmatrix}.$$

This interval matrix is irregular as it contains the singular matrix $\left(\begin{smallmatrix} 1 & 1 \\ -1 & -1 \end{smallmatrix}\right)$. On the other hand, there is no symmetric matrix inside that would be singular. That is, the symmetric interval matrix $\boldsymbol{A}^S$ is regular. Moreover, the conditions for testing regularity of standard interval matrices are not straightforwardly applicable. For example, both symmetric vertex matrices of the symmetric interval matrix

$$\boldsymbol{A}^S = \begin{pmatrix} 0 & [-1,1] \\ [-1,1] & 0 \end{pmatrix}^S$$

are nonsingular, and have the same determinant and spectrum, so the direct adaptation of Theorem 3.61(3) fails.

Enclosures of Σ_{sym}. Now, let us focus on computing an enclosure of the symmetric solution set. Considering the symmetric interval matrix $\boldsymbol{A}^S$ explicitly in the corresponding parametric form $A(p)$, $p \in \boldsymbol{p}$, is not convenient here because there would be a quadratic number of parameters and the matrices $A^{(k)}$ would be very sparse. Therefore, it is better to handle symmetry implicitly.

As in the previous section, we consider an enclosure in the form of $\boldsymbol{x} = x^* + \boldsymbol{y}$, where x^* is any real vector (e.g., the solution of $A^c x = b^c$) and $\boldsymbol{y} \in \mathbb{IR}^n$ encloses the solution set of

$$(C\boldsymbol{A}^S)y = C(\boldsymbol{b} - \boldsymbol{A}^S x^*).$$

Since in each column of $\boldsymbol{A}^S$ every parameter appears at most once, we have

$$C\boldsymbol{A} = \square\{CA;\ A \in \boldsymbol{A}^S\}.$$

Thus, the best way to evaluate the constraint matrix is simply to use interval arithmetic and calculate $C\boldsymbol{A}$. Utilizing dependencies in $\boldsymbol{A}^S$ cannot improve the enclosure of the constraint matrix.

What does, however, help is the utilization of dependencies in the evaluation the right-hand side vector of the residual form. The direct evaluation by interval arithmetic will ignore the dependencies, leading to needless overestimation. Taking into account the symmetry of $\boldsymbol{A}^S$, the ℓ-th entry of the right-hand side is evaluated as

$$\boldsymbol{z}_\ell \coloneqq \sum_{j=1}^{n} C_{\ell j}(\boldsymbol{b}_j - \boldsymbol{A}_{jj}x_j^*) - \sum_{i<j}(C_{\ell i}x_j^* + C_{\ell j}x_i^*)\boldsymbol{A}_{ij}.$$

Since each interval quantity appears at most once in the above expression, by Theorem 2.16, $\boldsymbol{z}_\ell$ is the optimal interval. That is,

$$\boldsymbol{z} = \square\big\{C(b - Ax^*);\ b \in \boldsymbol{b}, A \in \boldsymbol{A}, A = A^T\big\}.$$

Therefore, the enclosing interval vector $\boldsymbol{y}$ is computed by employing any solver for the standard interval linear system of equations

$$(C\boldsymbol{A})y = \boldsymbol{z}.$$

Interval Cholesky method. Interval Cholesky method for enclosing the symmetric solution set with a positive definite interval matrix $\boldsymbol{A}^S$ was introduced by Alefeld and Mayer (1993) and later dealt with in Alefeld and Mayer (2008); Garloff (2012); Mayer (2017), among others. The basic idea is straightforward – we replace the arithmetic operations by interval arithmetic, and the procedure yields a lower triangular interval matrix $\boldsymbol{L}$ such that for each $A \in \boldsymbol{A}^S$ there exists $L \in \boldsymbol{L}$ such that $A = LL^T$. An enclosure of Σ_{sym} is then obtained by forward and backward substitution.

Similarly as for the Gaussian elimination, the procedure may break down even for $\boldsymbol{A}^S$ positive definite. That is, the existence of Cholesky decomposition for each $A \in \boldsymbol{A}^S$ does not imply the existence of the interval Cholesky decomposition in general; for a counterexample see Mayer (2017); Reichmann (1979).

Mayer (2017) also shows an example, where the interval Cholesky decomposition yields an enclosure of Σ_{sym} that is tighter (in some coordinates) than $\square\Sigma$ (the solution set obtained by relaxing symmetry). On the other hand, it sometimes produces enclosures that are worse than those calculated by the interval Gaussian elimination.

Analogies of Theorems 3.24 and 3.27 are valid, too. In particular, if $\boldsymbol{A}^S$ is an H-matrix with positive diagonal entries, then the interval Cholesky decomposition succeeds and $\boldsymbol{L}$ is again an H-matrix. If $\boldsymbol{A}^S$ is an M-matrix, then we can say more. The interval Cholesky decomposition succeeds and produces an enclosure $\Sigma \subseteq \boldsymbol{x}^{Ch}$, $\boldsymbol{L}$ is again an M-matrix and it has the form of $\boldsymbol{L} = [L_\ell, L_u]$, where $\underline{A} = L_\ell L_\ell^T$ and $\overline{A} = L_u L_u^T$ are the corresponding Cholesky decompositions. Similarly as for the interval Gaussian elimination, we have $\boldsymbol{x}^{Ch} = \square\Sigma_{sym}$ if $\underline{b} \geq 0$, $\overline{b} \leq 0$ or $0 \in \boldsymbol{b}$ holds true.

Garloff (2012) investigated pivot tightening, which prevents the interval Cholesky method to break down and produces tighter Cholesky factors. For other references see Mayer (2017).

3.10 *AE* solution set

So far, we considered existentially quantified interval systems

$$\Sigma \coloneqq \{x \in \mathbb{R}^n;\ \exists A \in \boldsymbol{A}\, \exists b \in \boldsymbol{b} : Ax = b\}, \tag{3.37}$$

also referred to as *the united solution set.* In some applications, universal quantifiers may appear. Suppose that each interval parameter is quantified by some quantifier and universal quantifiers precede the existential ones; this is called *AE* quantification. The interval quantities that are universally quantified are denoted by $\boldsymbol{A}^\forall$, $\boldsymbol{b}^\forall$, and the existential ones by $\boldsymbol{A}^\exists$, $\boldsymbol{b}^\exists$. Thus, the interval system $\boldsymbol{A}x = \boldsymbol{b}$ can be written as $(\boldsymbol{A}^\forall + \boldsymbol{A}^\exists)x = \boldsymbol{b}^\forall + \boldsymbol{b}^\exists$, and the so-called *AE* solution set is defined

$$\Sigma_{AE} \coloneqq \big\{x \in \mathbb{R}^n;\ \forall A^\forall \in \boldsymbol{A}^\forall\, \forall b^\forall \in \boldsymbol{b}^\forall\, \exists A^\exists \in \boldsymbol{A}^\exists\, \exists b^\exists \in \boldsymbol{b}^\exists : (A^\forall + A^\exists)x = b^\forall + b^\exists\big\}.$$

Characterization of the *AE* solution set was developed by Shary (1995b).

Theorem 3.90 (Shary, 1995b)**.** *We have*

$$\Sigma_{AE} = \big\{x \in \mathbb{R}^n;\ \boldsymbol{A}^\forall x - \boldsymbol{b}^\forall \subseteq \boldsymbol{b}^\exists - \boldsymbol{A}^\exists x\big\}. \tag{3.38}$$

Proof.

$$\begin{aligned}\Sigma_{AE} &= \big\{x \in \mathbb{R}^n;\ \forall A^\forall \in \boldsymbol{A}^\forall\, \forall b^\forall \in \boldsymbol{b}^\forall\, \exists A^\exists \in \boldsymbol{A}^\exists\, \exists b^\exists \in \boldsymbol{b}^\exists : A^\forall x - b^\forall = b^\exists - A^\exists x\big\}\\ &= \big\{x \in \mathbb{R}^n;\ \forall A^\forall \in \boldsymbol{A}^\forall\, \forall b^\forall \in \boldsymbol{b}^\forall : A^\forall x - b^\forall \in \boldsymbol{b}^\exists - \boldsymbol{A}^\exists x\big\}\\ &= \big\{x \in \mathbb{R}^n;\ \boldsymbol{A}^\forall x - \boldsymbol{b}^\forall \subseteq \boldsymbol{b}^\exists - \boldsymbol{A}^\exists x\big\}.\end{aligned}$$

□

A reformulation by means of inequalities is due to Rohn; see Shary (2002).

Theorem 3.91 (Rohn, 1996)**.** *We have*

$$\Sigma_{AE} = \big\{x \in \mathbb{R}^n;\ |A^c x - b^c| \le \big(\operatorname{rad}\boldsymbol{A}^\exists - \operatorname{rad}\boldsymbol{A}^\forall\big)|x| + \operatorname{rad}\boldsymbol{b}^\exists - \operatorname{rad}\boldsymbol{b}^\forall\big\}.$$

Proof. By (2.1), two intervals $\boldsymbol{p}, \boldsymbol{q} \in \mathbb{IR}$ satisfy $\boldsymbol{p} \subseteq \boldsymbol{q}$ if and only if $|p^c - q^c| \le q^\Delta - p^\Delta$. In this way, the inclusion (3.38) is written as

$$\begin{aligned}\big|\operatorname{mid}\big(\boldsymbol{A}^\forall x - \boldsymbol{b}^\forall\big) - \operatorname{mid}\big(\boldsymbol{b}^\exists - \boldsymbol{A}^\exists x\big)\big| &\le \operatorname{rad}\big(\boldsymbol{A}^\exists x - \boldsymbol{b}^\exists\big) - \operatorname{rad}\big(\boldsymbol{b}^\forall - \boldsymbol{A}^\forall x\big)\\ &= \operatorname{rad}(\boldsymbol{A}^\exists)|x| + \operatorname{rad}\boldsymbol{b}^\exists - \operatorname{rad}(\boldsymbol{A}^\forall)|x| - \operatorname{rad}\boldsymbol{b}^\forall,\end{aligned}$$

or,

$$\big|\big(\operatorname{mid}\boldsymbol{A}^\forall + \operatorname{mid}\boldsymbol{A}^\exists\big)x - \big(\operatorname{mid}\boldsymbol{b}^\forall + \operatorname{mid}\boldsymbol{b}^\exists\big)\big| \le \big(\operatorname{rad}\boldsymbol{A}^\exists - \operatorname{rad}\boldsymbol{A}^\forall\big)|x| + \operatorname{rad}\boldsymbol{b}^\exists - \operatorname{rad}\boldsymbol{b}^\forall.$$

□

Based on this characterization, we see that the *AE* solution set Σ_{AE} has a similar geometric structure as the standard solution set Σ. In particular, Σ_{AE} is a non-convex polyhedral set, which is convex in each orthant.

Example 3.92. Consider the interval system with quantificators

$$\begin{pmatrix}[3,4]^\exists & [-2,1]^\exists\\ [0,2]^\forall & [3,4]^\forall\end{pmatrix} x = \begin{pmatrix}[-4,5]^\exists\\ [-4,5]^\exists\end{pmatrix}. \tag{3.39}$$

The corresponding *AE* solution set and the united solution set (3.37) are depicted in Figure 3.15.

For comparison, consider the same interval system with another quantification

$$\begin{pmatrix}[3,4]^\forall & [-2,1]^\forall\\ [0,2]^\forall & [3,4]^\forall\end{pmatrix} x = \begin{pmatrix}[-4,5]^\exists\\ [-4,5]^\exists\end{pmatrix}. \tag{3.40}$$

Now, the *AE* solution set is drawn in Figure 3.16. Notice that in this case we have the so-called tolerable solutions, which will be discussed in Section 3.10.2. □

The concept of *AE* solutions is quite general. In the following subsections, we inspect two particular cases of controllable and tolerable solutions. Another case of strong solutions, defined by ∀-quantifiers only, will be addressed later in Section 4.5.1.

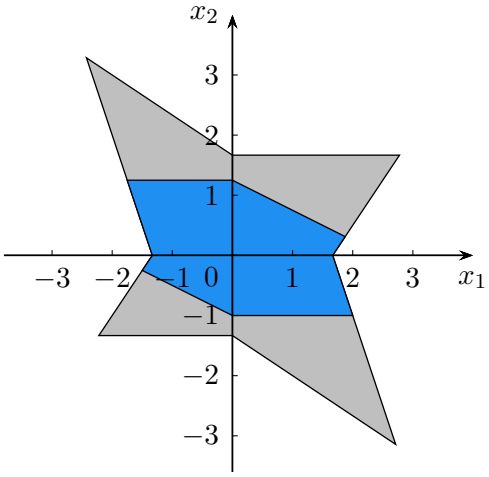


Figure 3.15: (Example 3.92) The *AE* solution set Σ_{AE} as a subset of the classical solution set Σ for (3.39).

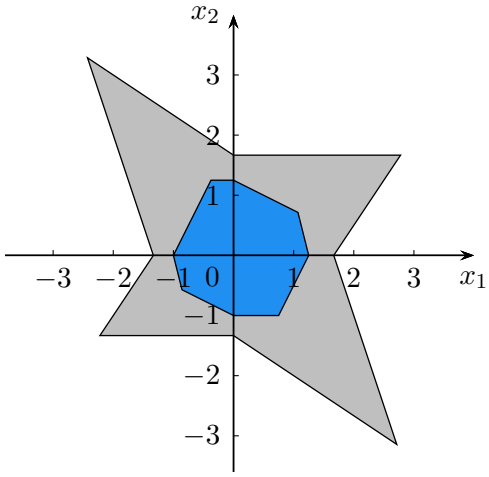


Figure 3.16: (Example 3.92) The *AE* solution set Σ_{AE} as a subset of the classical solution set Σ for (3.40).

3.10.1 Controllable solutions

Controllable solutions are a special case of *AE* solution. A vector $x \in \mathbb{R}^n$ is called a *controllable* solution to the interval system $\boldsymbol{A}x = \boldsymbol{b}$ if for each $b \in \boldsymbol{b}$ there is $A \in \boldsymbol{A}$ such that $Ax = b$. They were first introduced by Shary (1992) and later studied in Fiedler et al. (2006); Shary (2002).

Theorem 3.90 gives the characterization of a controllable solution by means of the inclusion $\boldsymbol{b} \subseteq \boldsymbol{A}x$. By Theorem 3.91 they are equivalently characterized by the system (Lakeev and Noskov, 1994)

$$|A^c x - b^c| \leq A^\Delta |x| - b^\Delta.$$

It is known that checking existence of a controllable solution is *NP*-hard (Lakeev and Noskov, 1994); see also (Fiedler et al., 2006, Thm. 2.30).

3.10.2 Tolerable solutions

Another special *AE* solution is a tolerable solution. A vector $x \in \mathbb{R}^n$ is a *tolerable* solution to $\boldsymbol{A}x = \boldsymbol{b}$ if for each $A \in \boldsymbol{A}$ there is $b \in \boldsymbol{b}$ such that $Ax = b$. By Theorem 3.90, they are characterized by the inclusion $\boldsymbol{A}x \subseteq \boldsymbol{b}$, and Theorem 3.91 provides a description by means of the system

$$|A^c x - b^c| \leq -A^\Delta |x| + b^\Delta. \tag{3.41}$$

At first sight, it may seem that checking existence of a tolerable solution is *NP*-hard, too, but the converse is true. Tolerable solutions can be described by linear inequalities and thus the problem becomes tractable (Rohn, 1986); cf. (Fiedler et al., 2006, Thm. 2.28). Geometrically, the tolerable solution set forms a convex polyhedron; see Figure 3.16.

Theorem 3.93 (Rohn, 1986)**.** *A vector $x \in \mathbb{R}^n$ is a tolerable solution if and only if $x = x_1 - x_2$, where*

$$\overline{A}x_1 - \underline{A}x_2 \leq \overline{b},\ \underline{A}x_1 - \overline{A}x_2 \geq \underline{b},\ x_1, x_2 \geq 0.$$

Proof. "⇐" Let $A \in \boldsymbol{A}$. Then

$$\begin{aligned} Ax = Ax_1 - Ax_2 &\leq \overline{A}x_1 - \underline{A}x_2 \leq \overline{b}, \\ Ax = Ax_1 - Ax_2 &\geq \underline{A}x_1 - \overline{A}x_2 \geq \underline{b} \end{aligned}$$

Thus, $Ax \in \boldsymbol{b}$ and $Ax = b$ for some $b \in \boldsymbol{b}$.

"⇒" Let $x \in \mathbb{R}^n$ be a tolerable solution. Define $x_1 \coloneqq \max\{x, 0\}$ and $x_2 \coloneqq \max\{-x, 0\}$ the positive and negative part of x, respectively. Then $x = x_1 - x_2$, $|x| = x_1 + x_2$, and the inequality (3.41) draws

$$\begin{aligned} A^c(x_1 - x_2) - b^c &\leq -A^\Delta(x_1 + x_2) + b^\Delta, \\ -A^c(x_1 - x_2) + b^c &\leq -A^\Delta(x_1 + x_2) + b^\Delta, \end{aligned}$$

which is easily transformed to the desired form. □

The early motivations for studying tolerable solutions came from the crane construction and from the input-output planning (Rohn, 1978). Further research was conveyed in Fiedler et al. (2006); Lakeev and Noskov (1994); Shary (1995a, 2002). The following application of tolerable solutions addresses the Leontief input–output model with uncertain data; see also Dymova et al. (2013); Li and Liu (2008).

Example 3.94 (Input–Output Model)**.** The input–output model is an economic technique that highlights the interdependencies between different branches of national or different economies. The model was developed by Wassily Leontief in 1930s. Later, he was awarded the Nobel Prize in 1973 "for the development of the input–output method and for its application to important economic problems".

Let us we have an economy with n sectors (e.g., agriculture, industry, transportation, etc.). Sector i produces a single commodity of amount x_i. Assume that the production of each unit of the jth commodity will require a_{ij} (amount) of the ith commodity. Denote by d_i the final demand in sector i. Now the model draws

$$x_i = a_{i1}x_1 + \cdots + a_{in}x_n + d_i,$$

or, in a matrix form

$$x = Ax + d.$$

Suppose that $\rho(A) < 1$. Then $I_n - A$ is an M-matrix and by Theorem 1.10the solution $x = (I_n - A)^{-1}d = \sum_{k=0}^{\infty} A^k d$ is nonnegative. This is essential in practice

The entries of A and d are often estimated and thus imprecise. Say we have lower and upper bounds for the true values. Hence A varies within some interval matrix $\boldsymbol{A}$ and d within $\boldsymbol{d}$. A natural question arises whether a solution x exists such that for each $A \in \boldsymbol{A}$ there is $d \in \boldsymbol{d}$ satisfying $(I_n - A)x = d$. In other words, whether the interval system $(I_n - \boldsymbol{A})x = \boldsymbol{d}$ has a tolerable solution.

A personal supplement to centrally planned economy in communist systems was given by Rohn (2019a). □

Notes and further reading. For some other motivations to tolerable and controllable solutions see Kreinovich (2008). Rzeżuchowski and Wąsowski (2017) characterized the AE solution set by the theory of convex sets, and they also extended the concept by considering more general domains for the parameters than intervals.

Chapter 4

Systems of interval linear inequalities

Recall that a real system od linear inequalities $Ax \leq b$ describes a convex polyhedron. Theory of convex polyhedra closely relates to linear programming. This chapter extends the theory to the case the coefficients are interval-valued.

We present various solution concepts, including weak solutions, strong solutions, AE solutions, and weak and strong solvability. The concepts concern not only interval systems of linear inequalities, but also the linear equations and mixed linear equations and inequalities.

4.1 Weak solutions of interval linear inequalities

Let $\boldsymbol{A} \in \mathbb{IR}^{m\times n}$, $\boldsymbol{b} \in \mathbb{IR}^m$, and consider an interval linear system of inequalities

$$\boldsymbol{A}x \leq \boldsymbol{b}, \tag{4.1}$$

which is a shortcut for the family

$$Ax \leq b, \quad A \in \boldsymbol{A},\ b \in \boldsymbol{b}.$$

Similarly as for interval equations, a vector $x \in \mathbb{R}^n$ is called *a solution* (or *a weak solution*) if $Ax \leq b$ for some $A \in \boldsymbol{A}$ and $b \in \boldsymbol{b}$. A basic characterization of solutions is due to Gerlach (1981).

Theorem 4.1 (Gerlach, 1981)**.** *A vector $x \in \mathbb{R}^n$ is a solution of $\boldsymbol{A}x \leq \boldsymbol{b}$ if and only if*

$$A^c x \leq A^\Delta |x| + \overline{b}. \tag{4.2}$$

Proof. If x is a solution of (4.1), then $Ax \leq b$ for some $A \in \boldsymbol{A}$ and $b \in \boldsymbol{b}$, and one has

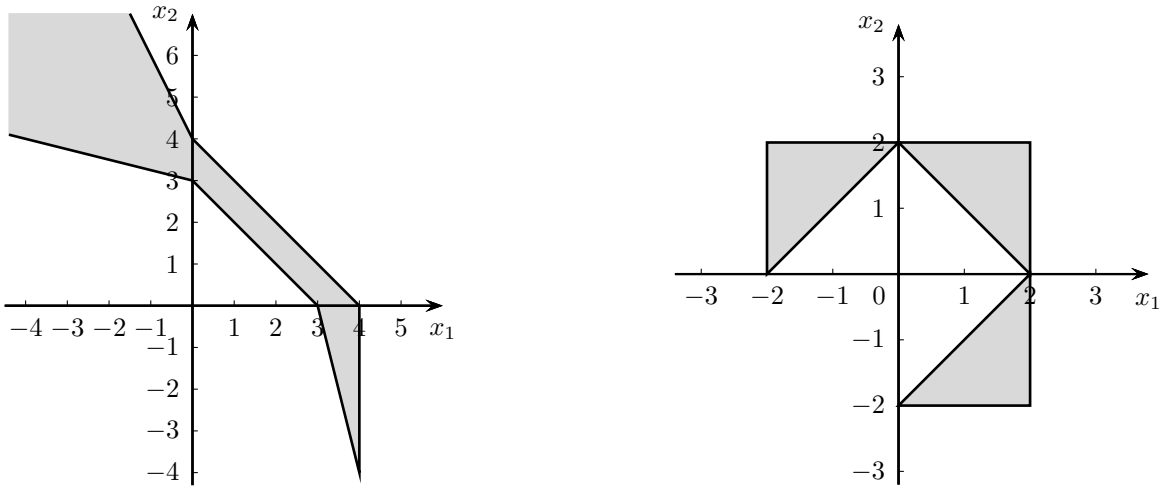
$$A^c x \leq A^c x + b - Ax = (A^c - A)x + b \leq |(A^c - A)||x| + b \leq A^\Delta |x| + \overline{b}.$$

Conversely, let x satisfy (4.2). Set $z \coloneqq \operatorname{sgn}(x)$, so $|x| = \operatorname{diag}(z)x$. Thus, (4.2) takes the form of $A^c x \leq A^\Delta \operatorname{diag}(z)x + \overline{b}$, or $(A^c - A^\Delta \operatorname{diag}(z))x \leq \overline{b}$. Hence x fulfills $Ax \leq b$ for $b \coloneqq \overline{b}$ and $A \coloneqq A_{ez} = A^c - A^\Delta \operatorname{diag}(z)$. □

By the above proof, we also have that $\boldsymbol{A}x \leq \boldsymbol{b}$ is solvable if and only if $A_{ez}x \leq \overline{b}$ is solvable for at least one $z \in \{\pm 1\}^n$. The number 2^n of instances can be reduced to 2^k provided the intervals in $\boldsymbol{A}$ are situated in only k columns; cf. Remark 3.13. In particular, if the matrix is real, then the solution set of $Ax \leq \boldsymbol{b}$ is described by $Ax \leq \overline{b}$.

In the same way as for interval equations (Theorem 3.5), the characterization (4.2) reveals that the set of all solutions forms a polyhedral set, which is convex in each orthant. In particular, for the nonnegative orthant we have the following consequence.

Corollary 4.2 (Vajda, 1961)**.** *A vector $x \in \mathbb{R}^n$ is a solution of $\boldsymbol{A}x \leq \boldsymbol{b}$, $x \geq 0$ if and only if $\underline{A}x \leq \overline{b}$, $x \geq 0$.*



(a) The first system. (b) The second system.

Figure 4.1: (Example 4.3) Two examples of interval systems of linear inequalities and their solution sets.

Proof. By the Gerlach theorem, the solution set is described $A^c x \leq A^\Delta |x| + \overline{b}$, $x \geq 0$, which is equivalent to $A^c x \leq A^\Delta x + \overline{b}$, $x \geq 0$, or $\underline{A}x \leq \overline{x}$, $x \geq 0$. □

Example 4.3. Figure 4.1 gives an illustration of two interval systems of linear inequalities, inspired by Chen et al. (2009). The first one is

$$\begin{aligned} [1,2]x_1 + [1,2]x_2 &\leq 4, \\ -[1,4]x_1 - [1,4]x_2 &\leq -12, \\ x_1 &\leq 4, \\ -x_2 &\leq 4, \end{aligned}$$

and it is depicted in Figure 4.1a. The solution set of the second system

$$\begin{aligned} [-1,1]x_1 + [-1,1]x_2 &\leq -2, \\ -x_1 - x_2 &\leq 2, \\ -2 \leq x_1 &\leq 2, \\ -2 \leq x_2 &\leq 2 \end{aligned}$$

is depicted in Figure 4.1b. □

Computational complexity. In the general case, however, the solution set need not be convex and checking non-emptiness is an NP-hard problem (Fiedler et al., 2006; Rohn, 1995).

Theorem 4.4 (Rohn, 1995)**.** *Checking solvability of $\boldsymbol{A}x \leq \boldsymbol{b}$ is an NP-hard problem.*

Proof. By Theorem 1.35 we know that checking solvability of $-e \leq Ax \leq e$, $e^T|x| \geq 1$ is *NP*-hard. The system is equivalent to

$$Ax \leq e,\ -Ax \leq e,\ 0x \leq e^T|x| - 1,$$

which is the Gerlach characterization (4.2) of the interval system $\boldsymbol{C}x \leq \boldsymbol{c}$ with

$$C^c = \begin{pmatrix} A \\ -A \\ 0 \end{pmatrix},\ C^\Delta = \begin{pmatrix} 0 \\ 0 \\ e^T \end{pmatrix},\ \boldsymbol{c} = \begin{pmatrix} e \\ e \\ -1 \end{pmatrix}.$$

□

In fact, the proof revealed that checking solvability of interval linear inequalities is *NP*-hard even if the intervals are situated in one inequality only. We show now a modified statement, which will be useful later on. Another computationally complexity result for related interval systems was presented in Hladík (2021).

Theorem 4.5 (Garajová et al., 2017)**.** *Checking solvability of $Ax \le 0$, $\boldsymbol{b}^T x < 0$ is an NP-hard problem.*

Proof. By Theorem 1.36 we know that checking solvability of

$$|Ax| \le e, \quad e^T|x| > 1 \tag{4.3}$$

is *NP*-hard. We now show that it is equivalent to checking solvability of the system

$$|Ax| \le ey, \quad y \ge 0, \quad e^T|x| > y. \tag{4.4}$$

Clearly, if x solves (4.3), then the pair x and $y = 1$ solves (4.4). Conversely, let x, y be a solution to (4.4). If $y > 0$, then $x' \coloneqq \frac{1}{y}x$ solves (4.3). When $y = 0$, then x satisfies $Ax = 0$, $e^T|x| > 0$. Consider the vector $x' \coloneqq \frac{1}{e^T|x|-\varepsilon}x$, where $0 < \varepsilon < e^T|x|$. Then x' solves (4.3) since $|Ax'| = 0 \le e$ and

$$e^T|x'| = \frac{1}{e^T|x| - \varepsilon} e^T|x| > 1.$$

Now, by the Gerlach characterization and its modification, system (4.4) describes the solution set of the interval system

$$Ax - ey \le 0, \quad -Ax - ey \le 0, \quad -y \le 0, \quad [-e, e]^T x + y < 0,$$

which has the desired form. □

In general, the standard transformations of interval systems are not possible for the interval case since they cause dependencies; see Section 4.3. There are a few of exceptions, one of them stated below. Notice that the interval system $\boldsymbol{A}x \le \boldsymbol{b}$, $\boldsymbol{A}x \ge \boldsymbol{b}$ is understood as a family of systems

$$A^1 x \le b^1, \quad A^2 x \ge b^2,$$

with $A^1, A^1 \in \boldsymbol{A}$ and $b^1, b^1 \in \boldsymbol{b}$ being chosen independently, which makes the statement nontrivial.

Theorem 4.6 (Rohn, 1985; Li, 2015)**.** *The solution set of $\boldsymbol{A}x = \boldsymbol{b}$ is the same as the solution set of $\boldsymbol{A}x \le \boldsymbol{b}$, $\boldsymbol{A}x \ge \boldsymbol{b}$.*

Proof. By the Oettli–Prager Theorem 3.2, the solution set of $\boldsymbol{A}x = \boldsymbol{b}$ is characterized by

$$\begin{aligned} A^c x - b^c &\le A^\Delta |x| + b^\Delta, \\ -A^c x + b^c &\le A^\Delta |x| + b^\Delta. \end{aligned}$$

By the Gerlach Theorem 4.1, the solution set of $\boldsymbol{A}x \le \boldsymbol{b}$, $-\boldsymbol{A}x \le -\boldsymbol{b}$ is characterized by

$$\begin{aligned} A^c x &\le A^\Delta |x| + \overline{b}, \\ -A^c x &\le A^\Delta |x| - \underline{b}, \end{aligned}$$

which is equivalent to the above one. □

Not only checking non-emptiness, but also boundedness of the solution set is an intractable problem.

Corollary 4.7 (Garajová and Hladík, 2019)**.** *It is co-NP-hard to decide whether the solution set of $\boldsymbol{A}x \le \boldsymbol{b}$ is bounded.*

Proof. From Corollary 3.68 we know that checking boundedness of the solution set of $\boldsymbol{A}x = \boldsymbol{b}$ is *co-NP*-hard. By Theorem 4.6, the system can be equivalently expressed as $\boldsymbol{A}x \le \boldsymbol{b}$, $-\boldsymbol{A}x \le -\boldsymbol{b}$, from which the result follows. □

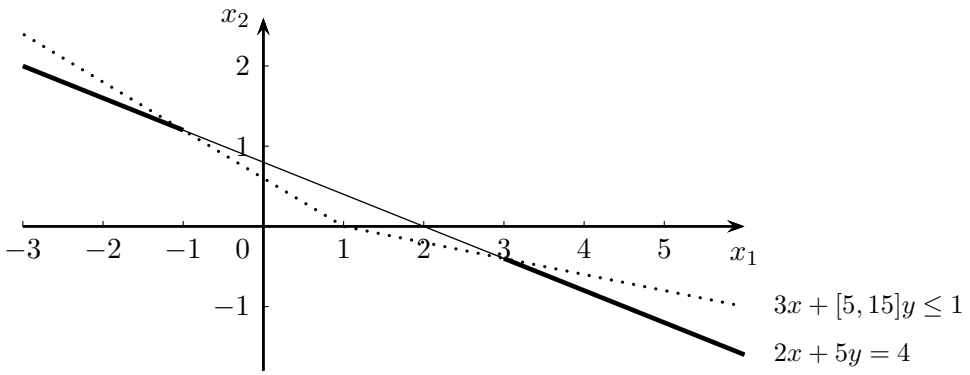


Figure 4.2: The solution set of interval system $2x + 5y = 4,\ 3x + [5, 15]y \leq 1$ consists of two disjoint rays.

Topological properties. We will now study more topological properties of the solution set

$$\Sigma \coloneqq \{x \in \mathbb{R}^n;\ Ax \leq b,\ A \in \boldsymbol{A},\ b \in \boldsymbol{b}\}.$$

We have already mentioned that Σ is a polyhedral set, which is convex in each orthant. However, it might be disconnected. For instance, the solution set of

$$[-1, 1]x \leq -1$$

consists of two disjoint sets $(-\infty, -1]$ and $[1, \infty)$. It might seem that disconnectedness is caused by the interval containing the zero, but it is not hard to find another example of a disconnected solution set without such an interval:

$$2x + 5y = 4,\quad 3x + [5, 15]y \leq 1.$$

The solution set consists of two disjoint rays and is depicted in Figure 4.2.

Below, we propose some sufficient conditions for connectedness.

Proposition 4.8. *If $\overline{b} \geq 0$, then Σ is connected.*

Proof. The condition $\overline{b} \geq 0$ implies $0 \in \Sigma$. Since Σ is connected in each orthant, it is connected as a whole via the origin. □

This condition is very cheap, but not very strong in general. The following condition is obviously stronger; consider, e.g., the interval system $-x \leq -1$ with degenerate intervals.

Proposition 4.9. *If the linear system of inequalities*

$$\overline{A}u - \underline{A}v \leq \overline{b},\ u, v \geq 0 \tag{4.5}$$

is solvable, then Σ is connected.

Proof. As we will see in Theorem 4.28, if u, v solves (4.5), then $x^* \coloneqq u - v$ is a solution to $Ax \leq \overline{b}$ for every $A \in \boldsymbol{A}$. Thus, every two points in Σ are connected via x^*. □

Proposition 4.9 gives only a sufficient condition for connectedness, but not necessary in general. For example, consider the interval linear system

$$-x \leq -1,\ [1, 2]x \leq 1.$$

Here, $\Sigma = \{1\}$ is connected, but the sufficient condition is not satisfied.

Example 4.10. This example shows that one additional inequality may split a connected solution set into an exponential number of components. Consider the interval linear inequalities

$$\begin{aligned} -K \leq x_i \leq K, &\quad i = 1, \ldots, n, \\ x_i + \sum_{j \neq i} [-\tfrac{1}{n-1}, \tfrac{1}{n-1}]x_j \leq 0, &\quad i = 1, \ldots, n, \end{aligned}$$

where $K > 0$ is large enough. Due to the symmetry, it is sufficient to investigate the non-negative orthant only. In this orthant, the restricted solution set is

$$\begin{aligned} 0 \le x_i \le K, \quad i = 1, \dots, n, \\ x_i - \sum_{j \ne i} \tfrac{1}{n-1} x_j \le 0, \quad i = 1, \dots, n, \end{aligned}$$

which describes the segment joining the origin and the point $(K, \dots, K)^T$. Thus, the solution set Σ is connected. Now, consider an additional constraint

$$\sum_{i=1}^{n} [-1, 1] x_i \ge 1.$$

In each orthant, it cuts off the closer-to-the-origin part of the segment. For instance, in the non-negative orthant, the restricted solution set will be the segment joining the points $(\frac{1}{n}, \dots, \frac{1}{n})^T$ and $(K, \dots, K)^T$. Hence the resulting solution set will consist of 2^n components. Notice again that this exponential increase of the components is not caused by the zero-containing intervals in the additional inequality. If the inequality reads

$$\sum_{i=1}^{n} [\varepsilon, 1] x_i \ge 1,$$

where $\varepsilon > 0$ is sufficiently small, then the solution set splits into $2^n - 1$ connected components. □

4.2 Systems of parametric linear inequalities

When the interval coefficients in the system do not take the values of their interval domains independently of each other, then we speak about dependencies and the system is called *a parametric interval system*; see Section 3.9 on parametric linear equations.

In this section, we consider merely a parametric system of inequalities with linear dependencies

$$A(p)x \le b(p),$$

where $A(p) = \sum_{k=1}^{K} A^k p_k$, $b(p) = \sum_{k=1}^{K} b^k p_k$ and $p \in \boldsymbol{p}$ for some given interval vector $\boldsymbol{p} \in \mathbb{IR}^K$, matrices $A^1, \dots, A^K \in \mathbb{R}^{m \times n}$ and vectors $b^1, \dots, b^K \in \mathbb{R}^m$. A vector $x \in \mathbb{R}^n$ is *a solution* of the parametric system if it satisfies $A(p)x \le b(p)$ for at least one $p \in \boldsymbol{p}$.

In general, it is hard to characterize the solution set. We state at least a necessary condition.

Theorem 4.11 (Popova, 2015)**.** *If $x \in \mathbb{R}^n$ is a solution, then*

$$A(p^c)x - b(p^c) \le \sum_{k=1}^{K} p_k^{\Delta} |A^k x - b^k|. \tag{4.6}$$

Proof. Let $x \in \mathbb{R}^n$ be a solution to $A(p)x \le b(p)$ for some $p \in \boldsymbol{p}$. Then,

$$\begin{aligned} A(p^c)x - b(p^c) &= \sum_{k=1}^{K} p_k^c (A^k x - b^k) \le \sum_{k=1}^{K} p_k^c (A^k x - b^k) - \sum_{k=1}^{K} p_k (A^k x - b^k) \\ &= \sum_{k=1}^{K} (p_k^c - p_k)(A^k x - b^k) \le \sum_{k=1}^{K} |p_k^c - p_k| |A^k x - b^k| \le \sum_{k=1}^{K} p_k^{\Delta} |A^k x - b^k|. \end{aligned}$$

□

Similarly as for linear interval parametric equations (Theorem 3.81) we can show that the above system fully characterizes the solution set provided each interval parameter influences at most one inequality.

Theorem 4.12 (Popova, 2015)**.** *Suppose that for every $k = 1, \dots, K$, the matrix $(A^k \mid b^k)$ has only one nonzero row. Then $x \in \mathbb{R}^n$ is a solution if and only if x solves (4.6).*

Proof. From the assumptions, the vector $A^k x - b^k$, $k \in \{1, \dots, K\}$, has at least one nonzero element. If $A^k x = b^k$, then put $s_k := 1$; otherwise, denote by s_k the sign of the nonzero element in $A^k x - b^k$. Now, we have

$$A(p^c)x - b(p^c) \le \sum_{k=1}^{K} p_k^\Delta s_k (A^k x - b^k),$$

from which

$$\sum_{k=1}^{K} (p_k^c - p_k^\Delta s_k)(A^k x - b^k) \le 0.$$

Therefore x is a solution of $A(p)x \le b(p)$ for $p_k := p_k^c - p_k^\Delta s_k \in \boldsymbol{p}_k$, $k \in \{1, \dots, K\}$. □

Notes and further reading. A characterization of the interval parametric system of equations and inequalities was derived by Popova (2015); she also considered various kinds of solutions. Hladík (2017) presented allowable transformations of (standard and) parametric systems of equations and inequalities for several types of solutions.

4.3 General interval linear systems

Now, we extend the previous ideas to an interval system of mixed equations and inequalities. Most of the results described below come from Hladík (2013b).

Let $\boldsymbol{A} \in \mathbb{IR}^{m_1 \times n_1}$, $\boldsymbol{B} \in \mathbb{IR}^{m_1 \times n_2}$, $\boldsymbol{C} \in \mathbb{IR}^{m_2 \times n_1}$, $\boldsymbol{D} \in \mathbb{IR}^{m_2 \times n_2}$, $\boldsymbol{b} \in \mathbb{IR}^{m_1}$, and $\boldsymbol{d} \in \mathbb{IR}^{m_2}$. Consider the general interval linear system

$$\boldsymbol{A}x + \boldsymbol{B}y = \boldsymbol{b}, \quad \boldsymbol{C}x + \boldsymbol{D}y \le \boldsymbol{d}, \ x \ge 0, \tag{4.7}$$

which is again a shortage for the family of linear systems

$$Ax + By = b, \quad Cx + Dy \le d, \ x \ge 0 \tag{4.8}$$

with $A \in \boldsymbol{A}$, $B \in \boldsymbol{B}$, $C \in \boldsymbol{C}$, $D \in \boldsymbol{D}$, $b \in \boldsymbol{b}$, and $d \in \boldsymbol{d}$.

Each interval linear system can be transformed into the form of (4.7). Inequalities of the form "$\ge$" are multiplied by -1, and the non-positivity condition $z \le 0$ is replaced by $z' \ge 0$ using the substitution $z' := -z$. None of these transformations causes dependencies in the interval system, and so does affect neither solvability nor the solution set of any type.

The solution set. We say that a pair $(x, y) \in \mathbb{R}^{n_1+n_2}$ is *a solution* of (4.7) if it satisfies (4.8) for some $A \in \boldsymbol{A}$, $B \in \boldsymbol{B}$, $C \in \boldsymbol{C}$, $D \in \boldsymbol{D}$, $b \in \boldsymbol{b}$, and $d \in \boldsymbol{d}$. The Oettli–Prager Theorem 3.2 characterizing interval linear equations and Gerlach's Theorem 4.1 for inequalities are generalized to the mixed interval linear system in the following theorem.

Theorem 4.13. *The solution set to (4.7) is described by*

$$\underline{A}x + B^c y \le B^\Delta |y| + \overline{b}, \tag{4.9a}$$

$$-\overline{A}x - B^c y \le B^\Delta |y| - \underline{b}, \tag{4.9b}$$

$$\underline{C}x + D^c y \le D^\Delta |y| + \overline{d}, \ x \ge 0. \tag{4.9c}$$

Proof. By the Oettli–Prager theorem, to be a weak solution to $\boldsymbol{A}x + \boldsymbol{B}y = \boldsymbol{b}$, $x \ge 0$, any $x \in \mathbb{R}^{n_1}$ and $y \in \mathbb{R}^{n_2}$, have to fulfill

$$|A^c x + B^c y - b^c| \le A^\Delta |x| + B^\Delta |y| + b^\Delta = A^\Delta x + B^\Delta |y| + b^\Delta,$$

that is,

$$A^c x + B^c y - b^c \le A^\Delta x + B^\Delta |y| + b^\Delta, \ -A^c x - B^c y + b^c \le A^\Delta x + B^\Delta |y| + b^\Delta.$$

By Gerlach's theorem, weak solutions to $\boldsymbol{C}x + \boldsymbol{D}y \le \boldsymbol{d}$, $x \ge 0$, are described

$$C^c x + D^c y \le C^\Delta x + D^\Delta |y| + \overline{d}, \ x \ge 0.$$

Thus, any solution to (4.7) must satisfy both conditions, which results in the characterization by (4.9). □

Since $|y| = \operatorname{diag}(\operatorname{sgn}(y))y$ for any $y \in \mathbb{R}^{n_2}$, we can linearize the absolute value in the description of weak solutions and obtain a characterization by means of 2^{n_2} linear systems.

Corollary 4.14. *A pair $(x, y) \in \mathbb{R}^{n_1+n_2}$ is a solution to (4.7) if and only if there is $s \in \{\pm 1\}^{n_2}$ such that*

$$\begin{aligned} \underline{A}x + B_{s,e}y &\leq \overline{b}, \\ -\overline{A}x - B_{-s,e}y &\leq -\underline{b}, \\ \underline{C}x + D_{s,e}y &\leq \overline{d}, \ x \geq 0. \end{aligned}$$

Obviously, if the intervals are situated in the right-hand side only, then the system (4.9) becomes linear, and the solution set is a convex polyhedron. Checking solvability is then easy. Similar reasoning holds provided the intervals are situated in a fixed number k of columns of the constraint matrix; see Remark 3.13. By the above corollary, the problem reduces to a fixed number 2^k of linear systems.

Corollary 4.15. *Checking solvability is a polynomial problem provided the constraint matrix is real or the number of variables with interval coefficients is fixed.*

Given a weak solution $(x, y) \in \mathbb{R}^{n_1+n_2}$, one may ask for a realization (4.8) of the interval system having (x, y) as a solution.

Theorem 4.16. *Let $(x, y) \in \mathbb{R}^{n_1+n_2}$ be a weak solution to (4.7). Then (x, y) solves (4.8) with*

$$\begin{aligned} A &= A_{u,e}, & B &= B_{u,s}, \\ C &= \underline{C}, & D &= D_{e,s}, \\ b &= b_u, & d &= \overline{d}, \end{aligned}$$

where $s = \operatorname{sgn}(y)$ and $u \in [-1, 1]^{m_1}$ is defined

$$u_i = \begin{cases} \frac{(A^c x + B^c y - b^c)_i}{(A^\Delta x + B^\Delta |y| + b^\Delta)_i} & \text{if } (A^\Delta x + B^\Delta |y| + b^\Delta)_i > 0, \\ 1 & \text{otherwise}, \end{cases} \qquad i = 1, \ldots, m_1.$$

Proof. The realizations concerning equations follow from Theorem 3.2. The realizations concerning inequalities easily follow from $|y| = \operatorname{diag}(\operatorname{sgn}(y))y$ and the inequality

$$\underline{C}x + D^c y \leq D^\Delta |y| + \overline{d}. \qquad \square$$

Transformations of interval systems. In the real case, one often applies various transformations yielding equivalent linear systems. In the interval case, one has to be careful since transformations can cause dependencies, and if one relaxes them, the transformation is no more equivalent. Below, we discuss two commonly used transformations. For extensions see Hladík (2017).

Splitting equations to double inequalities. This transformation replaces an interval system of linear equations by $\boldsymbol{A}x = \boldsymbol{b}$ with interval linear inequalities $\boldsymbol{A}x \leq \boldsymbol{b}$, $\boldsymbol{A}x \geq \boldsymbol{b}$. In Theorem 4.6, we observed that, surprisingly, this transformation does not affect the weak solution set. It can, however, change other types of solutions and solvabilities, for instance the strong solvability discussed later in Section 4.4.

Imposing nonnegativity. This transformation applies the substitution $x = x^1 - x^2$, $x^1, x^2 \geq 0$, which induces nonnegative variables x^1 and x^2 in the case x was a free variable. We can impose nonnegativity for both interval equations

$$\boldsymbol{A}x = \boldsymbol{b} \quad \to \quad \boldsymbol{A}x^1 - \boldsymbol{A}x^2 = \boldsymbol{b}, \ x^1, x^2 \geq 0,$$

and interval inequalities

$$\boldsymbol{A}x \leq \boldsymbol{b} \quad \to \quad \boldsymbol{A}x^1 - \boldsymbol{A}x^2 \leq \boldsymbol{b}, \ x^1, x^2 \geq 0.$$

Again, due to multiple occurrences of the interval coefficients, the transformation does not yield an equivalent system (an exception is strong solvability, as we will see in Theorem 4.20).

Example 4.17. Consider the interval system

$$x = 1, \quad [1,2]x = 3,$$

which has no weak solution. Imposing nonnegativity transforms the system to

$$x^1 - x^2 = 1, \quad [1,2]x^1 - [1,2]x^2 = 3, \quad x^1, x^2 \geq 0.$$

The solution set is nonempty now since the realization

$$x^1 - x^2 = 1, \quad 2x^1 - x^2 = 3, \quad x^1, x^2 \geq 0$$

has solution $x = (2,1)^T$. □

Imposing equations. The transformation of $\boldsymbol{A}x \leq \boldsymbol{b}$ to $\boldsymbol{A}x + z = \boldsymbol{b}$, $z \geq 0$ changes inequalities to equations with slack nonnegative variables. Since no dependencies occur, the transformation is equivalent. The transformed system, however, need not be more convenient. For example, we will see later that the conditions for checking strong solvability are more expensive for inequalities than for equations.

4.4 Strong solvability of interval systems

Herein, we introduce another solvability concept for interval linear systems of equations and inequalities. An interval linear system is *strongly solvable* if each realization is solvable. Notice that this concept differs from *AE* solutions discussed in Section 3.10 since strong solvability does not require all realizations to have a common solution.

Below, we state characterizations of strong solvability for various kinds of systems. More detailed exposition is given in Hladík (2013b); Rohn (2006a).

4.4.1 Interval equations

Let $\boldsymbol{A} \in \mathbb{IR}^{m\times n}$, $\boldsymbol{b} \in \mathbb{IR}^m$ and consider an interval system of linear equations $\boldsymbol{A}x = \boldsymbol{b}$. First, we characterize strong solvability for the case with nonnegative variables, and then we extend the result to a system without restriction on variables.

Theorem 4.18 (Rohn, 1981)**.** *An interval system $\boldsymbol{A}x = \boldsymbol{b}$, $x \geq 0$ is strongly solvable if and only if the system*

$$A_{s,e}x = b_s, \quad x \geq 0 \tag{4.10}$$

is solvable for each $s \in \{\pm 1\}^m$.

Proof. By negation, the interval system $\boldsymbol{A}x = \boldsymbol{b}$, $x \geq 0$ is not strongly solvable if and only if the system $Ax = b$, $x \geq 0$ is unsolvable for certain $A \in \boldsymbol{A}$ and $b \in \boldsymbol{b}$. Equivalently, by Farkas' lemma 1.33, the system $A^T u \geq 0$, $b^T u \leq -1$ is solvable. That is, the interval system $\boldsymbol{A}^T u \geq 0$, $\boldsymbol{b}^T u \leq -1$ is (weakly) solvable. By Gerlach's Theorem 4.1, the system

$$(A^c)^T u + (A^\Delta)^T |u| \geq 0, \ (b^c)^T u - (b^\Delta)^T |u| \leq -1$$

has a solution. Equivalently, there is $s \in \{\pm 1\}^m$ such that the linear system

$$(A^c)^T u + (A^\Delta)^T \operatorname{diag}(s) u \geq 0, \ (b^c)^T u - (b^\Delta)^T \operatorname{diag}(s) u \leq -1$$

has a solution. Eventually, by Farkas' lemma again, it is equivalent to unsolvability of

$$(A^c + \operatorname{diag}(s) A^\Delta)x = b^c - \operatorname{diag}(s) b^\Delta, \quad x \geq 0$$

for some $s \in \{\pm 1\}^m$. Substitution $s \coloneqq -s$ leads to (4.10). □

Theorem 4.19 (Rohn, 2003)**.** *Let $\boldsymbol{A} \in \mathbb{IR}^{m\times n}$ and $\boldsymbol{b} \in \mathbb{IR}^m$. An interval system $\boldsymbol{A}x = \boldsymbol{b}$ is strongly solvable if and only if the system*

$$A_{s,e}x^1 - A_{-s,e}x^2 = b_s, \quad x^1, x^2 \geq 0 \tag{4.11}$$

is solvable for each $s \in \{\pm 1\}^m$.

We omit the proof since the statement directly follows from Theorem 4.18 applied to the following result from Hladík (2017).

Theorem 4.20. *Let $\boldsymbol{A} \in \mathbb{IR}^{m\times n}$ and $\boldsymbol{b} \in \mathbb{IR}^m$. An interval system $\boldsymbol{A}x = \boldsymbol{b}$ is strongly solvable if and only if the system*

$$\boldsymbol{A}x^1 - \boldsymbol{A}x^2 = \boldsymbol{b}, \quad x^1, x^2 \geq 0$$

is strongly solvable.

Proof. The system $\boldsymbol{A}x = \boldsymbol{b}$ is not strongly solvable if and only if there are $A \in \boldsymbol{A}$ and $b \in \boldsymbol{b}$ such that $Ax = b$ is unsolvable. By Farkas' lemma 1.33, equivalently, the system

$$A^T u = 0, \quad b^T u \leq -1$$

is solvable. This means that the system

$$\boldsymbol{A}^T u = 0, \quad \boldsymbol{b}^T u \leq -1$$

is solvable. By Theorem 4.6, we can equivalently rewrite the system to

$$\boldsymbol{A}^T u \geq 0, \quad \boldsymbol{A}^T u \leq 0, \quad \boldsymbol{b}^T u \leq -1.$$

That is, there are $A, A' \in \boldsymbol{A}$ and $b \in \boldsymbol{b}$ such that the system

$$A^T u \geq 0, \quad A'^T u \leq 0, \quad b^T u \leq -1$$

is solvable. Calling Farkas' lemma once again,

$$Ax^1 - A'x^2 = b, \quad x^1, x^2 \geq 0$$

is unsolvable. □

The solutions of systems (4.11) have an interesting interpretation for the particular realizations of $\boldsymbol{A}x = \boldsymbol{b}$. Even thought the points $x_s^1 - x_s^2$, $s \in \{\pm 1\}^m$ need not belong to the solution set Σ of $\boldsymbol{A}x = \boldsymbol{b}$ (see Example 4.22 below), they do provide an enclosure of its (selected) solutions.

Theorem 4.21 (Rohn, 2003)**.** *If the interval system $\boldsymbol{A}x = \boldsymbol{b}$ is strongly solvable, then each realization has a solution in the set*

$$\operatorname{conv}_{s\in\{\pm 1\}^m}\{x_s^1 - x_s^2\}, \tag{4.12}$$

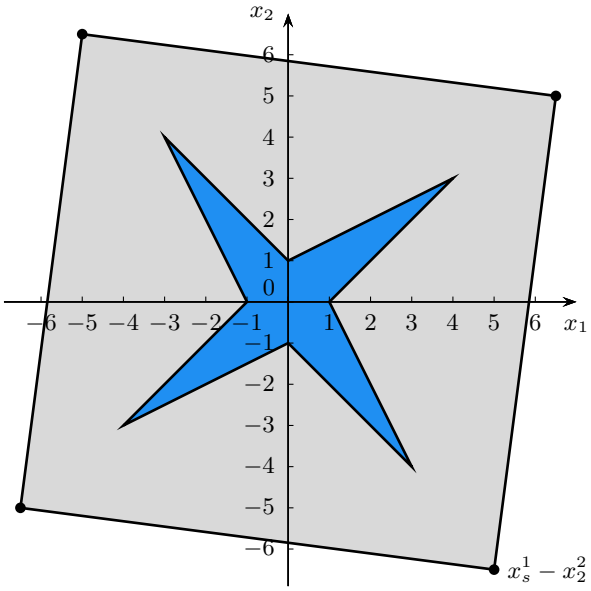
where x_s^1, x_s^2, $s \in \{\pm 1\}^m$, is a solution to (4.11).

Proof. Let $A \in \boldsymbol{A}$ and $b \in \boldsymbol{b}$. We want to find $\lambda_s \geq 0$, $s \in \{\pm 1\}^m$, such that $\sum_s \lambda_s = 1$ and $Ax = b$ for $x = \sum_s \lambda_s(x_s^1 - x_s^2)$. In other words, we want to show that the linear system

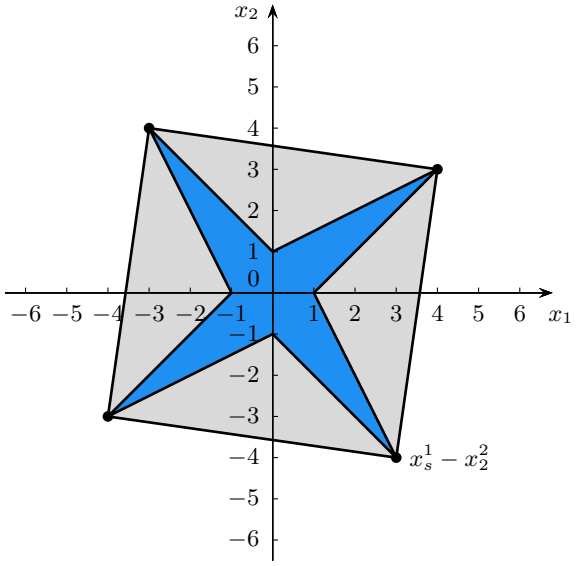
$$\sum_{s\in\{\pm 1\}^m} \lambda_s A(x_s^1 - x_s^2)) = b, \quad \sum_{s\in\{\pm 1\}^m} \lambda_s = 1, \ \lambda_s \geq 0, \ s \in \{\pm 1\}^m$$

has a solution with respect to λ_s, $s \in \{\pm 1\}^m$. Equivalently, by Farkas' theorem, for each $p \in \mathbb{R}^m$ and $q \in \mathbb{R}$ it should hold that

$$\big(\forall s \in \{\pm 1\}^m : p^T A(x_s^1 - x_s^2)) + q \geq 0\big) \quad \Rightarrow \quad p^T b + q \geq 0.$$



(a) One possible choice of the points $x_s^1 - x_s^2$, $s \in \{\pm 1\}^2$, which do not belong to Σ.



(b) One possible choice of the points $x_s^1 - x_s^2$, $s \in \{\pm 1\}^2$, which do belong to Σ.

Figure 4.3: (Example 4.22): Enclosure of the solution set Σ by means of the convex hull of the four points $x_s^1 - x_s^2$, $s \in \{\pm 1\}^2$.

Let $p \in \mathbb{R}^m$, $q \in \mathbb{R}$ and put $s \coloneqq \operatorname{sgn}(p)$. Now,

$$\begin{aligned} 0 &\leq p^T(Ax_s^1 - Ax_s^2) + q \\ &\leq p^T\left(A^c + \operatorname{diag}(s)A^\Delta\right)x_s^1 - \left(A^c - \operatorname{diag}(s)A^\Delta\right)x_s^2) + q \\ &= p^T(b^c - \operatorname{diag}(s)b^\Delta) + q \leq p^T b + q, \end{aligned}$$

which closes the proof. □

Example 4.22. Consider the Barth & Nuding system of interval linear equations (3.6). For particular $s \in \{\pm 1\}^2$ we compute solutions to (4.11) as follows

$$\begin{aligned} s = (1,1)^T: &\quad x_s^1 = (7,5)^T, &\quad x_s^2 = (0.5,0)^T, \\ s = (1,-1)^T: &\quad x_s^1 = (5,0.5)^T, &\quad x_s^2 = (0,7)^T, \\ s = (-1,1)^T: &\quad x_s^1 = (0,7)^T, &\quad x_s^2 = (5,0.5)^T, \\ s = (-1,-1)^T: &\quad x_s^1 = (0.5,0)^T, &\quad x_s^2 = (7,5)^T. \end{aligned}$$

Since $\boldsymbol{A}$ is regular, the set (4.12) provides an enclosure of the whole solution set Σ. This is illustrated in Figure 4.3a. We see that the particular points $x_s^1 - x_s^2$ do not lie in Σ. The systems (4.11), however, do not possess unique solutions. Consider another choice of solutions, which were indeed calculated by a linear programming solver,

$$\begin{aligned} s = (1,1)^T: &\quad x_s^1 = (4,3)^T, &\quad x_s^2 = (0,0)^T, \\ s = (1,-1)^T: &\quad x_s^1 = (3,0)^T, &\quad x_s^2 = (0,4)^T, \\ s = (-1,1)^T: &\quad x_s^1 = (0,4)^T, &\quad x_s^2 = (3,0)^T, \\ s = (-1,-1)^T: &\quad x_s^1 = (0,0)^T, &\quad x_s^2 = (4,3)^T. \end{aligned}$$

Now, the points $x_s^1 - x_s^2$ belong to Σ and, even more, they form the spires of the solution set. Hence $\operatorname{conv}\Sigma = \operatorname{conv}_{s\in\{\pm 1\}^2}\{x_s^1 - x_s^2\}$ in this case; see Figure 4.3b. □

The following characterization by Rohn (2014) on strong solvability might seem more complicated than the above one, but it gives a useful technical tool and also a nice geometric insight.

Theorem 4.23. *An interval system $\boldsymbol{A}x = \boldsymbol{b}$ is strongly solvable if and only if the system*

$$\operatorname{diag}(s)\boldsymbol{A}x \leq \operatorname{diag}(s)\boldsymbol{b}$$

is strongly solvable for each $s \in \{\pm 1\}^m$.

Proof. By negation, suppose that $Ax = b$ is not solvable for some $A \in \boldsymbol{A}$ and $b \in \boldsymbol{b}$. By Farkas' lemma 1.33, it is equivalent to solvability of

$$A^T u = 0, \quad b^T u \leq -1.$$

This means that there is $u \in \mathbb{R}^m$ being a weak solution of the interval system

$$\boldsymbol{A}^T u = 0, \quad \boldsymbol{b}^T u \leq -1.$$

Yet equivalently, there is $s \in \{\pm 1\}^m$ such that $v = \operatorname{diag}(s)u$ is a weak solution of the interval system

$$\boldsymbol{A}^T \operatorname{diag}(s)v = 0, \quad \boldsymbol{b}^T \operatorname{diag}(s)v \leq -1, \quad v \geq 0.$$

By Farkas' lemma again, there is $s \in \{\pm 1\}^m$ such that the interval system

$$\operatorname{diag}(s)\boldsymbol{A}x \leq \operatorname{diag}(s)\boldsymbol{b}$$

is not strongly solvable. □

Eventually, we show that checking strong solvability of interval linear equations is *co-NP*-hard, as well as checking strong nonnegative solvability (Rohn, 1998b).

Theorem 4.24. *Checking strong solvability of a square system $\boldsymbol{A}x = \boldsymbol{b}$ is co-NP-hard.*

Proof. We know (Theorem 3.60) that checking regularity of interval matrices is *co-NP*-hard. We show that a matrix $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$ is regular if and only if the interval system $\boldsymbol{A}x = [-e, e]$ is strongly solvable. If $\boldsymbol{A}$ is regular, then the system must be strongly solvable. If $\boldsymbol{A}$ is not regular, then there is a singular $A \in \boldsymbol{A}$. The system $Ax = b$ is unsolvable for any b, $\|b\|_\infty \leq 1$, lying outside the column space of A. □

Theorem 4.25 (Rohn, 1998b)**.** *Checking strong solvability of an interval system $\boldsymbol{A}x = \boldsymbol{b}$, $x \geq 0$ is co-NP-hard.*

Proof. By Theorem 4.24, checking strong solvability of $\boldsymbol{A}x = \boldsymbol{b}$ is *co-NP*-hard. By Theorem 4.20, strong solvability of $\boldsymbol{A}x = \boldsymbol{b}$ is equivalent to strong solvability of $\boldsymbol{A}x^1 - \boldsymbol{A}x^2 = \boldsymbol{b}$, $x^1, x^2 \geq 0$. □

Interval right-hand side. Now, let us focus on the case when the intervals are situated in the right-hand side vector only. It turns out that checking strong solvability is hard if we restrict to the nonnegative variables, but the problem becomes polynomial provided the variables are free.

Theorem 4.26 (Garajová et al., 2017)**.** *Checking strong solvability of an interval system $Ax = \boldsymbol{b}$, $x \geq 0$ is co-NP-hard.*

Proof. By Farkas' lemma 1.33, the system $Ax = b$, $x \geq 0$ is infeasible if and only if the system $A^T y \geq 0$, $b^T y < 0$ is feasible. Therefore, $Ax = \boldsymbol{b}$, $x \geq 0$ is strongly solvable if and only if the system $A^T y \geq 0$, $\boldsymbol{b}^T y < 0$ has no solution. However, checking solvability of this interval system is *NP*-hard by Theorem 4.5. □

Theorem 4.27. *Checking strong solvability of an interval system $Ax = \boldsymbol{b}$ is a polynomial problem.*

Proof. Suppose that the ith row of A is linearly dependent on the others. If $b_i^\Delta > 0$, then the system is not strongly solvable. Thus, we can assume that each such row has real right-hand side. If the ith row of $(A \mid \boldsymbol{b})$ is linearly dependent on the others, then we can omit it; otherwise, the system is not strongly solvable. In this way, we end up with a matrix that has full row rank, and so its columns generate the whole space. Therefore the system is strongly solvable. □

4.4.2 Interval inequalities

In contrast to equations, strong solvability of interval linear inequalities is characterized by linear constraints and so the decidability problem is polynomial.

Theorem 4.28 (Rohn and Kreslová, 1994)**.** *An interval system $\boldsymbol{A}x \leq \boldsymbol{b}$ is strongly solvable if and only if the system*

$$\overline{A}x^1 - \underline{A}x^2 \leq \underline{b},\ x^1 \geq 0,\ x^2 \geq 0 \tag{4.13}$$

is solvable.

Proof. The interval system $\boldsymbol{A}x \leq \boldsymbol{b}$ is not strongly solvable if and only if there are $A \in \boldsymbol{A}$ and $b \in \boldsymbol{b}$ such that $Ax \leq b$ is unsolvable. By Farkas' lemma 1.33, equivalently, the system

$$A^T u = 0,\quad b^T u \leq -1,\quad u \geq 0$$

is solvable. Thus, we have that the interval system

$$\boldsymbol{A}^T u = 0,\quad \boldsymbol{b}^T u \leq -1,\quad u \geq 0$$

is weakly solvable. By the Oettli–Prager and Gerlach theorems (or Theorem 4.13), the solution se is described by

$$\overline{A}^T u \geq 0,\quad -\underline{A}^T u \geq 0,\quad \underline{b}^T u \leq -1,\quad u \geq 0.$$

By Farkas' lemma again, the system (4.13) is unsolvable. □

Theorem 4.29 (Rohn and Kreslová, 1994)**.** *Suppose that the interval system $\boldsymbol{A}x \leq \boldsymbol{b}$ is strongly solvable, and define $x^* := x^1 - x^2$, where x^1, x^2 solves (4.13). Then x^* is a solution to $Ax \leq b$ for every $A \in \boldsymbol{A}$ and $b \in \boldsymbol{b}$.*

Proof. Let $A \in \boldsymbol{A}$ and $b \in \boldsymbol{b}$ be arbitrary. Then

$$Ax^* = A(x^1 - x^2) = Ax^1 - Ax^2 \leq \overline{A}x^1 - \underline{A}x^2 \leq \underline{b} \leq b.$$ □

Proof. One direction is trivial since $\overline{A}x \leq \underline{b},\ x \geq 0$ is a realization of the interval system. Conversely, let x^* be a solution to $\overline{A}x \leq \underline{b},\ x \geq 0$. Whenever $A \in \boldsymbol{A}$ and $b \in \boldsymbol{b}$, we have $Ax^* \leq \overline{A}x^* \leq \overline{b} \leq b$. □

Theorem 4.29 says an interesting property of interval linear inequalities. If they are strongly solvable, then they possess the so-called strong solution, which is a solution common to all realizations; see also Corollary 4.37. Naturally, interval equations cannot have a strong solution so easily. For example, the interval system

$$x + y = [1, 2],\quad x - y = [2, 3]$$

is strongly solvable, but there is no common solution.

In the case of non-negative variables, strong solvability is even more easy to check for. As a consequence, we get that the transformation of imposing nonnegativity does not change strong solvability, so that one can substitute $x = x^1 - x^2$, $x^1, x^2 \geq 0$, for a free variable x and can ignore dependencies.

Theorem 4.30 (Vajda, 1961)**.** *An interval system $\boldsymbol{A}x \leq \boldsymbol{b},\ x \geq 0$ is strongly solvable if and only if the system $\overline{A}x \leq \underline{b},\ x \geq 0$ is solvable.*

Corollary 4.31. *An interval system $\boldsymbol{A}x \leq \boldsymbol{b}$ is strongly solvable if and only if the interval system $\boldsymbol{A}x^1 - \boldsymbol{A}x^2 \leq \boldsymbol{b},\ x^1, x^2 \geq 0$, is strongly solvable.*

Other transformations, however, can change strong solvability (cf. Section 4.3). A typical example is splitting equations to double inequalities. For concreteness, consider the interval system $x = [1, 2]$, which is strongly solvable. The transformation yields the interval system $x \leq [1, 2],\ x \geq [1, 2]$. It is not strongly solvable since it involves the infeasible realization $x \leq 1,\ x \geq 2$.

4.4.3 General interval system

We extend the previous results to an interval system of mixed linear equations and inequalities (4.7) in the lines of Hladík (2013b).

Theorem 4.32. *The system (4.7) is strongly solvable if and only if the system*

$$A_{s,e}x + B_{s,e}y^1 - B_{-s,e}y^2 = b_s, \tag{4.14a}$$
$$\overline{C}x + \overline{D}y^1 - \underline{D}y^2 \le \underline{d}, \tag{4.14b}$$
$$x, y^1, y^2 \ge 0. \tag{4.14c}$$

is solvable for every $s \in \{\pm 1\}^{m_1}$.

Proof. By negation, suppose that (4.8) is not solvable for some realization of interval data. By Farkas' theorem (Corollary 1.34), there are $u \in \mathbb{R}^{m_1}$ and $v \in \mathbb{R}^{m_2}$, $v \ge 0$, such that

$$A^T u + C^T v \ge 0,\ B^T u + D^T v = 0,\ b^T u + d^T v \le -1.$$

That is, (u, v) forms a weak solution to the interval system

$$\boldsymbol{A}^T u + \boldsymbol{C}^T v \ge 0,\ \boldsymbol{B}^T u + \boldsymbol{D}^T v = 0,\ \boldsymbol{b}^T u + \boldsymbol{d}^T v \le -1,\ v \ge 0$$

and by Corollary 4.14 there is $s \in \{\pm 1\}^{m_1}$ such that

$$\begin{aligned}(A_{-s,e})^T y + \overline{C}^T v &\ge 0,\\ (B_{-s,e})^T y + \overline{D}^T v &\ge 0,\\ -(B_{s,e})^T y - \underline{D}^T v &\ge 0,\\ (b_{-s})^T y + \underline{d}^T v &\le -1,\ v \ge 0.\end{aligned}$$

Again, we utilize Farkas' theorem, saying that there is no non-negative x, y^1, y^2 satisfying

$$\begin{aligned}A_{-s,e}x + B_{-s,e}y^1 - B_{s,e}y^2 &= b_{-s},\\ \overline{C}x + \overline{D}y^1 - \underline{D}y^2 &\le \underline{d}.\end{aligned}$$

This is equivalent to (4.14) by substituting $s \coloneqq -s$. □

Provided we check strong solvability by Theorem 4.32, we can also output a bounded set in which each realization (4.8) has at least one solution; cf. Rohn (2006a). This generalizes the result provided by Theorem 4.21.

Theorem 4.33. *If the system (4.7) is strongly solvable, then each realization has a solution in the set*

$$\operatorname{conv}_{s\in\{\pm 1\}^{m_1}}\{(x_s, y_s^1 - y_s^2)\},$$

where x_s, y_s^1, y_s^2, $s \in \{\pm 1\}^{m_1}$, *is a solution to (4.14).*

Proof. Let $A \in \boldsymbol{A}$, $B \in \boldsymbol{B}$, $C \in \boldsymbol{C}$, $D \in \boldsymbol{D}$, $b \in \boldsymbol{b}$, and $d \in \boldsymbol{d}$. Then for any convex combination

$$x = \sum_{s\in\{\pm 1\}^{m_1}} \lambda_s x_s, \quad y = \sum_{s\in\{\pm 1\}^{m_1}} \lambda_s (y_s^1 - y_s^2),$$

with $\sum_{s\in\{\pm 1\}^{m_1}} \lambda_s = 1$ and $\lambda_s \ge 0$, $s \in \{\pm 1\}^{m_1}$, we have

$$\begin{aligned}Cx + Dy &= \sum_{s\in\{\pm 1\}^{m_1}} \lambda_s C x_s + \sum_{s\in\{\pm 1\}^{m_1}} \lambda_s D(y_s^1 - y_s^2)\\ &\le \sum_{s\in\{\pm 1\}^{m_1}} \lambda_s \overline{C} x_s + \sum_{s\in\{\pm 1\}^{m_1}} \lambda_s (\overline{D} y_s^1 - \underline{D} y_s^2)\\ &\le \sum_{s\in\{\pm 1\}^{m_1}} \lambda_s \underline{d} = \underline{d} \le d.\end{aligned}$$

Thus, it suffices to find an appropriate convex combination satisfying the equations. That is, we want to show that the system

$$\sum_{s\in\{\pm1\}^{m_1}} \lambda_s(Ax_s + B(y_s^1 - y_s^2)) = b, \quad \sum_{s\in\{\pm1\}^{m_1}} \lambda_s = 1,\ \lambda_s \geq 0,\ s \in \{\pm1\}^{m_1}$$

has a solution with respect to λ_s, $s \in \{\pm1\}^{m_1}$. Equivalently, by Farkas' theorem, for each $u \in \mathbb{R}^{m_1}$ and $v \in \mathbb{R}$ it should hold that

$$\big(\forall s \in \{\pm1\}^{m_1} : u^T(Ax_s + B(y_s^1 - y_s^2)) + v \geq 0\big) \quad \Rightarrow \quad u^T b + v \geq 0.$$

Let $u \in \mathbb{R}^{m_1}$, $v \in \mathbb{R}$ and put $s \coloneqq -\operatorname{sgn}(u)$. Now,

$$\begin{aligned} 0 &\leq u^T\big(Ax_s + B(y_s^1 - y_s^2)\big) + v \\ &\leq u^T\big(A_{s,e}x_s + B_{s,e}y_s^1 - B_{-s,e}y_s^2\big) + v \\ &= u^T b_s + v \leq u^T b + v, \end{aligned}$$

which closes the proof. □

A sufficient condition for strong solvability

In view of intractability of checking strong solvability, a sufficient condition is welcome. We propose the following method:

1: Solve a linear program to obtain a suitable initial solution.

2: Based on the previous result, transform the equations to a square system.

3: Compute an enclosure $\boldsymbol{x}, \boldsymbol{y}$ to the square system.

4: Check whether $\boldsymbol{x}, \boldsymbol{y}$ fulfills the inequalities $\boldsymbol{x} \geq 0$ and $\boldsymbol{C}\boldsymbol{x} + \boldsymbol{D}\boldsymbol{y} \leq \boldsymbol{d}$.

Now, we describe particular steps in detail. In Step 1, we solve the linear program

$$\max \alpha \ \text{ subject to } \ A^c x + B^c y = b^c,\ C^c x + D^c y + \alpha e \leq \underline{d},\ x \geq 0.$$

In the case there are no inequalities, we solve

$$\max \alpha \ \text{ subject to } \ A^c x + B^c y = b^c, \quad x \geq \alpha e.$$

instead. The idea is to determine a sufficiently robust feasible solution. Let x^*, y^* be an optimal solution. If the linear program is infeasible, then strong solvability is not satisfied. If the linear program is unbounded, then we take any solution on its unbounded edge.

In Step 2, we transform the system of equations $A^c x + B^c y = b^c$ to a square one. Suppose $m_1 < n_1 + n_2$. For each i such that $x_i^* = 0$ we remove the ith column of $\boldsymbol{A}$ and $\boldsymbol{C}$, and the ith entry of x^*. Thus, the number of variables decreases. If there are more vanishing entries of x^* than the value of $m + n - k$, then we remove only $n_1 + n_2 - m_1$ columns. After the dimension reduction, if the system is not square yet, we add a suitable number of artificial equations $A'x + B'y = b'$. The natural choice for $(A'\ B')$ is the orthogonal basis of the null space to $(A^c\ B^c)$, and the right-hand side is calculated $b' \coloneqq A'x^* + B'y^*$. The idea behind this choice is that the solution of the resulting system is supposed to be less sensitive to perturbations of $A \in \boldsymbol{A}$, $B \in \boldsymbol{B}$ and $b \in \boldsymbol{b}$ (the orthogonal matrices have the smallest condition number).

In Step 3, we solve the interval linear system of equations

$$\boldsymbol{A}x + \boldsymbol{B}y = \boldsymbol{b}, \quad A'x + B'y = b'.$$

We can employ any of the methods discussed in Sections 3.2–3.5. Even though $\boldsymbol{x}, \boldsymbol{y}$ do not contain all weak solutions of $\boldsymbol{A}x + \boldsymbol{B}y = \boldsymbol{b}$, they do contain at least one solution of any realization $Ax + By = b$ with $A \in \boldsymbol{A}$, $B \in \boldsymbol{B}$ and $b \in \boldsymbol{b}$.

In Step 4, we check whether $\boldsymbol{x} \geq 0$ and $\boldsymbol{Cx} + \boldsymbol{Dy} \leq \boldsymbol{d}$, that is, whether $\underline{x} \geq 0$ and $\sup(\boldsymbol{Cx} + \boldsymbol{Dy}) \leq \underline{d}$. If yes, then each realization of interval data has at least one solution fulfilling the constraints and we output that the system is strongly solvable. If no, then we cannot decide.

The method requires solving one linear program, an orthogonal basis of the null space and an enclosure to the solution set of interval linear equations, so the overall computational cost is very low. Moreover, it is easy to implement it in a reliable way. Since the linear program and orthogonal basis play a role of a heuristic only, they need not be computed verified. The other computation is done by using interval arithmetic.

Notice that the method gives a new and strong sufficient condition for strong solvability of specific cases of interval linear systems, namely, for $\boldsymbol{A}x = \boldsymbol{b}$, $x \geq 0$ and for $\boldsymbol{B}y = \boldsymbol{b}$.

Example 4.34. Let

$$A = \begin{pmatrix} -3 \end{pmatrix}, \; B = \begin{pmatrix} 2 & 1 \end{pmatrix}, \; b = \begin{pmatrix} 4 \end{pmatrix}, \; C = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \; D = \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix}, \; d = \begin{pmatrix} 10 \\ 5 \end{pmatrix}.$$

For a given parameter $\delta \geq 0$, we enlarge real quantities to intervals having radius δ. Thus, $\boldsymbol{A}$ is defined as $A \coloneqq [A - \delta ee^T, A + \delta ee^T]$ and analogously for $\boldsymbol{B}, \boldsymbol{C}, \boldsymbol{D}, \boldsymbol{b}, \boldsymbol{d}$. The maximal value of δ, for which our method confirms strong solvability of (4.7), is $\delta \approx 0.6929$. The maximal value of δ, for which (4.7) is really strong solvable, is $\delta \approx 0.7391$. □

4.5 *AE* solutions of general interval linear systems

Herein, we again consider a general interval linear system (cf. (4.7))

$$\boldsymbol{A}x + \boldsymbol{B}y = \boldsymbol{b}, \quad \boldsymbol{C}x + \boldsymbol{D}y \leq \boldsymbol{d}, \; x \geq 0, \tag{4.15}$$

where $\boldsymbol{A} \in \mathbb{IR}^{m_1 \times n_1}$, $\boldsymbol{B} \in \mathbb{IR}^{m_1 \times n_2}$, $\boldsymbol{C} \in \mathbb{IR}^{m_2 \times n_1}$, $\boldsymbol{D} \in \mathbb{IR}^{m_2 \times n_2}$, $\boldsymbol{b} \in \mathbb{IR}^{m_1}$, and $\boldsymbol{d} \in \mathbb{IR}^{m_2}$. Similarly as in Section 3.10 we suppose that each interval is associated with quantifier $\forall$ or $\exists$. First, we focus on strong solutions, which are solutions defined by $\forall$-quantification.

4.5.1 Strong solutions

Recall that interval system (4.15) is called strongly solvable if each realization is solvable (not necessarily by the same pair (x, y)). Now, we introduce *a strong solution* of (4.7) as any $(x, y) \in \mathbb{R}^{n_1+n_2}$ such that it solves (4.8) for every realization $A \in \boldsymbol{A}$, $B \in \boldsymbol{B}$, $C \in \boldsymbol{C}$, $D \in \boldsymbol{D}$, $b \in \boldsymbol{b}$, and $d \in \boldsymbol{d}$. Clearly, existence of a strong solution implies strong solvability, but not the conversely (an exception are pure inequalities; see Corollary 4.37 below).

The strong solution set can be viewed as the intersection of all feasible solution sets of all realizations. Therefore, the strong solution set is always convex.

Theorem 4.35. *A pair $(x, y) \in \mathbb{R}^{n_1+n_2}$ is a strong solution to (4.7) if and only if it satisfies*

$$b^\Delta = 0, \; A^\Delta x = 0, \; B^\Delta |y| = 0, \; A^c x + B^c y = b^c, \; \overline{C}x + D^c y + D^\Delta |y| \leq \underline{d}, \; x \geq 0.$$

Proof. The pair $(x, y) \in \mathbb{R}^{n_1+n_2}$ is a strong solution only if it satisfies

$$Ax + By - b = 0, \; Cx + Dy - d \leq 0, \; x \geq 0$$

for every realization. Equivalently,

$$\boldsymbol{A}x + \boldsymbol{B}y - \boldsymbol{b} = 0, \; \boldsymbol{C}x + \boldsymbol{D}y - \boldsymbol{d} \leq 0, \; x \geq 0,$$

where the left-hand sides of equations and inequalities are calculated by interval arithmetic. Two intervals are identical if and only if they have the same midpoint and radius. So the interval condition $\boldsymbol{A}x + \boldsymbol{B}y - \boldsymbol{b} = 0$ is equivalent to the equality of their centers, i.e., $A^c x + B^c y = b^c$, and equality of their radii, i.e., $A^\Delta x + B^\Delta |y| + b^\Delta = 0$. The latter can be written as $b^\Delta = 0$, $A^\Delta x = 0$, $B^\Delta |y| = 0$. The condition $\boldsymbol{C}x + \boldsymbol{D}y - \boldsymbol{d} \leq 0$ holds true if and only if $\overline{\boldsymbol{C}x} + \overline{\boldsymbol{D}y} - \underline{d} \leq 0$. Since $\overline{\boldsymbol{C}x} = \overline{C}x$ and $\overline{\boldsymbol{D}y} = D^c y + D^\Delta |y|$ (see Proposition 2.11), we have the characterization in question. □

The strong solutions of an interval system of linear equations $\boldsymbol{A}x = \boldsymbol{b}$ in particular are characterized by

$$b^{\Delta} = 0,\ A^{\Delta}|x| = 0,\ A^c x = b^c.$$

Therefore, $\boldsymbol{b}$ must be a real vector and the strong solutions are equivalently described by linear equations

$$A^c x = b^c,\quad x_i = 0\ \forall i : (A^{\Delta})_{*i} \neq 0.$$

This shows the the strong solution set is geometrically an affine space.

For an intervl system of equations, a strong solution is unlikely to exist, however, for inequalities, this concept is more useful. Therefore, we summarize various characterizations of strong solutions of interval linear inequalities. The second item comes from Theorem 4.35 and the third one from Theorem 4.29.

Corollary 4.36. *For a vector $x \in \mathbb{R}^n$, the following are equivalent:*

(1) x is a strong solution to $\boldsymbol{A}x \leq \boldsymbol{b}$,

(2) $A^c x + A^{\Delta}|x| \leq \underline{b}$,

(3) $x = x^1 - x^2,\ \overline{A}x^1 - \underline{A}x^2 \leq \underline{b},\ x^1, x^2 \geq 0$.

Eventually, we state the observation we have already mentioned.

Corollary 4.37. *An interval system $\boldsymbol{A}x \leq \boldsymbol{b}$ is strongly solvable if and only if it has a strong solution.*

The solution set characterized by the Gerlach Theorem 4.1 can also be viewed as the union of all polyhedra described by $Ax \leq b$, $A \in \boldsymbol{A}$, $b \in \boldsymbol{b}$, that is,

$$\cup_{A \in \boldsymbol{A},\, b \in \boldsymbol{b}} \{x;\ Ax \leq b\},$$

whereas the set of strong solutions is the intersection of all polyhedra described by $Ax \leq b$, $A \in \boldsymbol{A}$, $b \in \boldsymbol{b}$, that is,

$$\cap_{A \in \boldsymbol{A},\, b \in \boldsymbol{b}} \{x;\ Ax \leq b\},$$

This geometric viewpoint is also illustrated by the following example.

Example 4.38. Consider the interval system of linear inequalities $\boldsymbol{A}x \leq \boldsymbol{b}$ in the form

$$\begin{pmatrix} [2,4] & [4,6] \\ -5 & -[1,1] \\ [2,4] & -[4,6] \\ -[1,3] & -[2,4] \end{pmatrix} x \leq \begin{pmatrix} [20,25] \\ [20,25] \\ [20,25] \\ 18 \end{pmatrix}.$$

Figure 4.4 illustrates the weak solution set, which is a non-convex polyhedron, and the strong solution set, which is always a convex polyhedron. Each realization $Ax \leq b$ of $\boldsymbol{A}x \leq \boldsymbol{b}$ describes a convex polyhedron lying in between these two sets w.r.t. inclusion. The figure depicts separately two realizations, $\underline{A}x \leq \overline{b}$ and $\overline{A}x \leq \underline{b}$, respectively. □

Even though the linear characterization given by Corollary 4.36(3) gives an efficient tool for handling strong solutions, some sideway problems can still be tough. The characterization describes the strong solutions by using auxiliary variables x^1 and x^2. To project the polyhedron in the space of x and to obtain a description of the strong solutions by linear inequalities in x cannot be done effectively. As Example 4.39 illustrates, the number of inequalities can be exponential. Therefore, Corollary 4.36(3) gives a convenient extended formulation of the strong solution set (Avis and Tiwary, 2015).

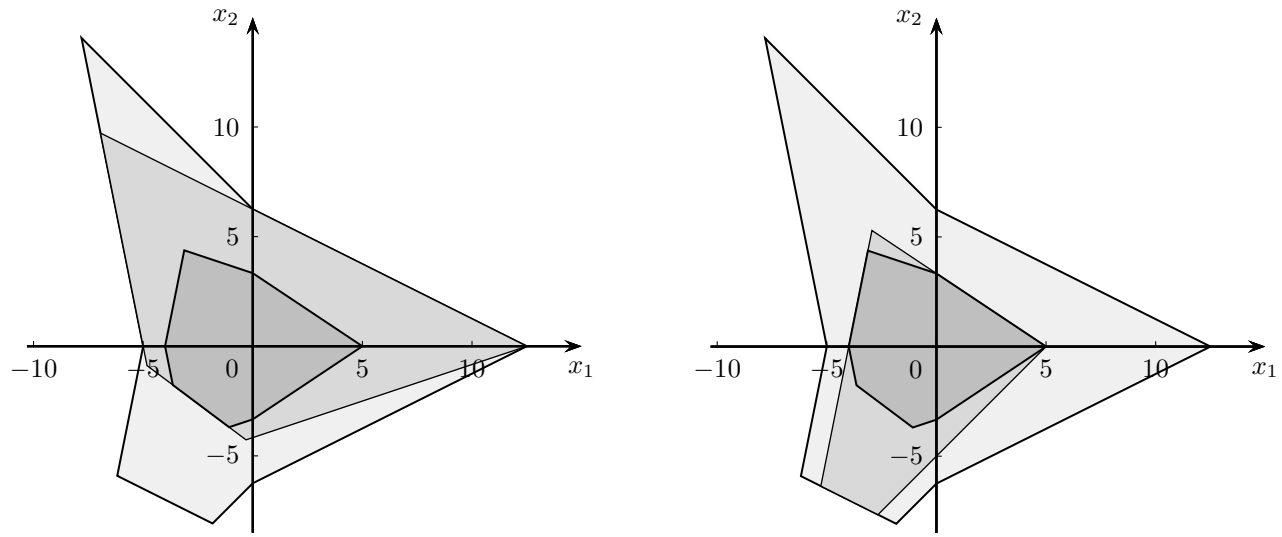
Example 4.39. Consider the interval linear inequality in n variables

$$[-e, e]^T x \leq e$$

By Corollary 4.36(2), the strong solution set is described by $e^T|x| \leq 1$. This inequality characterizes the unit ball in the Manhattan norm, which has 2^n faces. Thus, the minimum number of linear inequalities is 2^n, and they have the form of $\pm x_1 \pm \cdots \pm x_n \leq 1$, considering all combinations of $\pm$ sign. □

Domes and Neumaier (2016) use a similar concept for feasibility of uncertain optimization problem.

Notes and further reading. Huang et al. (2021) addressed the reversed quantification – *EA* solutions of general interval systems. Sharaya (2014a); Vatolin (1984) considered even more generally quantified interval problems; surprisingly, they can be characterized by a simple formula.



(a) The intermediate realization is $\underline{A}x \leq \overline{b}$.

(b) The intermediate realization is $\overline{A}x \leq \underline{b}$.

Figure 4.4: (Example 4.38) The strong solution set in dark gray and the weak solution set in light gray. The intermediate polyhedron in medium gray represents one particular realization of $\boldsymbol{A}x \leq \boldsymbol{b}$.

Chapter 5

Properties of interval matrices

5.1 Eigenvalues of interval matrices

5.1.1 Symmetric case

Recall that a symmetric matrix A has always real eigenvalues and we sort them non-increasingly

$$\lambda_1(A) \geq \cdots \geq \lambda_n(A).$$

Now, let $\boldsymbol{A}^S$ be a symmetric interval matrix and introduce the eigenvalue sets

$$\boldsymbol{\lambda}_i(\boldsymbol{A}^S) \coloneqq \{\lambda_i(A);\ A \in \boldsymbol{A}^S\}, \quad i = 1, \ldots, n.$$

Each eigenvalue set $\boldsymbol{\lambda}_i(\boldsymbol{A}^S)$ consists of ith eigenvalues of all symmetric matrices in $\boldsymbol{A}$. By the continuity of eigenvalues (Horn and Johnson, 1985; Meyer, 2000) and compactness and convexity of $\boldsymbol{A}^S$ it is easy to see that $\boldsymbol{\lambda}_i(\boldsymbol{A}^S)$, $i = 1, \ldots, n$ are compact intervals. They may be disjoint or they may overlap, but one interval can never lie in the interior of another one.

Example 5.1. For

$$\boldsymbol{A} = \begin{pmatrix} [1,2] & 0 & 0 \\ 0 & [7,8] & 0 \\ 0 & 0 & [4,10] \end{pmatrix}$$

we have $\boldsymbol{\lambda}_1(\boldsymbol{A}^S) = [7,10]$, $\boldsymbol{\lambda}_2(\boldsymbol{A}^S) = [4,8]$ and $\boldsymbol{\lambda}_3(\boldsymbol{A}^S) = [1,2]$; see Figure 5.1.

□

Calculation of the eigenvalue sets is intractable. We show that just checking if zero lies in some eigenvalue set is *NP*-hard.

Theorem 5.2. *Checking whether* $0 \in \boldsymbol{\lambda}_i(\boldsymbol{A}^S)$ *for some* $i = 1, \ldots, n$ *is NP-hard.*

Proof. By Theorem 3.60 we know that checking regularity of $\boldsymbol{A} \in \mathbb{IR}^{n \times n}$ is a *co-NP*-hard problem. Put

$$\boldsymbol{M} \coloneqq \begin{pmatrix} 0 & \boldsymbol{A} \\ \boldsymbol{A}^T & 0 \end{pmatrix}.$$

The interval matrix $\boldsymbol{A}$ is regular if and only if $\boldsymbol{M}^S$ is regular. Now, $\boldsymbol{M}^S$ is regular if and only if no $M \in \boldsymbol{M}^S$ has zero eigenvalue. □

A simple enclosure of the eigenvalue set is obtained by a Weyl theorem; cf. Hladík et al. (2010); Rohn (2012a).

Theorem 5.3. *We have*

$$\boldsymbol{\lambda}_i(\boldsymbol{A}^S) \subseteq [\lambda_i(A^c) - \rho(A^\Delta), \lambda_i(A^c) + \rho(A^\Delta)], \quad i = 1, \ldots, n.$$

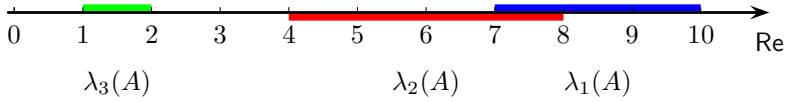


Figure 5.1: (Example 5.1) Eigenvalue sets of a symmetric interval matrix.

Proof. Let $A \in \boldsymbol{A}^S$, so $|A - A^c| \leq A^\Delta$. By Weyl's Theorem 1.12 and Theorem 1.8 we have

$$\lambda_i(A) = \lambda_i(A^c + (A - A^c)) \leq \lambda_i(A^c) + \lambda_1(A - A^c) \leq \lambda_i(A^c) + \rho(|A - A^c|) \leq \lambda_i(A^c) + \rho(A^\Delta).$$

Similarly for the lower bound. □

The largest and smallest eigenvalues of $\boldsymbol{A}^S$ can be calculated by the exponential formula by Hertz (1992). These two extremal eigenvalues are attained by matrices of the form $A_{-z,z}$ and A_{zz}. The other boundary points of the eigenvalue sets need not be attained at these matrices, moreover, they need not be attained at vertex matrices. The situation here is more complex and still challenging; cf. Hladík et al. (2011c).

Theorem 5.4 (Hertz, 1992)**.** *We have*

$$\overline{\lambda}_1(\boldsymbol{A}^S) = \max_{z \in \{\pm 1\}^n} \lambda_1(A_{-z,z}), \quad \underline{\lambda}_n(\boldsymbol{A}^S) = \min_{z \in \{\pm 1\}^n} \lambda_n(A_{zz}).$$

Proof. Notice that $(A_{-z,z})_{ij} = \overline{a}_{ij}$ if $z_i = z_j$ and $(A_{-z,z})_{ij} = \underline{a}_{ij}$ otherwise. Suppose to the contrary that there is $A \in \boldsymbol{A}^S$ such that $\lambda_1(A) > \max_{z \in \{\pm 1\}^n} \lambda_1(A_{-z,z})$. Thus, $Ax = \lambda_1(A)x$ for some $x \neq 0$ with $\|x\|_2 = 1$. Put $z^* \coloneqq \operatorname{sgn}(x)$, and by the Rayleigh–Ritz Theorem 1.11 we have

$$\begin{aligned} \lambda_1(A) = x^T A x = \sum_{i,j=1}^n x_i x_j A_{ij} \leq \sum_{i,j=1}^n x_i x_j (A_{-z^*,z^*})_{ij} = x^T A_{-z^*,z^*} x \\ \leq \max_{y : \|y\|_2 = 1} y^T A_{-z^*,z^*} y = \lambda_1(A_{-z^*,z^*}). \end{aligned}$$

This is a contradiction. The lower limit is proved accordingly. □

Notice that $A_{-z,z} = A_{-(-z),-z}$, so in the Hertz formula it is sufficient to go through $z \in 1 \times \{\pm 1\}^{n-1}$ instead of $z \in \{\pm 1\}^n$. This saves half the execution time. The exponentiality of the Hertz formula is justified by intractability of computation of the extremal eigenvalues.

Theorem 5.5. *All the conditions* $\overline{\lambda}_1(\boldsymbol{A}^S) > 0$, $\overline{\lambda}_1(\boldsymbol{A}^S) \geq 0$, $\underline{\lambda}_n(\boldsymbol{A}^S) < 0$ *and* $\underline{\lambda}_n(\boldsymbol{A}^S) \leq 0$ *are NP-hard to check for.*

Proof. The strict inequalities follow from *co-NP*-hardness of checking positive semidefiniteness (Section 5.2). The others follow from *co-NP*-hardness of checking positive definiteness (Theorem 5.17), and hence *NP*-hardness remains true even on the class of interval matrices with A^c non-negative positive definite rational and $A^\Delta = ee^T$. □

In some specific cases, the extremal eigenvalues are polynomially solvable.

Theorem 5.6.

(1) If A^c *is* essentially non-negative*, i.e.,* $A^c_{ij} \geq 0 \ \forall i \neq j$*, then* $\overline{\lambda}_1(\boldsymbol{A}^S) = \lambda_1(\overline{A})$.

(2) If A^Δ *is* diagonal*, then* $\overline{\lambda}_1(\boldsymbol{A}^S) = \lambda_1(\overline{A})$ *and* $\underline{\lambda}_n(\boldsymbol{A}^S) = \lambda_n(\underline{A})$.

Proof.

(1) For the sake of simplicity, suppose $A^c \geq 0$ first. Then for any $A \in \boldsymbol{A}^S$ we have

$$|A| \leq |A - A^c| + A^c \leq A^\Delta + A^c = \overline{A},$$

whence by the Perron–Frobenius Theorem 1.7 and Theorem 1.8,

$$\lambda_1(A) = \rho(A) \leq \rho(\overline{A}) = \lambda_1(\overline{A}).$$

If A^c is not nonnegative, then $A^c + \alpha I_n \geq 0$ for some $\alpha \in \mathbb{R}$. Thus, we shift the interval matrix $\boldsymbol{A}$ by αI_n, and the proof follows from

$$\overline{\lambda}_1(\boldsymbol{A}^S) = \overline{\lambda}_1(\boldsymbol{A}^S + \alpha I_n) - \alpha = \lambda_1(\overline{A} + \alpha I_n) - \alpha = \lambda_1(\overline{A}).$$

(2) By Hertz's theorem,

$$\overline{\lambda}_1(\boldsymbol{A}^S) = \max_{z \in \{\pm 1\}^n} \lambda_1(A^c + \operatorname{diag}(z) A^\Delta \operatorname{diag}(z)) = \lambda_1(A^c + A^\Delta) = \lambda_1(\overline{A}),$$

and similarly for $\underline{\lambda}_n(\boldsymbol{A}^S)$. □

Remark 5.7. Surprisingly, $\underline{\lambda}_1(\boldsymbol{A}^S)$ and $\overline{\lambda}_n(\boldsymbol{A}^S)$ are polynomially computable by calling a suitable semidefinite program. For simplicity, we discuss $\overline{\lambda}_n(\boldsymbol{A}^S)$ only. It may be formulated as an optimal value of the optimization problem

$$\overline{\lambda}_n(\boldsymbol{A}^S) = \max\ \alpha \ \text{ subject to } \ A - \alpha I_n \text{ is positive semidefinite, } \ A \in \boldsymbol{A}^S.$$

The constraint $A \in \boldsymbol{A}^S$ can be expressed in terms of positive semidefiniteness, too. Consider a block diagonal matrix $M(A, \alpha)$ with blocks

$$A - \alpha I_n,\ a_{ij} - \underline{a}_{ij},\ \overline{a}_{ij} - a_{ij},\ i \leq j.$$

Then the optimization problem reads

$$\overline{\lambda}_n(\boldsymbol{A}^S) = \max\ \alpha \ \text{ subject to } \ M(A, \alpha) \text{ is positive semidefinite.}$$

Since the objective function is linear and $M(A, \alpha)$ depends linearly (or affinely) on variables a_{ij}, $i \leq j$, and α, the optimization problem has the structure of a semidefinite program and can be solved in polynomial time with arbitrary precision (Gärtner and Matoušek, 2012; Nesterov and Nemirovskii, 1994; Vandenberghe and Boyd, 1996).

5.1.2 General case

We have shown that calculation eigenvalue bounds for symmetric interval matrices is a computationally hard problem. Bounding (complex) eigenvalues for general interval matrix $\boldsymbol{A} \in \mathbb{IR}^{n \times n}$ is much more difficult.

The following enclosure for all eigenvalues of all $A \in \boldsymbol{A}$ is by Hladík (2013a), and it is a special version of a more general result for complex interval matrices.

Theorem 5.8. *For any $A \in \boldsymbol{A}$ and its eigenvalue $\nu = \lambda + i\mu$ we have*

$$\underline{\lambda}_n\big(\tfrac{1}{2}(\boldsymbol{A} + \boldsymbol{A}^T)^S\big) \leq \lambda \leq \overline{\lambda}_1\big(\tfrac{1}{2}(\boldsymbol{A} + \boldsymbol{A}^T)^S\big), \tag{5.1a}$$

$$\underline{\lambda}_n \begin{pmatrix} 0 & \tfrac{1}{2}(\boldsymbol{A} - \boldsymbol{A}^T) \\ \tfrac{1}{2}(\boldsymbol{A}^T - \boldsymbol{A}) & 0 \end{pmatrix}^S \leq \mu \leq \overline{\lambda}_1 \begin{pmatrix} 0 & \tfrac{1}{2}(\boldsymbol{A} - \boldsymbol{A}^T) \\ \tfrac{1}{2}(\boldsymbol{A}^T - \boldsymbol{A}) & 0 \end{pmatrix}^S. \tag{5.1b}$$

Proof. Basically, it follows from Bendixson's Theorem 1.13, but we present a more elementary proof here. Let $A \in \boldsymbol{A}$, its eigenvalue $\nu = \lambda + i\mu$ and the corresponding eigenvector $x + iy$ normalized such that $1 = \|x + iy\|_2^2 = x^T x + y^T y$. The eigenvalue equation $A(x + iy) = (\lambda + i\mu)(x + iy)$ reads

$$Ax + iAy = \lambda x - \mu y + i(\lambda y + \mu x)$$

or, by splitting the real and imaginary parts

$$Ax = \lambda x - \mu y, \quad Ay = \lambda y + \mu x. \tag{5.2}$$

Multiplying the first equation by x^T, the second by y^T, and summing we get

$$x^T Ax + y^T Ay = \lambda x^T x - \mu x^T y + \lambda y^T y + \mu y^T x = \lambda(x^T x + y^T y) = \lambda.$$

Now, by the Rayleigh–Ritz Theorem 1.11

$$\begin{aligned}
\lambda_1(\tfrac{1}{2}(A + A^T)) &= \lambda_1 \begin{pmatrix} \frac{1}{2}(A + A^T) & 0 \\ 0 & \frac{1}{2}(A + A^T) \end{pmatrix} \\
&= \max_{u,v:\, u^T u + v^T v = 1} \begin{pmatrix} u^T & v^T \end{pmatrix} \begin{pmatrix} \frac{1}{2}(A + A^T) & 0 \\ 0 & \frac{1}{2}(A + A^T) \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} \\
&= \max_{u,v:\, u^T u + v^T v = 1} \frac{1}{2} u^T (A + A^T) u + \frac{1}{2} v^T (A + A^T) v \\
&= \max_{u,v:\, u^T u + v^T v = 1} u^T A u + v^T A v \geq x^T A x + y^T A y = \lambda.
\end{aligned}$$

Similarly for the lower bound.

Multiplying the first equation in (5.2) by $-y^T$, the second by x^T, and summing we get

$$x^T Ay - y^T Ax = -\lambda y^T x + \mu y^T y + \lambda x^T y + \mu x^T x = \mu(x^T x + y^T y) = \mu.$$

Hence $x^T(A - A^T)y = \mu$. Now,

$$\begin{aligned}
\lambda_1 \begin{pmatrix} 0 & \frac{1}{2}(A - A^T) \\ \frac{1}{2}(A^T - A) & 0 \end{pmatrix} &= \max_{u,v:\, u^T u + v^T v = 1} \begin{pmatrix} u^T & v^T \end{pmatrix} \begin{pmatrix} 0 & \frac{1}{2}(A - A^T) \\ \frac{1}{2}(A^T - A) & 0 \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} \\
&= \max_{u,v:\, u^T u + v^T v = 1} \frac{1}{2} u^T (A - A^T) v + \frac{1}{2} v^T (A^T - A) u \\
&= \max_{u,v:\, u^T u + v^T v = 1} u^T (A - A^T) v \geq x^T (A - A^T) y = \mu.
\end{aligned}$$

Similarly for the lower bound. □

Theorem 5.8 reduces the problem of enclosing complex eigenvalues of interval matrices to bounding real eigenvalues of symmetric interval matrices. Thus, we can apply any enclosing method for the symmetric case. In particular, Theorem 5.3 yields

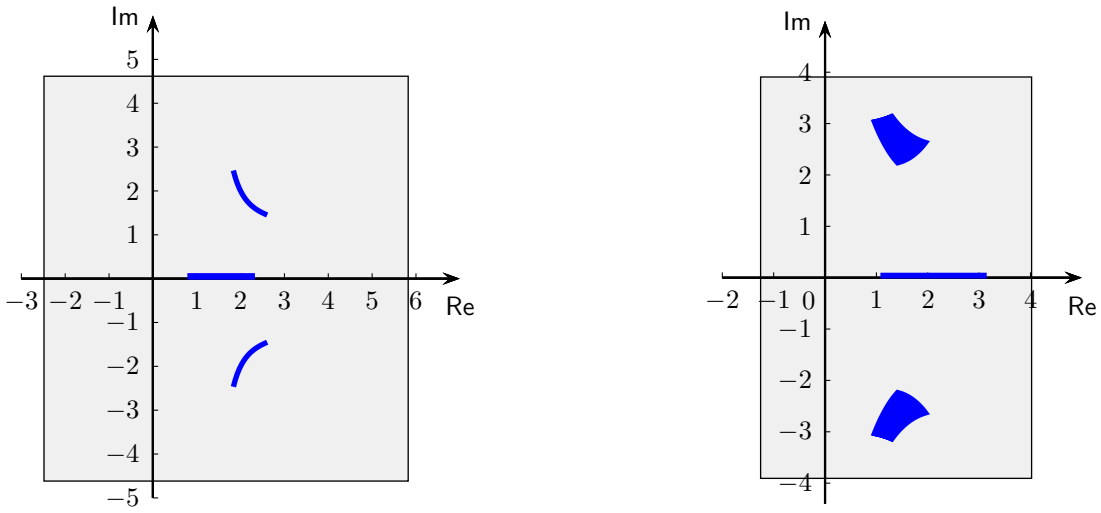
$$\begin{aligned}
&\lambda \leq \lambda_n(\tfrac{1}{2}(A^c + A^{cT})) - \rho(\tfrac{1}{2}(A^\Delta + A^{\Delta T})), \\
&\lambda \geq \lambda_1(\tfrac{1}{2}(A^c + A^{cT})) + \rho(\tfrac{1}{2}(A^\Delta + A^{\Delta T})), \\
&\mu \leq \lambda_n \begin{pmatrix} 0 & \frac{1}{2}(A^c - A^{cT}) \\ \frac{1}{2}(A^{cT} - A^c) & 0 \end{pmatrix} - \rho \begin{pmatrix} 0 & \frac{1}{2}(A^\Delta + A^{\Delta T}) \\ \frac{1}{2}(A^{\Delta T} + A^\Delta) & 0 \end{pmatrix}, \\
&\mu \geq \lambda_1 \begin{pmatrix} 0 & \frac{1}{2}(A^c - A^{cT}) \\ \frac{1}{2}(A^{cT} - A^c) & 0 \end{pmatrix} + \rho \begin{pmatrix} 0 & \frac{1}{2}(A^\Delta + A^{\Delta T}) \\ \frac{1}{2}(A^{\Delta T} + A^\Delta) & 0 \end{pmatrix}.
\end{aligned}$$

These are exactly the bounds developed by Rohn (1998a). Better bounds for the symmetric case would result in a tighter enclosure to the general case.

Notice that the bounds (5.1b) can be slightly sharpened by realizing the dependency in the interval matrix

$$\begin{pmatrix} 0 & \frac{1}{2}(\boldsymbol{A} - \boldsymbol{A}^T) \\ \frac{1}{2}(\boldsymbol{A}^T - \boldsymbol{A}) & 0 \end{pmatrix}.$$

Concretely, the diagonal of $\boldsymbol{A} - \boldsymbol{A}^T$ is overestimated due to the properties of interval arithmetic. However, for each $A \in \boldsymbol{A}$ the diagonal of $A - A^T$ is zero, so we can put to zero the diagonal of the resulting interval matrix $\boldsymbol{A} - \boldsymbol{A}^T$, too.



(a) Matrix $\boldsymbol{A}$. (b) Matrix $\boldsymbol{B}$.

Figure 5.2: (Example 5.9) The eigenvalue sets of interval matrices $\boldsymbol{A}$ and $\boldsymbol{B}$.

Formula (5.1b) can also be reformulated to the form similar to (5.1a), reducing the size of the matrix by the factor of 2

$$\underline{\lambda}_n\big(\tfrac{1}{2i}(\boldsymbol{A}-\boldsymbol{A}^T)\big) \le \mu \le \overline{\lambda}_1\big(\tfrac{1}{2i}(\boldsymbol{A}-\boldsymbol{A}^T)\big).$$

In this case, however, the matrix $\frac{i}{2}(\boldsymbol{A}^T-\boldsymbol{A})$ is handled as an interval Hermitian matrix, so the results for symmetric interval matrices have to be adapted accordingly.

Example 5.9. Let

$$\boldsymbol{A} = \begin{pmatrix} 3 & 1 & [2,3] \\ 5 & 2 & 3 \\ -5 & 1 & 1 \end{pmatrix}.$$

Figure 5.2a illustrates the eigenvalue set. One eigenvalue is always real and the other two are complex conjugates. The gray rectangle determines the initial bounds calculated by means of Theorem 5.8; the bounds on the eigenvalues of the symmetric interval matrices therein were computed by the basic enclosure of Theorem 5.3.

As the second example, consider matrix

$$\boldsymbol{B} = \begin{pmatrix} 1 & 2 & [-1,0] \\ -4 & 3 & [0,1] \\ 3 & 2 & 1 \end{pmatrix}.$$

The eigenvalue set is depicted in Figure 5.2b. □

5.1.3 Spectral radius

Let $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$. Our focus is on the range of spectral radii

$$\rho(\boldsymbol{A}) = \{\rho(A);\ A \in \boldsymbol{A}\},$$

which forms a compact real interval. Computational complexity of determining the left or the right endpoint of $\rho(\boldsymbol{A})$ is an open problem. Indeed, there are only a few of results known for $\rho(\boldsymbol{A})$.

Proposition 5.10. *If* $\underline{A} \ge 0$, *then* $\rho(\boldsymbol{A}) = [\rho(\underline{A}), \rho(\overline{A})]$.

Proof. Let $A \in \boldsymbol{A}$. By Theorem 1.8, $0 \le \underline{A} \le A \le \overline{A}$ implies $\rho(\underline{A}) \le \rho(A) \le \rho(\overline{A})$. By continuity of eigenvalues, all intermediate values are attained. In particular, the range $\rho(\boldsymbol{A})$ is obtained by the set of spectral radii of the matrices of the form $(1-\alpha)\underline{A} + \alpha\overline{A}$, $\alpha \in [0,1]$. □

Several bounds on the spectral radius of $\boldsymbol{A}$ were proposed in Hladík (2011b).

5.1.4 Eigenvectors

The surprising fact on eigenvectors of interval matrices is that while checking whether $\lambda \in \mathbb{R}$ is an eigenvalue of $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$ is NP-hard, checking whether $x \in \mathbb{R}^n$ is an eigenvector of $\boldsymbol{A}$ is effectively decidable. The explanation might be that for a given λ, knowledge of the corresponding n-dimensional eigenvector x is missing, while knowing x, only one-dimensional parameter λ must be found.

Theorem 5.11 (Rohn, 1993b)**.** *A vector $o \neq x \in \mathbb{R}^n$ is an eigenvector of some matrix in $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$ if and only if*

$$(\operatorname{diag}(z)A^c x - A^\Delta|x|)|x|^T \leq |x|(\operatorname{diag}(z)A^c x + A^\Delta|x|)^T, \tag{5.3}$$

where $z \coloneqq \operatorname{sgn}(x)$.

Proof. The vector x is an eigenvector corresponding to an eigenvalue $\lambda \in \mathbb{R}$ iff $(A - \lambda I_n)x = o$ for some $A \in \boldsymbol{A}$. By the Oettli–Prager Theorem 3.2, x and λ have to satisfy

$$|A^c x - \lambda x| \leq A^\Delta|x|,$$

or,

$$A^c x - A^\Delta|x| \leq \lambda x \leq A^c x + A^\Delta|x|.$$

Feasibility of this system is easily checked by linear programming. Since there is only one variable λ, we can eliminate it and obtain an explicit characterization. Denoting $\boldsymbol{y} \coloneqq [A^c x - A^\Delta|x|, A^c x + A^\Delta|x|]$, the ith inequality reads

$$y_i^c - y_i^\Delta \leq \lambda x_i \leq y_i^c + y_i^\Delta.$$

Supposing $x_i \neq 0$ (which case can be easily treated), we get

$$(y_i^c - z_i y_i^\Delta)/x_i \leq \lambda \leq (y_i^c + z_i y_i^\Delta)/x_i.$$

The Fourier–Motzkin elimination of λ yields inequalities by combining all left and right hand sides:

$$(y_i^c - z_i y_i^\Delta)/x_i \leq (y_j^c + z_i y_j^\Delta)/x_j, \quad \forall i,j.$$

By multiplying $x_i x_j$, we get

$$z_i(y_i^c - z_i y_i^\Delta)z_j x_j \leq z_j(y_j^c + z_i y_j^\Delta)z_i x_i, \quad \forall i,j.$$

A compact matrix form (slightly redundant) is

$$\operatorname{diag}(z)(A^c x - \operatorname{diag}(z)A^\Delta|x|)|x|^T \leq |x|\operatorname{diag}(z)(A^c x + \operatorname{diag}(z)A^\Delta|x|)^T,$$

which is equivalent to (5.3). □

In the symmetric case, the problem is also polynomially solvable by linear programming techniques, however, there is not an explicit formula such as (5.3).

Observation 5.12. *A vector $o \neq x \in \mathbb{R}^n$ is an eigenvector of some matrix in $\boldsymbol{A}^S$ if and only if the linear system*

$$Ax = \lambda x, \quad \underline{A} \leq A \leq \overline{A}, \quad A = A^T$$

is solvable w.r.t. variables λ and A.

Proof. A vector $o \neq x \in \mathbb{R}^n$ is an eigenvector of some matrix in $\boldsymbol{A}^S$ if and only if there is $A \in \boldsymbol{A}^S$ such that $Ax = \lambda x$, from which the final form follows. □

An adaptation for the general case of a complex eigenvector follows. Below, $\lambda + i\mu$ represents the eigenvalue associated to the eigenvector $x + iy$.

Observation 5.13. *A vector $o \neq x + iy \in \mathbb{C}^n$ is an eigenvector of some matrix in $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$ if and only if the linear system*

$$Ax = \lambda x - \mu y, \quad Ay = \lambda y + \mu x, \quad \underline{A} \leq A \leq \overline{A}$$

is solvable w.r.t. variables λ, μ and A.

Notes

One of the first results on computing the ranges of eigenvalues of interval matrices are due to Deif (1991a) and Rohn and Deif (1992). They derived formulae for determining exact bounds for the complex and the real eigenvalues, respectively. However, these results apply only under a restrictive assumption on the sign pattern invariancy of the corresponding eigenvectors. Other theoretical properties of real eigenvalues are in Rohn (1993b). Various methods for enclosing eigenvalues of $\boldsymbol{A}$ and/or $\boldsymbol{A}^S$ were studied in Hladík et al. (2010); Kolev (2006a). A filtering method for making the enclosures tighter was proposed in Hladík et al. (2011b). Kolev and Petrakieva (2005) estimated the real parts of the eigenvalues by solving a specific nonlinear system.

The generalized interval eigenvalue problem was dealt with in Leng and He (2010). The related problem of bounding singular values of an interval matrix was discussed, e.g., in Ahn and Chen (2007); Deif (1991b); Hladík et al. (2010).

Applications of the interval eigenvalue problem are found in diverse areas such as robotics (Chablat et al., 2004), mass structures (Qiu et al., 1996), automobile suspension system (Qiu et al., 2001), vibrating systems (Dimarogonas, 1995) or principal component analysis (Gioia and Lauro, 2006).

For more on Perron vectors of nonnegative interval matrices see, e.g., (Rohn, 2012b, Sec. 3.4.4).

There are many papers studying (Hurwitz or Schur) stability of interval matrices: an algorithm giving a sufficient condition of instability of a symmetric interval matrix (Rohn, 1996). Gerschgorin-like enclosures were investigated in Franzè et al. (2006). Mansour (1989) reviews diverse approaches to stability of interval matrices. Sufficient conditions by Lyapunov approach are in Petkovski (1991). More exhausting algorithms based on the branch & bound framework with some useful sufficient conditions were proposed in Ghosh et al. (2000); Wang et al. (1994). There are related problems of stability of linear time-invariant systems Ahn and Chen (2008) and iterative control systems Ahn et al. (2007). Stability checking by two-dimensional exposed faces was studied in Xiao and Unbehauen (2000).

The bounds for the symmetric case were utilized in Euclidean distance problems (Dias da Silva, 2017). They are often used in constructing convex lower bounds of nonconvex functions (Section 8.3.1).

5.2 Positive semidefiniteness

We say that a symmetric interval matrix $\boldsymbol{A}^S$ is *positive semidefinite* if every $A \in \boldsymbol{A}^S$ is positive semidefinite. The following theorem gives necessary and sufficient conditions for $\boldsymbol{A}^S$ to be positive semidefinite.

Theorem 5.14 (Rohn, 1994b)**.** *The following statements are equivalent*

(1) $\boldsymbol{A}^S$ *is positive semidefinite;*

(2) A_{zz} *is positive semidefinite for each* $z \in \{\pm 1\}^n$*;*

(3) $x^T A^c x - |x|^T A^\Delta |x| \geq 0$ *for each* $x \in \mathbb{R}^n$*.*

Proof.

"(1) $\Rightarrow$ (2)" Obvious from $A_{zz} \in \boldsymbol{A}^S$.

"(2) $\Rightarrow$ (3)" Let $x \in \mathbb{R}^n$ and put $z := \operatorname{sgn}(x)$. Now,

$$x^T A^c x - |x|^T A^\Delta |x| = x^T A^c x - x^T \operatorname{diag}(z) A^\Delta \operatorname{diag}(z) x = x^T A_{zz} x \geq 0.$$

"(3) $\Rightarrow$ (1)" Let $A \in \boldsymbol{A}^S$ and $x \in \mathbb{R}^n$. Now,

$$x^T A x = x^T A^c x + x^T (A - A^c) x \geq x^T A^c x - |x^T (A - A^c) x| \geq x^T A^c x - |x|^T A^\Delta |x| \geq 0. \qquad \square$$

Kreinovich (2005) showed that the finite reduction in Theorem 5.14 is optimal, that is, the set of matrices A_{zz}, $z \in \{\pm 1\}^n$, cannot be further reduced for checking positive semidefiniteness of $\boldsymbol{A}^S$. It was proved by Nemirovskii (1993) that checking positive semidefiniteness of $\boldsymbol{A}^S$ is *co-NP*-hard; cf. Kreinovich et al. (1998). Thus, any sufficient or necessary condition may be useful. The following is a direct consequence of Theorem 5.3 using the fact that $\boldsymbol{A}^S$ is positive semidefinite if and only if $\underline{\lambda}_n(\boldsymbol{A}^S) \geq 0$.

Proposition 5.15 (Rohn, 1994b)**.** $\boldsymbol{A}^S$ *is positive semidefinite if* $\lambda_n(A^c) \geq \rho(A^\Delta)$.

5.3 Positive definiteness

In analogy to positive semidefiniteness, $\boldsymbol{A}^S$ is called *positive definite* if every $A \in \boldsymbol{A}^S$ is positive definite. There are also similar characterizations and properties of positive definiteness, with an additional one.

Theorem 5.16 (Rohn, 1994b)**.** *The following are equivalent*

(1) $\boldsymbol{A}^S$ *is positive definite;*

(2) A_{zz} *is positive definite for each* $z \in \{\pm 1\}^n$*;*

(3) $x^T A^c x - |x|^T A^\Delta |x| > 0$ *for each* $o \neq x \in \mathbb{R}^n$*;*

(4) A^c *is positive definite and* $\boldsymbol{A}$ *is regular.*

Proof.
"(1) $\Leftrightarrow$ (2) $\Leftrightarrow$ (3)" is similar to the proof of Theorem 5.14.
"(1) $\Rightarrow$ (4)" Suppose to the contrary that $\boldsymbol{A}$ is not regular, that is, there are $A \in \boldsymbol{A}$ and $x \neq o$ such that $Ax = o$. This implies $0 = x^T Ax = x^T \frac{1}{2}(A + A^T)x$, and so $\frac{1}{2}(A + A^T) \in \boldsymbol{A}^S$ is not positive definite.
"(4) $\Rightarrow$ (1)" Positive definiteness of A^c implies that for each $i = 1, \ldots, n$ the interval $\boldsymbol{\lambda}_i(\boldsymbol{A}^S)$ contains a positive value, and regularity of $\boldsymbol{A}$ implies that it does not contain zeros, so it must be a positive interval. □

Intractability of checking positive definiteness is a consequence of intractability of checking regularity.

Theorem 5.17 (Rohn, 1994a)**.** *Checking positive definiteness of* $\boldsymbol{A}^S$ *is a co-NP-hard problem on the class of interval matrices with* A^c *non-negative positive definite rational and* $A^\Delta = ee^T$*.*

Proof. By Theorem 3.60, checking regularity of $\boldsymbol{A}$ is *co-NP*-hard even on the class of interval matrices with A^c non-negative positive definite rational and $A^\Delta = ee^T$. By Theorem 5.16, such an interval matrix is regular if and only if $\boldsymbol{A}^S$ is positive definite. □

An analogy of Proposition 5.15 follows.

Proposition 5.18 (Rohn, 1994b)**.** $\boldsymbol{A}^S$ *is positive definite if* $\lambda_n(A^c) > \rho(A^\Delta)$*.*

Remark 5.19. We say that $\boldsymbol{A}^S$ is *weakly positive definite* if there is at least one positive definite matrix in $\boldsymbol{A}^S$. In contrast to checking positive definiteness, checking the weak version is a polynomial time problem (Jaulin and Henrion, 2005). Similarly as in Remark 5.7, we utilize a semidefinite programming problem

$$\max\ \alpha \ \text{ subject to } \ A - \alpha I_n \text{ is positive semidefinite, } \ A \in \boldsymbol{A}^S.$$

Then $\boldsymbol{A}^S$ is weakly positive definite if and only if the optimal value is positive.

For weak positive semidefiniteness, this is an open problem as semidefinite programming problems can be solved efficiently but with an (arbitrarily small) precision. For weak positive definiteness, this does not matter, since if $\boldsymbol{A}^S$ includes a positive definite matrix, then there is one with a polynomially large gap between the smallest eigenvalue and zero. This is easy to see since if $\boldsymbol{A}^S$ is weakly positive definite, then for every $A \in \boldsymbol{A}^S$ we have $\lambda_n(A) = \lambda_1^{-1}(A^{-1}) \geq \|A^{-1}\|^{-1}$. The value $\|A^{-1}\|^{-1}$ is efficiently computable for most of the matrix norms, and also the limits of the entries of the inverses $\{A^{-1};\ A \in \boldsymbol{A}\}$ have polynomial size from the properties of interval linear systems (cf. the proof of Theorem 3.5 or Theorem 3.52).

5.3.1 Application: Convexity testing

It is well known (Bazaraa et al., 2006; Boyd and Vandenberghe, 2004; Luenberger and Ye, 2008) that a function $f\colon \mathbb{R}^n \to \mathbb{R}$ is convex on a nondegenerate $\boldsymbol{x} \in \mathbb{IR}^n$ if and only if its Hessian $\nabla^2 f(x)$ is positive semidefinite $\forall x \in \operatorname{int} \boldsymbol{x}$. Thus, we have the following sufficient condition for checking convexity of a function on a box.

Proposition 5.20. *Let* $f\colon \mathbb{R}^n \to \mathbb{R}$*,* $\boldsymbol{x} \in \mathbb{IR}^n$*, and* $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$ *such that* $\nabla^2 f(\boldsymbol{x}) \subseteq \boldsymbol{A}^S$*. Then* $f(x)$ *is convex on* $\boldsymbol{x} \in \mathbb{IR}^n$ *if* $\boldsymbol{A}^S$ *is positive semidefinite.*

Example 5.21. Let

$$f(x,y,z) = x^3 + 2x^2y - xyz + 3yz^2 + 8y^2,$$

where $x \in \boldsymbol{x} = [2,3]$, $y \in \boldsymbol{y} = [1,2]$ and $z \in \boldsymbol{z} = [0,1]$. The Hessian of $f(x)$ reads

$$\nabla^2 f(x,y,z) = \begin{pmatrix} 6x+4y & 4x-z & -y \\ 4x-z & 16 & -x+6z \\ -y & -x+6z & 6y \end{pmatrix}.$$

The evaluation of the Hessian matrix by interval arithmetic gives

$$\nabla^2 f(\boldsymbol{x},\boldsymbol{y},\boldsymbol{z}) \subseteq \boldsymbol{A} = \begin{pmatrix} [16,26] & [7,12] & -[1,2] \\ [7,12] & 16 & [-3,4] \\ -[1,2] & [-3,4] & [6,12] \end{pmatrix}.$$

By Proposition 5.15 or 5.16, we check $\lambda_n(A^c) = 7.5 > 7.1218 = \rho(A^\Delta)$, from which positive definiteness of $\boldsymbol{A}^S$ follows. Thus, we can conclude that $f(x)$ is convex on the interval domain.

It is interesting to check the minimal coefficient of y^2 for which the interval Hessian is still positive semidefinite. Proposition 5.15 has the threshold of about 7.49156. Beeck's sufficient condition for checking regularity of $\boldsymbol{A}$ has the threshold of about 7.13044. Surprisingly, condition (2) of Theorem 5.16 gives the same boundary point. The reason is, however, simple as Beeck's condition is both sufficient and necessary in this case by Corollary 3.73. □

In general, we cannot hope for a stronger test for convexity checking since the problem is *co-NP*-hard, even when restricted to polynomials of degree three.

Proposition 5.22. *Checking convexity of a polynomial of degree three in n variables and on an interval domain is a co-NP-hard problem.*

Proof. Consider the polynomial in variables $x \in \mathbb{R}^n$ and $y \in \mathbb{R}$

$$p(x,y) = x^T A x + dy^2 + x^T xy,$$

where $A \in \mathbb{R}^{n\times n}$ is positive definite and $d \in \mathbb{R}$. Let $\boldsymbol{x} \in \mathbb{IR}^n$, $\boldsymbol{y} = [0,1]$ be a domain, where we want to test convexity of $p(x,y)$. The Hessian of $p(x,y)$ reads

$$\nabla^2 p(x,y) = 2\begin{pmatrix} A + yI_n & x \\ x^T & d \end{pmatrix}.$$

It is sufficient to check positive definiteness for $y = 0$. Then the Hessian matrix is positive definite if and only if $d - x^T A^{-1} x > 0$. In order to check positive definiteness for each $x \in \boldsymbol{x}$, we have to compute $\max_{x\in\boldsymbol{x}} x^T A^{-1} x$ and compare with d. However, computing the maximum of a convex quadratic form on an interval box is an *NP*-hard problem (Vavasis, 1991). More precisely, it is *NP*-hard to decide whether the maximum value is at least d. □

Notes and further reading. Positive definiteness and regularity of interval matrices play also an important role in testing a generalized concept of convexity, called pseudoconvexity; see Hladík (2018a) for standard interval matrices and Hladík et al. (2021) for linear interval parametric matrices. Positive definiteness of parametric interval matrices was also considered in Fazzolari and Ferreira (2021) in the context of solving interval Lyapunov equations.

5.4 Further topics

distance of interval vectors Mayer (1970).

5.4.1 Linear independence

Linear independence of interval vectors will be presented in terms of a full column rank of an interval matrix.

Definition 5.23. An interval matrix $\boldsymbol{A} \in \mathbb{IR}^{n\times k}$ has the *full column rank* if $\operatorname{rank}(A) = k$ for each $A \in \boldsymbol{A}$.

Checking full column rank is straightforwardly reduced to solving interval linear equations as $\boldsymbol{A}$ has a full rank if and only if $\boldsymbol{A}x = 0$ has no nontrivial solution. The equivalent linear formulation avoiding nontriviality of x states that $\boldsymbol{A}x = 0$, $[-e, e]^T x = 1$ has no solution.

Sufficient conditions on the full column rank are obtained by a reduction to the square case and utilizing regularity of an interval matrix.

Proposition 5.24. *An interval matrix $\boldsymbol{A} \in \mathbb{IR}^{n\times k}$ has a full column rank if $(A^c)^\dagger \boldsymbol{A}$ is regular.*

Proof. Multiply $\boldsymbol{A}$ from the left by $(A^c)^\dagger$, which preserves the rank of every every $A \in \boldsymbol{A}$. □

For instance, applying the Beeck's sufficient regularity condition (Theorem 3.65), we get a simple sufficient condition from Rohn (2012a).

Corollary 5.25. *An interval matrix $\boldsymbol{A} \in \mathbb{IR}^{n\times k}$ has a full column rank if* $\operatorname{rank}(A^c) = k$ *and $\rho(|(A^c)^\dagger|A^\Delta) < 1$.*

Proof. The midpoint of $(A^c)^\dagger \boldsymbol{A}$ is $(A^c)^\dagger A^c = I_k$, and the radius is $|(A^c)^\dagger|A^\Delta$. □

Another sufficient condition is via regularity of a submatrix.

Proposition 5.26. *An interval matrix $\boldsymbol{A} \in \mathbb{IR}^{n\times k}$ has a full column rank if it contains a regular submatrix of size k, but not conversely.*

The proof is easy. The surprising point here is that the implication does not hold conversely (in contrast to real matrices). A counterexample is given below.

Example 5.27. Consider an interval matrix by Irene Sharaya (see Shary, 2014)

$$\begin{pmatrix} 1 & [0,1] \\ -1 & [0,1] \\ [-1,1] & 1 \end{pmatrix}.$$

By Corollary 5.25, it has rank 2. Nevertheless, it contains no regular submatrix of size 2. □

Notice, however, that extending the above interval matrix to a square matrix by including the column $(1,1,0)^T$ produces a regular interval matrix.

5.4.2 Rank

In view of intractability of checking regularity of an interval matrix (Theorem 3.60) we have that certain problems related to the rank of an interval matrix will be NP-hard as well. Nevertheless, there are also some easy tasks.

Proposition 5.28 (). *It is a polynomial problem to check if an interval matrix $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ contains a rank one matrix (and find it if it exists).*

Proof. We want to find $A \in \boldsymbol{A}$ of rank one, that is, in the form $A = xy^T$ for some $x, y \in \mathbb{R}^n$. If there is j such that $0 \in \boldsymbol{a}_{ij}$ for every i, then we can put $y_j = 0$. So we can assume that this is not the case and $y_j \neq 0$ for every j. Now, we can formulate the problem as the feasibility problem

$$\underline{a}_{ij} \leq x_i y_j \leq \overline{a}_{ij}, \quad \forall i, j,$$

or, equivalently

$$\underline{a}_{ij} \frac{1}{y_j} \leq x_i \leq \overline{a}_{ij} \frac{1}{y_j}, \quad \forall i, j.$$

Substituting $y'_j := \frac{1}{y_j}$, we can state the problem as a linear program

$$\underline{a}_{ij} y'_j \leq x_i \leq \overline{a}_{ij} y'_j, \quad \forall i, j.$$

□

5.4.3 Determinants

Given $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$, the task is to determine the image of the determinant, that is, $\det(\boldsymbol{A}) = \{\det(A);\ A \in \boldsymbol{A}\}$. Since the determinant is a continuous functions w.r.t. matrix entries, the image $\det(\boldsymbol{A})$ is always a compact interval.

Proposition 5.29 (Rohn, 1985)**.** *The endpoints of* $\det(\boldsymbol{A})$ *are attained for matrices* $A \in \boldsymbol{A}$ *satisfying* $\boldsymbol{a}_{ij} \in \{\underline{a}_{ij}, \overline{a}_{ij}\}$ *for all* i, j*.*

Proof. Since the determinant is a linear function in each entry, we can fix any entry either at the left or the right endpoint. □

This characterization requires to compute the determinants of 2^{n^2} real matrices, so it is not useful for practical testing. An efficient method, however, hardly exists due to intractability (Kreinovich et al., 1998).

Proposition 5.30. *Computing the range* $\det(\boldsymbol{A})$ *is an NP-hard problem. In particular, checking* $0 \in \det(\boldsymbol{A})$ *is NP-hard on the class of interval matrices* $\boldsymbol{A}$ *with* A^c *non-negative positive definite rational and* $A^\Delta = ee^T$*.*

Proof. By Theorem 3.60, checking regularity of $\boldsymbol{A}$ is *co-NP*-hard. Now, we use the fact that $\boldsymbol{A}$ is not regular if and only if $0 \in \det(\boldsymbol{A})$. □

Obviously, it is also *NP*-hard to compute just one of the endpoints $\underline{\det}(\boldsymbol{A})$ or $\overline{\det}(\boldsymbol{A})$. This is because once we could efficiently determine $\underline{\det}(\boldsymbol{A})$, then we can efficiently calculate $\overline{\det}(\boldsymbol{A})$ as $\overline{\det}(\boldsymbol{A}) = -\overline{\det}(\boldsymbol{A}')$, where $\boldsymbol{A}'$ originates from $\boldsymbol{A}$ by multiplying the first row by -1.

For a symmetric interval matrix $\boldsymbol{A}^S$, computing the range $\det(\boldsymbol{A}^S)$ is an *NP*-hard problem, too (Hladík, 2018a).

Proposition 5.29 concerns the endpoints of $\det(\boldsymbol{A})$. The intermediate values in $\det(\boldsymbol{A})$ are also attained for matrices of a special form.

Theorem 5.31 (Rohn, 1985)**.** *Let* $A \in \boldsymbol{A}$*. Then there is* $A' \in \boldsymbol{A}$ *such that* $\det(A) = \det(A)'$ *and for at most one entry we have* $a'_{ij} \notin \{\underline{a}_{ij}, \overline{a}_{ij}\}$*.*

Proof. Let $A^1, A^2 \in \boldsymbol{A}$ be the vertex matrices, for which $\underline{\det}(\boldsymbol{A})$ and $\overline{\det}(\boldsymbol{A})$ are attained, respectively. We take the entries in which the matrices differ and one-by-one continuously move A^1_{ij} to A^2_{ij}. Since the determinant is a continuous function, we eventually obtain the value of $\det(A)$ for some intermediate value between A^1_{ij} and A^2_{ij}. The other entries are at endpoints. □

Corollary 5.32 (Rohn, 1985)**.** *Let* $A \in \boldsymbol{A}$ *be irregular. Then there exists a singular* $A \in \boldsymbol{A}$ *such that for at most one entry we have* $a_{ij} \notin \{\underline{a}_{ij}, \overline{a}_{ij}\}$*.*

Enclosures. When computing an enclosure to $\det(\boldsymbol{A})$, one may follow Smith (1969) and employ the LU factorization $\boldsymbol{A} \subseteq \boldsymbol{LU}$; see Section 3.3.1 on Interval Gaussian elimination. As the diagonal of $\boldsymbol{L}$ consists of ones, we have $\det(\boldsymbol{A}) \subseteq \prod_{i=1}^n \boldsymbol{u}_{ii}$.

According to the numerical experiments carried out by Matějka (2017), it is convenient to precondition $\boldsymbol{A}$ by the numerically computed $C \approx (A^c)^{-1}$, and they using Gaussian elimination or Cramer's rule (Theorem 1.1). To compensate the preconditioning, we have to multiply the result by $\det(C)$.

The Cramer's rule method works here as follows. Suppose $\boldsymbol{A}$ is regular, and consider the interval linear system $\boldsymbol{A}x = e_1$. Let $\boldsymbol{x}$ be an enclosure to the solution set, and let $\boldsymbol{B}$ be the interval matrix $\boldsymbol{A}$ after removing the first row and column. By Cramer's rule for the first entry,

$$\det(\boldsymbol{A}) \subseteq \frac{\det(\boldsymbol{B})}{\boldsymbol{x}_1}.$$

This gives us an enclosing recursive formula for $\det(\boldsymbol{A})$, where $\det(\boldsymbol{B})$ is calculated recursively.

Determinant of a matrix is equal to the product of their eigenvalues. This observation can be also used for bounding the range of the determinant. By Matějka (2017), this approach is competitive with the above mentioned ones, particularly for symmetric interval matrices. Let $\boldsymbol{A}^S$ be a symmetric interval matrix and $\boldsymbol{\lambda}_1, \ldots, \boldsymbol{\lambda}_n$ enclosures for its eigenvalue sets (cf Section 5.1). Then $\det(\boldsymbol{A}) \subseteq \prod_{i=1}^n \boldsymbol{\lambda}_i$.

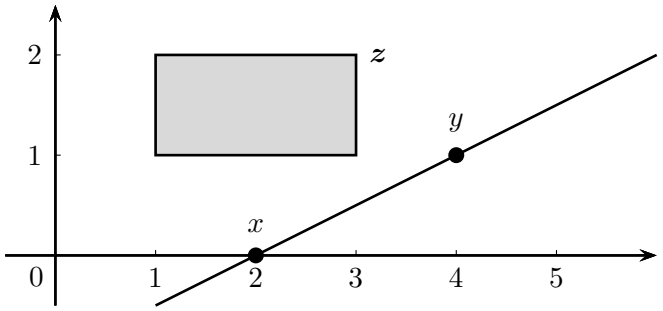


Figure 5.3: (Example 5.33) Intersection test of a line and a rectangle.

Applications. Determinants often appear in geometric computation. Therefore determinants of interval matrices are useful for numerically rigorous handling of geometrical objects and some positioning test; cf. Section 10.5.2 and Ratschek and Rokne (2003). The positioning tests include, for instance, the position and intersection of a box and a line, a plane, a circle or a sphere.

Example 5.33 (Ratschek and Rokne (2003))**.** Let three points $x, y, z \in \mathbb{R}^2$ in the plane be given. It is well known that the absolute value of the determinant

$$d(x,y,z) = \frac{1}{2}\begin{vmatrix} x_1 & x_2 & 1 \\ y_1 & y_2 & 1 \\ z_1 & z_2 & 1 \end{vmatrix}$$

gives the area of the triangle x, y, z. Thus, the points are collinear if and only if the determinant vanishes. Moreover, the sign of $d(x, y, z)$ coincides with the orientation of the three points. Now, suppose that we have a rectangle $\boldsymbol{z} \in \mathbb{IR}^2$ and the question is whether the line passing through x, y intersects the rectangle or not. From the above observation, it is clear that the line intersects the rectangle if and only if $0 \in d(x, y, \boldsymbol{z})$. Thus, it is sufficient to evaluate the determinant of the interval matrix. Luckily, it can be done exactly in this case since

$$2d(x,y,\boldsymbol{z}) = \begin{vmatrix} x_1 & x_2 & 1 \\ y_1 & y_2 & 1 \\ \boldsymbol{z}_1 & \boldsymbol{z}_2 & 1 \end{vmatrix} = \boldsymbol{z}_1(x_2 - y_2) - \boldsymbol{z}_2(x_1 - y_1) + x_1y_2 - x_2y_1.$$

As a concrete example, let $x = (2, 0)^T$, $y = (4, 1)^T$ and $\boldsymbol{z} = ([1, 3], [1, 2])^T$ as depicted in Figure 5.3. The line does not intersect the rectangle, which coincides with the value of the determinant

$$2d(x,y,\boldsymbol{z}) = [1,3](0-1) - [1,2] * (2-4) + 2 - 0 = [1,5]. \qquad \square$$

Remark 5.34. Determinants of interval matrices found applications in robotics. In the design of a parallel manipulator, an important step is to detect singularities. Singularities are configurations that may cause deadlocks or infinite forces, leading to a breakdown. A natural way to detect singularities is as such configurations for which the determinant of the inverse Jacobian matrix vanishes. Considering all configurations in the workspace, we end up with a problem of computing the determinant of an interval matrix Merlet (1998, 2007); Merlet and Donelan (2006); Oetomo et al. (2009). The interval matrices are often affected by nonlinear dependencies.

Notes and further reading. Determinants of interval matrices were also utilized in testing pseudoconvexity (a generalized convexity) of a function on an interval domain (Hladík, 2018a).

5.4.4 Inverse interval matrices

Numerical analysts recommend to avoid the computation of the matrix inverse. However, there are situations, where it is necessary; see (Higham, 1996, pp. 262) for references. Even in the interval environment,

we need sometimes to compute the inverse of an interval matrix. For example, some methods for solving an interval system of linear equations need an enclosure of the interval matrix inverse (Kolev, 2016; Skalna and Hladík, 2019).

Let $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$ be regular. Since the set of all inverses of the matrices in $\boldsymbol{A}$ does not form an interval matrix, we define the *inverse interval matrix* of $\boldsymbol{A}$ as its interval hull, that is,

$$\boldsymbol{A}^{-1} := \square\{A^{-1};\, A \in \boldsymbol{A}\}.$$

The i-th column of $\boldsymbol{A}^{-1}$ is the interval hull of the solution set of the interval linear system $\boldsymbol{A}x = e_i$, so the methods from Section 3.2 are directly applicable to compute $\boldsymbol{A}^{-1}$ column by column. Note that computing $\boldsymbol{A}^{-1}$ is NP-hard on a class of strongly regular interval matrices Coxson (1999).

More compact forms and specific results are as follows. A simple initial enclosure to $\boldsymbol{A}^{-1}$ is found in (Neumaier, 1990, Thm. 4.1.11); see also references on page 166 therein. A method to tightly enclose $\boldsymbol{A}^{-1}$ is proposed in (Rohn, 1989a, Thm. 2.40), and exact calculations of $\boldsymbol{A}^{-1}$ are discussed in Rohn (1989a, 2012a). Inverse stability and its effect on determining $\boldsymbol{A}^{-1}$ is dealt with in Rohn (1993c), as well as an inversion of $\boldsymbol{A}$ with A^{Δ} having rank one. A particular case with $A^c = I_n$, which can be also used to enclose $\boldsymbol{A}^{-1}$ in the general case after preconditioning, is considered in Rohn (1993a, 2011). An easy case when $\boldsymbol{A}$ is an interval M-matrix was dealt with in Section 3.5.5; an economic application of interval M-matrices and their inverses for determining the so-called *multipliers* was presented by Jerrell (1996). Finally, Rohn and Farhadsefat (2011) survey recent results.

5.4.5 P-matrices

An interval matrix $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$ is a *P-matrix* if every $A \in \boldsymbol{A}$ is a P-matrix. The characterization below is by Białas and Garloff (1984); cf. Kreinovich et al. (1998); Rohn and Rex (1996).

Theorem 5.35. *An interval matrix $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$ is a* P-matrix *if and only if A_{yy} is a P-matrix for every $y \in \{\pm 1\}^n$.*

By Theorem 5.16 we immediately have the following: As long as $\underline{A}$ and $\overline{A}$ are symmetric, $\boldsymbol{A}$ is a P-matrix if and only if $\boldsymbol{A}^S$ is positive definite.

5.4.6 Others

Powers of interval matrices.

Exponential of interval matrices. Let $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$, and the aim is to compute as tight as possible enclosure for the range of exponentials $\exp(\boldsymbol{A})$. Goldsztejn and Neumaier (2014) showed that a tight enclosure is *NP*-hard to compute and they also suggested some methods. In particular, scaling and squaring, together with an appropriate bound for a remainder, was a promising approach.

Chapter 6

Range of a function

In Section 2.4, we stated the fundamental problem of interval analysis—evaluation of a function over intervals. We introduced the basic method, the natural interval extension, to enclose the image of a function. In this chapter, we present other methods and discuss the related topics.

Let $f\colon \mathbb{R}^n \to \mathbb{R}$ be a real function and $\boldsymbol{x} \in \mathbb{IR}^n$. Our aim is to enclose the image $f(\boldsymbol{x}) = \{f(x);\ x \in \boldsymbol{x}\}$ as tightly as possible. Nevertheless, there are complexity issues; we mention some of them. The problem of determining the range of a function is *NP*-hard, even for polynomial functions; this was proved already in Gaganov (1981, 1985). Moreover, an ε-approximation of the range is *NP*-hard to calculate, too. As was shown by Kreinovich et al. (2003), checking whether a natural interval extension produces exact bounds for given interval domains and a function f is *NP*-hard as well. For general functions (it suffices to involve trigonometric functions), the problem becomes undecidable; see page 144 in Chapter 8.

Monotonicity test

Detecting monotonicity of $f(x)$ with respect to any variable is an important issue that often notably reduces overestimations done by particular enclosing methods. Suppose that $f(x)$ is non-decreasing with respect to x_i. The largest value of $f(x)$ over $\boldsymbol{x}$ is attained for $x_i = \overline{x}_i$ and the smallest one for $x_i = \underline{x}_i$. Thus, we can fixed the ith variable and solve two sub-problems with one less dimensions. Put $\boldsymbol{x}^1 := (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{i-1}, \underline{x}_i, \boldsymbol{x}_{i+1}, \ldots, \boldsymbol{x}_n)^T$ and $\boldsymbol{x}^2 := (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{i-1}, \overline{x}_i, \boldsymbol{x}_{i+1}, \ldots, \boldsymbol{x}_n)^T$, calculate enclosures $f(\boldsymbol{x}^1) \subseteq \boldsymbol{y}^1$ and $f(\boldsymbol{x}^2) \subseteq \boldsymbol{y}^2$ by any (inclusion monotonic) method and return the enclosure $\boldsymbol{y} := [\underline{y}^1, \overline{y}^2] \supseteq f(\boldsymbol{x})$.

Monotonicity of $f(x)$ with respect to x_i can be checked by calculating an enclosure $\boldsymbol{f}'_{x_i}(\boldsymbol{x})$ to the derivative $f'_{x_i}(x)$ over $\boldsymbol{x}$. If $\boldsymbol{f}'_{x_i}(\boldsymbol{x})$ lies in $[0, \infty)$, then $f(x)$ is nondecreasing on $\boldsymbol{x}$, and similarly if $\frac{\partial \boldsymbol{f}}{\partial x_i}(\boldsymbol{x})$ lies in $(-\infty, 0]$, then $f(x)$ is nonincreasing on $\boldsymbol{x}$. An enclosure of the derivative is calculated, e.g., by an interval extension of symbolic or automatic differentiation; see section 6.6.

Checking monotonicity by using an endpoint analysis was presented by Hansen (1997).

Endpoint analysis

The endpoint analysis serves not to calculate an enclosure to $f(\boldsymbol{x})$ but to check if the enclosure $\boldsymbol{f}(\boldsymbol{x})$ computed by a natural interval extension is optimal, that is, $\boldsymbol{f}(\boldsymbol{x}) = f(\boldsymbol{x})$ (Hansen, 1997; Hansen and Walster, 2004).

Example 6.1. Let $\boldsymbol{x} \in \mathbb{IR}$. The interval $\boldsymbol{x} + \boldsymbol{x} = [2\underline{x}, 2\overline{x}]$ calculated in interval arithmetic is optimal despite the dependency. However, $\boldsymbol{x} - \boldsymbol{x} = [\underline{x} - \overline{x}, \overline{x} - \underline{x}]$ is not optimal in general. The reason is that the lower (and upper) bound of the limit depends on both $\underline{x}$ and $\overline{x}$. □

Example 6.2. Let $f(x) = x(x-3)$. For $\boldsymbol{x} = [11, 12]$ we have $\underline{f}(\boldsymbol{x}) = \underline{x}(\underline{x} - 3)$. For $\boldsymbol{x} = [2, 5]$ we have $\underline{f}(\boldsymbol{x}) = \overline{x}(\underline{x} - 3)$. Thus, the resulting bounds depend on the concrete data. □

Theorem 6.3. *If* $\inf(\boldsymbol{f}(\boldsymbol{x}))$ *is calculated in interval arithmetic by using at most one limit from each of the intervals* $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$*, then* $\inf(\boldsymbol{f}(\boldsymbol{x})) = \inf(f(\boldsymbol{x}))$*. Analogously for the upper bound.*

Proof. Take $x \in \mathbb{R}^n$ such that x_i is the value of $\boldsymbol{x}_i$ that was used to calculate $\inf(\boldsymbol{f}(\boldsymbol{x}))$; if $\boldsymbol{x}_i$ was not used at all, then take any $x_i \in \boldsymbol{x}_i$. Now, $\inf(\boldsymbol{f}(\boldsymbol{x})) = f(x)$, so the lower bound $\inf(\boldsymbol{f}(\boldsymbol{x}))$ is attained at x and hence is exact. □

Refinement

Let $\boldsymbol{f}$ be an inclusion isotonic interval extension of f. The basic enclosure $\boldsymbol{f}(\boldsymbol{x})$ of $f(\boldsymbol{x})$ can be improved by a splitting of $\boldsymbol{x}$ into several sub-boxes and evaluation on them (Skelboe, 1974). Let us split $\boldsymbol{x}$ into k sub-boxes $\boldsymbol{x}^1, \ldots, \boldsymbol{x}^k$ so that $\boldsymbol{x} = \cup_{i=1}^k \boldsymbol{x}^i$. Due to inclusion isotonicity of $\boldsymbol{f}$,

$$f(\boldsymbol{x}) \subseteq \bigcup_{i=1}^{k} \boldsymbol{f}(\boldsymbol{x}^i) \subseteq \boldsymbol{f}(\boldsymbol{x}).$$

Thus, we obtain an enclosure that is never worse than $\boldsymbol{f}(\boldsymbol{x})$. The larger k, the tighter enclosure, but the larger k also needs more computations. This is a trade-off that is user or application dependent. The order of approximation and convergence to the image $f(\boldsymbol{x})$ is discussed, e.g., by Mayer (2017).

Overestimation

Let us introduce the following property of a function $f\colon \mathbb{R}^n \to \mathbb{R}$ and a given box $\boldsymbol{u} \in \mathbb{IR}^n$:

> **Property L.** There is given an expression for $f(x)$. For each variable x_i in the expression, let us replace all appearances of x_i by mutually different variables with the same domain $\boldsymbol{u}_i$. Then Property L is valid if the new expression defines a Lipschitz function on the corresponding domain.

This property is satisfied, for example, for real arithmetical operations $+, -, \cdot$ and / (division not by zero). Below is an example of a Lipschitz function, which gives rise to a non-Lipschitz one.

Example 6.4. Consider $f(x) = x^{1/3} - x^{1/3}$ and $x \in \boldsymbol{x} = [0, 1]$. Then $f(x)$ is Lipschitz on $\boldsymbol{x}$. After distinguishing particular occurrences of the variable x, we obtain the function $g(x, y) = x^{1/3} - y^{1/3}$, which is not Lipschitz on $[0, 1]^2$. □

Theorem 6.5. *Let $f\colon \mathbb{R}^n \to \mathbb{R}$ with $\boldsymbol{u} \in \mathbb{IR}^n$ satisfy Property L, and let $\boldsymbol{f}(\boldsymbol{x})$ be the natural extension of $f(x)$. Then there is $\gamma > 0$ such that for each $\boldsymbol{x} \in \mathbb{IR}^n$, $\boldsymbol{x} \subseteq \boldsymbol{u}$, we have*

$$\operatorname{rad}(\boldsymbol{f}(\boldsymbol{x})) - \operatorname{rad}(f(\boldsymbol{x})) \leq \gamma \|x^{\Delta}\|.$$

Proof. Let $g(z)$ be the function made from $f(x)$ when all appearances of all variables are renamed by $z_1, \ldots, z_m$. The interval vector $\boldsymbol{z}$ is created from the corresponding interval domains of variables. Then we have $\boldsymbol{f}(\boldsymbol{x}) = g(\boldsymbol{z})$. For any $z \in \boldsymbol{z}$ we can change the entries such that those corresponding to the same variable x_i are equal. This change affects the function value, but due to Lipschitz property no more than some constant multiple of x_i^{Δ}, which in total is $2\gamma\|x^{\Delta}\|$ for some constant γ. Thus, any value in $g(\boldsymbol{z})$ is not far from a value in $f(\boldsymbol{x})$ than $2\gamma\|x^{\Delta}\|$. □

The above theorem quantifies overestimation in terms of an absolute deviation. A relative overestimation can hardly be bounded as the following example illustrates. Consider $f(x) = x - x$ and $x \in \boldsymbol{x} = [0, 1]$. Then $f(\boldsymbol{x}) = 0$, but $\boldsymbol{f}(\boldsymbol{x}) = [-1, 1]$. So, $\operatorname{rad}(\boldsymbol{f}(\boldsymbol{x})) = 2$ is infinitely larger than $\operatorname{rad}(f(\boldsymbol{x})) = 0$.

As a consequence of Theorem 6.5, we have a linear convergence not only of the overestimation, but also of the enclosure itself:

Theorem 6.6. *Let $f\colon \mathbb{R}^n \to \mathbb{R}$ with $\boldsymbol{u} \in \mathbb{IR}^n$ satisfy Property L, and let $\boldsymbol{f}(\boldsymbol{x})$ be the natural extension of $f(x)$. Then there is $\gamma > 0$ such that for each $\boldsymbol{x} \in \mathbb{IR}^n$, $\boldsymbol{x} \subseteq \boldsymbol{u}$, we have*

$$\operatorname{rad}(\boldsymbol{f}(\boldsymbol{x})) \leq \gamma \|x^{\Delta}\|.$$

Proof. Since $f(x)$ satisfies Property L, it must be also Lipschitz. That is, there is a constant $c > 0$ such that $|f(x) - f(y)| \leq c\|x - y\|$ for all $x, y \in \boldsymbol{x}$. Thus, $\operatorname{rad}(f(\boldsymbol{x})) \leq c\|x^{\Delta}\|$. Since $\boldsymbol{f}(\boldsymbol{x})$ overestimates the true image $f(\boldsymbol{x})$ also linearly (Theorem 6.5), the statement follows. □

For more results on overestimation see (Alefeld and Herzberger, 1983, Chap. 3).

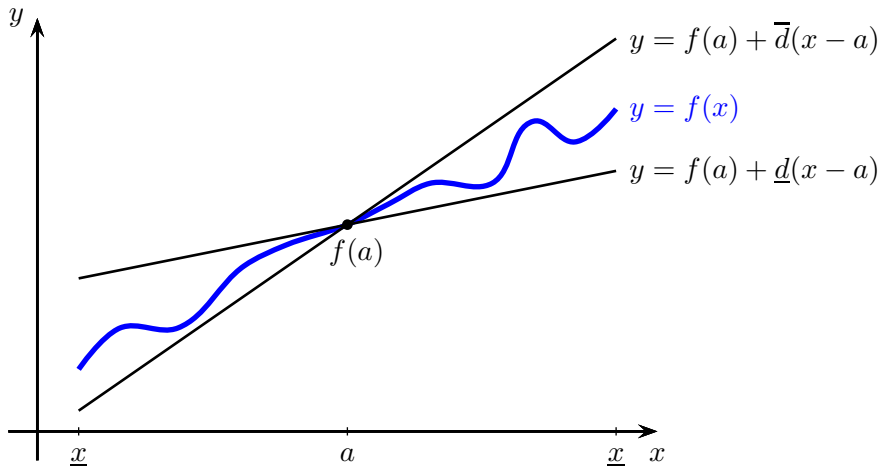


Figure 6.1: Enclosing a function $f(x)$ by an interval linear function by means of (6.4).

6.1 Mean value forms

Apart from the natural interval extension, the mean value form (and slopes) is another commonly used method to enclose the range of functions. It was originally proposed by Moore (1966).

While the natural interval extension converges with linear order (Theorem 6.5), the mean value form enclosure converges with order two (Theorem 6.12). Thus, roughly speaking, the former is more suitable for wide intervals, while the latter for more narrow intervals.

6.1.1 One-dimensional case

Let $f\colon \mathbb{R} \to \mathbb{R}$ be differentiable on an open superset of $\boldsymbol{x} \in \mathbb{IR}$. Let $a \in \boldsymbol{x}$; usually we set $a \coloneqq x^c$, but other selections may be used as well (and sometimes it is better). The mean value theorem implies that for any $x \in \boldsymbol{x}$ there is $c \in \boldsymbol{x}$ such that $f'(c) = \frac{f(x)-f(a)}{x-a}$, or $f(x) = f(a) + f'(c)(x - a)$. Thus, we have the following enclosure $f(\boldsymbol{x}) \subseteq f(a) + f'(\boldsymbol{x})(\boldsymbol{x} - a)$. In practice, naturally, we enclose the image of the derivative $f'(\boldsymbol{x})$ by its interval extension $\boldsymbol{f}'(\boldsymbol{x})$.

Theorem 6.7. *For $\boldsymbol{x} \in \mathbb{IR}$ and $a \in \boldsymbol{x}$ we have*

$$f(\boldsymbol{x}) \subseteq f(a) + \boldsymbol{f}'(\boldsymbol{x})(\boldsymbol{x} - a), \tag{6.1}$$
$$f(x) \subseteq f(a) + \boldsymbol{f}'(\boldsymbol{x})(x - a) \quad \forall x \in \boldsymbol{x}. \tag{6.2}$$

Throughout this section, we use the notation $\boldsymbol{d} \coloneqq \boldsymbol{f}'(\boldsymbol{x})$ and $\boldsymbol{y}$ for the mean value form enclosure. Thus, formulae (6.1)–(6.2) read

$$f(\boldsymbol{x}) \subseteq f(a) + \boldsymbol{d}(\boldsymbol{x} - a) \equiv \boldsymbol{y}, \tag{6.3}$$
$$f(x) \subseteq f(a) + \boldsymbol{d}(x - a) \quad \forall x \in \boldsymbol{x}. \tag{6.4}$$

The expression (6.3) provides an interval enclosure of the image $f(\boldsymbol{x})$, while (6.4) encloses function $f(x)$ on $x \in \boldsymbol{x}$ by an interval linear function.. Example 6.8 illustrates the former and Figure 6.1 the latter.

Example 6.8. Let $f(x) = x^5 - 2x^3$ and $\boldsymbol{x} = [1.8, 2]$. The natural interval extension yields $f(\boldsymbol{x}) \subseteq [2.8956, 20.3361]$. For $a \coloneqq x^c = 1.9$, the mean value theorem gives $f(\boldsymbol{x}) \subseteq [4.9869, 17.0990]$, and for $a \coloneqq \underline{x} = 1.8$, we get a different enclosure $f(\boldsymbol{x}) \subseteq [7.2316, 19.3437]$ with the same width.

However, for $\boldsymbol{x} = [1, 2]$ we get quite poor results. The natural interval extension yields $f(\boldsymbol{x}) \subseteq [-15, 30]$. The mean value theorem enclosure are worse. For $a \coloneqq x^c = 1.9$ we have $f(\boldsymbol{x}) \subseteq [-36.1563, 37.8438]$, and for $a \coloneqq \underline{x} = 1.8$, we have $f(\boldsymbol{x}) \subseteq [-20, 73]$. This shows that the mean value theorem is useful mainly for the narrow intervals. □

How to choose an appropriate $a \in \boldsymbol{x}$? The center $a \coloneqq x^c$ may not be the best choice, however, the radius of the enclosure is $\mathrm{mag}(\boldsymbol{d})x^\Delta$, which is minimal for all selections. More generally, we have:

Proposition 6.9. *If $0 \notin \boldsymbol{d}$, then y^Δ is constant for all $a \in \boldsymbol{x}$. If $0 \in \boldsymbol{d}$, then y^Δ is minimal for any $a \in x^c + x^\Delta \frac{d^c}{d^\Delta}[-1,1]$.*

Proof. If $\boldsymbol{d} > 0$, then $\overline{y} = f(a) + \overline{d}(\overline{x} - a)$, $\underline{y} = f(a) + \overline{d}(\underline{x} - a)$, and therefore $y^\Delta = \overline{d}x^\Delta = \mathrm{mag}(\boldsymbol{d})x^\Delta$. If $\boldsymbol{d} < 0$, then analogously $y^\Delta = \mathrm{mag}(\boldsymbol{d})x^\Delta$.

For the second part, we first show that $a \coloneqq x^c$ yields the most narrow enclosure. This is easy to see since

$$y^\Delta = \mathrm{rad}(\boldsymbol{d}(\boldsymbol{x} - a)) \geq \mathrm{rad}(\mathrm{mag}(\boldsymbol{d})(\boldsymbol{x} - a)) = \mathrm{mag}(\boldsymbol{d})x^\Delta = \mathrm{rad}(\boldsymbol{d}(\boldsymbol{x} - x^c)).$$

Now, let $\alpha \in [-1,1]$, and denote $a = x^c + \alpha x^\Delta \frac{d^c}{d^\Delta}$. Then,

$$\begin{aligned}\boldsymbol{f}'(\boldsymbol{x})(\boldsymbol{x} - a) &= \boldsymbol{d}\left(\boldsymbol{x} - x^c - \alpha x^\Delta \frac{d^c}{d^\Delta}\right) = -\frac{x^\Delta}{d^\Delta}\boldsymbol{d}[\alpha d^c - d^\Delta, \alpha d^c + d^\Delta] \\ &= -\frac{x^\Delta}{d^\Delta}\,\mathrm{mag}(\boldsymbol{d})[\alpha d^c - d^\Delta, \alpha d^c + d^\Delta].\end{aligned}$$

The radius of this interval is also $\mathrm{mag}(\boldsymbol{d})x^\Delta$. □

If $0 \notin \boldsymbol{d}$, then the function is increasing. Thus, the optimal choice is to take both $a \coloneqq \underline{x}$ and $a \coloneqq \overline{x}$ and intersect the results since it produces the exact lower and upper bounds on $f(\boldsymbol{x})$. The additional cost of this bi-centered form is low as the interval derivative $\boldsymbol{f}'(\boldsymbol{x})$ is calculated just once.

If $0 \in \boldsymbol{d}$, a useful selection of the center is the Baumann point (Baumann, 1988). For this center, we achieve the tightest lower bound of $f(\boldsymbol{x})$ by maximizing the left endpoint of the enclosure (6.1). In a similar manner, we find a center a minimizing the right-end point of the enclosure.

Theorem 6.10 (Baumann, 1988)**.** *If $0 \in \boldsymbol{d}$, then $\underline{y}$ has the maximal value for*

$$a = \frac{\overline{d}\underline{x} - \underline{d}\overline{x}}{2d^\Delta} = x^c - x^\Delta \frac{d^c}{d^\Delta} \in \boldsymbol{x}. \tag{6.5}$$

Proof. First, we show that the left endpoint of the interval $\boldsymbol{d}(\boldsymbol{x} - a) = \boldsymbol{d}(\boldsymbol{x} - a)$ is maximal for the value of a from (6.5). Since both intervals $\boldsymbol{d}$ and $\boldsymbol{x} - a$ contain the zero, we have

$$\inf(\boldsymbol{d}(\boldsymbol{x} - a)) = \min\{\underline{d}(\overline{x} - a), \overline{d}(\underline{x} - a)\}.$$

Since $\underline{d}(\overline{x} - a)$ nondecreasing in a and $\overline{d}(\underline{x} - a)$ nonincreasing in a, the minimum value is maximal when both functions have the same value, that is

$$\underline{d}(\overline{x} - a) = \overline{d}(\underline{x} - a)$$

Expressing a from this equation, we arrive at the final form.

Now, we incorporate into considerations also the absolute term $f(a)$. The idea is to show that a change of the center to a point a' may improve $f(a')$, but the decrease in the second term is superior. If $a' > a$, then

$$\begin{aligned}\inf(\boldsymbol{d}(\boldsymbol{x} - a)) &= \overline{d}(\underline{x} - a), \\ \inf(\boldsymbol{d}(\boldsymbol{x} - a')) &= \overline{d}(\underline{x} - a'),\end{aligned}$$

whence

$$\inf(\boldsymbol{d}(\boldsymbol{x} - a)) - \inf(\boldsymbol{d}(\boldsymbol{x} - a')) = \overline{d}(a' - a).$$

On the other hand, by the mean value theorem,

$$f(a') - f(a) \leq \overline{d}(a' - a),$$

from which the statement follows. Similarly, if $a' < a$, then

$$\inf(\boldsymbol{d}(\boldsymbol{x}-a)) = \underline{d}(\overline{x}-a),$$
$$\inf(\boldsymbol{d}(\boldsymbol{x}-a')) = \underline{d}(\overline{x}-a').$$

Hence

$$\inf(\boldsymbol{d}(\boldsymbol{x}-a)) - \inf(\boldsymbol{d}(\boldsymbol{x}-a')) = \underline{d}(a'-a).$$

On the other hand, by the mean value theorem,

$$f(a') - f(a) \le \underline{d}(a'-a),$$

proving the second case. □

This result is straightforwardly extended to a multi-dimensional case. In the original paper, Baumann (1988) also considered a general Lipschitz function instead of the particular case of the derivative $\boldsymbol{f}'$.

We can also check quality of this enclosure. In general, quality of enclosures is measured via *inner estimations* (Krawczyk, 1980; Markov, 1990). An inner estimation of $f(\boldsymbol{x})$ is simply any interval $\boldsymbol{a} \in \mathbb{IR}$ such that $\boldsymbol{a} \subseteq f(\boldsymbol{x})$.

The inclusion $f(x) \in f(a) + \boldsymbol{d}(x-a)$ gives us lower and upper approximations of $f(x)$. Function $f(x)$ lies between two lines $\ell_1(x) = f(a) + \overline{d}(x-a)$ and $\ell_2(x) = f(a) + \underline{d}(x-a)$. The maximum spread between the lines is

$$\max_{x\in\boldsymbol{x}} |\ell_1(x) - \ell_2(x)| = \max_{x\in\boldsymbol{x}} 2d^{\Delta}|x-a| = 2d^{\Delta}\operatorname{mag}(\boldsymbol{x}-a).$$

Therefore $\underline{y}$ and $\overline{y}$ overestimate the exact minimum and maximum of the range by no more than $2d^{\Delta}\operatorname{mag}(\boldsymbol{x}-a)$, yielding:

Theorem 6.11 (Krawczyk and Neumaier, 1986; Neumaier, 1990)**.** *We have*

$$\sup(f(\boldsymbol{x}) \ge \overline{y} - 2d^{\Delta}\operatorname{mag}(\boldsymbol{x}-a),$$
$$\inf(f(\boldsymbol{x}) \le \underline{y} + 2d^{\Delta}\operatorname{mag}(\boldsymbol{x}-a).$$

As a consequence, we have a quadratic convergence of the mean value form, which is true also for the multi-dimensional case (Alefeld and Herzberger, 1983; Chuba and Miller, 1972; Krawczyk and Nickel, 1982; Miller, 1972).

Theorem 6.12. *Suppose that the derivative $f'(x)$ with some $\boldsymbol{u} \in \mathbb{IR}$ satisfies Property L (page 118), and its enclosure is calculated by the natural extension. Denote by $\boldsymbol{f}(\boldsymbol{x})$ the mean value extension of $f(x)$. Then there is $\gamma > 0$ such that for each $\boldsymbol{x} \in \mathbb{IR}^n$, $\boldsymbol{x} \subseteq \boldsymbol{u}$, we have*

$$\operatorname{rad}(\boldsymbol{f}(\boldsymbol{x})) - \operatorname{rad}(f(\boldsymbol{x})) \le \gamma(x^{\Delta})^2.$$

Proof. By Theorem 6.11 the overestimation of the mean value form $\boldsymbol{y}$ over the image $f(\boldsymbol{x})$ is at most

$$4\operatorname{rad}(\boldsymbol{f}'(\boldsymbol{x}))\operatorname{mag}(\boldsymbol{x}-a).$$

We estimate $\operatorname{mag}(\boldsymbol{x}-a) \le 2x^{\Delta}$, and by Theorem 6.6, we bound also $\operatorname{rad}(\boldsymbol{f}'(\boldsymbol{x}))$ from above by a constant multiple of x^{Δ}. □

Example 6.13. Consider the function

$$f(x) = \frac{\exp(x-1) + x^3 - 4x}{x^2+2}$$

and the interval $\boldsymbol{x} = [1-\delta, 1+\delta]$, where δ is a parameter. Table 6.1 displays the enclosures (and their radii) computed by the natural interval extension and by the mean value form for six values of parameter δ.

From the table, we see that the natural extension produces a tighter enclosure for wider input intervals, whereas the mean value form yields a tighter enclosure for narrow input intervals. This is in correspondence with the linear convergence of the natural extension (Theorem 6.5) and quadratic convergence of the mean value form (Theorem 6.12). □

Table 6.1: (Example 6.13) Comparison of natural interval extension and mean value form.

δ	natural interval extension		mean value form	
	enclosure	radius	enclosure	radius
0.1	$[-0.9844, -0.3625]$	0.3109	$[-0.7737, -0.5597]$	0.1070
0.2	$[-1.3142, -0.0728]$	0.6206	$[-1.0578, -0.2755]$	0.3911
0.3	$[-1.6531, 0.3000]$	0.9765	$[-1.6166, 0.2833]$	0.9499
0.4	$[-1.9974, 0.7779]$	1.3876	$[-2.5735, 1.2402]$	1.9068
0.5	$[-2.3416, 1.3439]$	1.8427	$[-4.0741, 2.7408]$	3.4074
0.6	$[-2.6793, 1.9992]$	2.3392	$[-6.2748, 4.9415]$	5.6081

Higher order expansion. In one-dimensional case, we can effectively use higher order derivatives and the corresponding enclosures. Let $f(x)$ be $m+1$ times differentiable on $\boldsymbol{x}$, and $a \in \boldsymbol{x}$. For each $x \in \boldsymbol{x}$ we have by Taylor's Theorem 1.24

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \cdots + \frac{f^{(m)}(a)}{m!}(x-a)^m + \frac{f^{(m+1)}(c)}{(m+1)!}(x-a)^{m+1}$$

for some $c \in \boldsymbol{x}$. Thus, we have the enclosure

$$f(\boldsymbol{x}) \subseteq \sum_{k=1}^{m} \frac{f^{(k)}(a)}{k!}(\boldsymbol{x}-a)^k + \frac{f^{(m+1)}(\boldsymbol{x})}{(m+1)!}(\boldsymbol{x}-a)^{m+1}.$$

Notice that this approach can be used not only to enclose the range of $f(x)$ on $\boldsymbol{x}$ but also to approximate $f(x)$ from above and from below by a polynomial of degree $m+1$ since

$$f(x) \in \sum_{k=1}^{m} \frac{f^{(k)}(a)}{k!}(x-a)^k + \frac{f^{(m+1)}(\boldsymbol{x})}{(m+1)!}(x-a)^{m+1}. \tag{6.6}$$

Denote $\boldsymbol{y} \coloneqq \frac{f^{(m+1)}(\boldsymbol{x})}{(m+1)!}$. Now,

$$\sum_{k=1}^{m} \frac{f^{(k)}(a)}{k!}(x-a)^k + \underline{y}(x-a)^{m+1} \le f(x) \le \sum_{k=1}^{m} \frac{f^{(k)}(a)}{k!}(x-a)^k + \overline{y}(x-a)^{m+1}$$

for $x \in \boldsymbol{x}$ provided m is odd. As long as m is even, we have to distinguish the sub-intervals $[\underline{x}, a]$ and $[a, \overline{x}]$.

Example 6.14. Let $f(x) = e^x$, $\boldsymbol{x} = [0,1]$, and $a = 0$. For $m = 1$ we have

$$f(x) \in 1 + t + \frac{1}{2}[1, e]t^2,$$

which encloses the exponential between two polynomials of degree two on $\boldsymbol{x}$. For $m = 9$ we obtain

$$f(x) \in 1 + t + \frac{1}{2}t^2 + \cdots + \frac{1}{10!}[1, e]t^{10},$$

which gives a tight enclosure to $f(x)$. Notice that, however, higher degree expansion need not result in a more tighter enclosure. For instance, take the exponential function on an interval with large values. □

6.1.2 Multi-dimensional case

Suppose that $f\colon \mathbb{R}^n \to \mathbb{R}$ is differentiable on an open superset of $\boldsymbol{x} \in \mathbb{IR}^n$, and let $a \in \boldsymbol{x}$. By the mean value theorem, we have for any $x \in \boldsymbol{x}$

$$f(x) = f(a) + \nabla f(c)^T(x-a)$$

for some $c \in \boldsymbol{x}$. The naive intervalization results in the enclosure

$$\begin{aligned} f(\boldsymbol{x}) &\subseteq f(a) + \nabla \boldsymbol{f}(\boldsymbol{x})^T (\boldsymbol{x} - a) \\ &= \sum_{i=1}^{n} f'_{x_i}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)(\boldsymbol{x}_i - a_i), \end{aligned} \tag{6.7}$$

where $\nabla \boldsymbol{f}(\boldsymbol{x})$ is an interval extension of the gradient $\nabla f(x)$ on $\boldsymbol{x}$. Caprani and Madsen (1980) proved that when $a = x^c$ and $\nabla \boldsymbol{f}(\boldsymbol{x})$ is inclusion isotone, then the mean value form is inclusion isotone; cf. Neumaier (1990).

Better enclosure is obtained when we apply the mean value theorem component-wisely as proposed by Hansen (1968); cf. Hansen and Walster (2004). For each $x \in \boldsymbol{x}$ there is some $c \in \boldsymbol{x}$ such that

$$\begin{aligned} f(x) &= f(x_1, \ldots, x_{n-1}, a_n) + f'_{x_n}(x_1, \ldots, x_{n-1}, c_n)(x_n - a_n) \\ &= f(x_1, \ldots, x_{n-2}, a_{n-1}, a_n) + f'_{x_{n-1}}(x_1, \ldots, x_{n-2}, c_{n-1}, a_n)(x_{n-1} - a_{n-1}) + \\ &\quad + f'_{x_n}(x_1, \ldots, x_{n-1}, c_n)(x_n - a_n) = \cdots = \\ &= f(a) + f'_{x_1}(c_1, a_2, \ldots, a_n)(x_1 - a_1) + f'_{x_2}(x_1, c_2, a_3, \ldots, a_n)(x_2 - a_2) + \\ &\quad + \cdots + f'_{x_n}(x_1, \ldots, x_{n-1}, c_n)(x_n - a_n). \end{aligned}$$

Hence

$$\begin{aligned} f(\boldsymbol{x}) \subseteq f(a) &+ f'_{x_1}(\boldsymbol{x}_1, a_2, \ldots, a_n)(\boldsymbol{x}_1 - a_1) + f'_{x_2}(\boldsymbol{x}_1, \boldsymbol{x}_2, a_3, \ldots, a_n)(\boldsymbol{x}_2 - a_2) + \\ &+ \cdots + f'_{x_n}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{n-1}, \boldsymbol{x}_n)(\boldsymbol{x}_n - a_n). \end{aligned}$$

Obviously, this is a tighter enclosure than (6.7) since many arguments of the partial derivatives are fixed to some a_i, $i = 2, \ldots, n$.

6.2 Slopes

Slopes were introduced by Herzberger and then studied by Krawczyk, Neumaier and others to replace derivatives in the mean value theorem approach by something more subtle. For the sake of simplicity, we consider the one-dimensional case with a function $f\colon \mathbb{R} \to \mathbb{R}$ on $\boldsymbol{x} \in \mathbb{IR}$ and an arbitrary but fixed $a \in \boldsymbol{x}$. The multiobjective case is studied, e.g., in Ratz (2001).

The basic idea of slopes goes as follows. For $x \neq a$,

$$f(x) = f(a) + \frac{f(x) - f(a)}{x - a}(x - a).$$

Define

$$S_f(x, a) \coloneqq \begin{cases} \frac{f(x)-f(a)}{x-a} & \text{if } x \neq a, \\ f'(x) & \text{otherwise.} \end{cases}$$

whence $f(x) = f(a) + S_f(x, a)(x - a)$. If we can find a tight enclosure to $S_f(\boldsymbol{x}, a)$, then we have an efficient enclosure to $f(\boldsymbol{x})$ by

$$f(\boldsymbol{x}) \subseteq f(a) + S_f(\boldsymbol{x}, a)(\boldsymbol{x} - a).$$

Thus, the problem is to compute a tight enclosure to $S_f(\boldsymbol{x}, a)$. By the mean value theorem, the value of $S_f(x, a)$ can be approximated by the derivative of f at some $c \in \boldsymbol{x}$. But estimating $S_f(\boldsymbol{x}, a) \subseteq f'(\boldsymbol{x})$ is too conservative and results in the traditional mean value theorem enclosure. Since $S_f(\boldsymbol{x}, a) \subseteq S_f(\boldsymbol{x}, \boldsymbol{x}) = f'(\boldsymbol{x})$, often with the proper inclusion, the slopes are more efficient to derivatives.

Example 6.15. Let $f(x) = \frac{1}{4}x^2 - x + \frac{1}{2}$ and $\boldsymbol{x} = [1, 7]$. Since $f'(x) = \frac{1}{2}x - 1$, we have $f'(\boldsymbol{x}) = [-\frac{1}{2}, \frac{5}{2}]$. For slopes,

$$S_f(x, a) = \frac{\frac{1}{4}x^2 - x + \frac{1}{2} - (\frac{1}{4}a^2 - a + \frac{1}{2})}{x - a} = \frac{1}{4}(x + a) - 1.$$

For $a \coloneqq x^c = 4$, we have $S_f(\boldsymbol{x}, a) = [\frac{1}{4}, \frac{7}{4}]$. The enclosure to $f(\boldsymbol{x})$ by slopes is

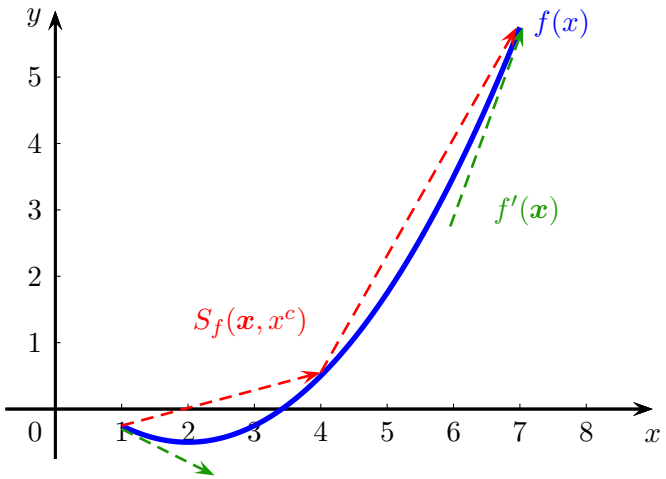


Figure 6.2: (Example 6.15) The range of slopes (in red) and the range of derivatives (in green).

Table 6.2: Chain rules for computing slopes.

function	its slope $S(x,a)$
constant	0
x	1
$\alpha f(x)$	$\alpha S_f(x,a)$
$f(x) \pm g(x)$	$S_f(x,a) \pm S_g(x,a)$
$f(x) \cdot g(x)$	$S_f(x,a)g(a) + f(x)S_g(x,a)$
$\dfrac{f(x)}{g(x)}$	$\dfrac{S_f(x,a) - \frac{f(a)}{g(a)}S_g(x,a)}{g(x)}$
$e^{f(x)}$	$e^{f(x)}S_f(x,a)$
$\ln(f(x))$	$\dfrac{S_f(x,a)}{f(x)}$
$\sin(f(x))$	$\cos(f(x))S_f(x,a)$
$\cos(f(x))$	$-\sin(f(x))S_f(x,a)$

$$f(a) - S_f(\boldsymbol{x},a)(\boldsymbol{x}-a) = \tfrac{1}{2} - [\tfrac{1}{4},\tfrac{7}{4}]\,([1,7]-4) = [-\tfrac{19}{4},\tfrac{23}{4}],$$

while the mean value theorem approach gives only $[-7,8]$; see Figure 6.2. For comparison, the natural interval extension yields $f(\boldsymbol{x}) \subseteq [-\frac{25}{4},\frac{47}{4}]$. Rewriting the expression to the form of $f(x) = (\frac{1}{2}x-1)^2 - \frac{1}{2}$, we get an expression with no dependencies and the natural interval extension gives the true range $f(\boldsymbol{x}) = [-\frac{1}{2},\frac{23}{4}]$. □

Calculation of slopes. Slopes can be calculated by rules similar to that for derivatives. Table 6.2 summarizes them for elementary operations and basic functions. Notice the slight differences to derivatives in the multiplication and division rules. This makes the slopes more efficient to derivatives.

It is an easy exercise to prove the chain rules. For instance, take the multiplication rule. Herein,

$$S_{fg}(x,a) = \frac{f(x)g(x) - f(a)g(a)}{x-a},$$

and

$$S_f(x,a)g(a) + f(x)S_g(x,a) = \frac{f(x)-f(a)}{x-a}g(a) + f(x)\frac{g(x)-g(a)}{x-a} = \frac{f(x)g(x)-f(a)g(a)}{x-a}.$$

Notice that the chain rules for multiplication and division can be replaced by their equivalent forms

$$S_{fg}(x,a) = S_f(x,a)g(x) + f(a)S_g(x,a),$$

and

$$S_{f/g}(x,a) = \frac{S_f(x,a) - \frac{f(x)}{g(x)}S_g(x,a)}{g(a)}.$$

Indeed, INTLAB implements both alternatives and results the intersection of both in order to keep things as tight as possible.

As for classical differentiation, slopes can be evaluated either automatically, or symbolically; cf. Section 6.6.

Example 6.16 (Polynomials)**.** Consider a polynomial $p(x) = \sum_{k=0}^{n} c_k x^k$. The slope of $p(x)$ can be computed in diverse ways, but the following form is particularly convenient for interval evaluation (Alefeld, 1981; Alefeld and Herzberger, 1983)

$$S_p(x,a) = \frac{p(x)-p(a)}{x-a} = \sum_{k=1}^{n} c_k \frac{x^k - a^k}{x-a} = \sum_{k=1}^{n} c_k \sum_{i=1}^{k} a^{k-i}x^{i-1} = \sum_{k=1}^{n} \left(\sum_{i=k}^{n} c_i a^{i-k} \right) x^{k-1}.$$

In this way, $S_p(x,a)$ is expressed as a real polynomial and so an enclosure of the image $S_p(\boldsymbol{x},a)$ is computed by a suitable method for enclosing polynomials; see Section 6.5. □

Overestimation. In a similar way as in Theorem 6.11 we can bound the overestimation of the slope form; see Krawczyk and Neumaier (1985); Neumaier (1990).

Theorem 6.17 (Krawczyk and Neumaier, 1985)**.** *Let $f\colon \mathbb{R}^n \to \mathbb{R}$, $\boldsymbol{x} \in \mathbb{IR}^n$ and $a \in \boldsymbol{a}$. Denote*

$$\boldsymbol{y} \coloneqq f(a) - S_f(\boldsymbol{x},a)(\boldsymbol{x}-a).$$

Then

$$\overline{f}(x) \geq \overline{y} - 2S_f(\boldsymbol{x},a)^{\Delta}\operatorname{mag}(\boldsymbol{x}-a),$$
$$\underline{f}(x) \leq \underline{y} + 2S_f(\boldsymbol{x},a)^{\Delta}\operatorname{mag}(\boldsymbol{x}-a).$$

In particular, putting $a \coloneqq x^c$ we get

$$\overline{f}(x) \geq \overline{y} - 2S_f(\boldsymbol{x},x^c)^{\Delta}x^{\Delta},$$
$$\underline{f}(x) \leq \underline{y} + 2S_f(\boldsymbol{x},x^c)^{\Delta}x^{\Delta}.$$

Monotonicity test. Even though slopes outperform derivatives in many ways, they cannot be used for monotonicity checking. As Example 6.15 or Example 6.18 below illustrate, it may happen that $0 \notin S_f(\boldsymbol{x},a)$, but $f(x)$ is not monotone on $\boldsymbol{x}$.

Example 6.18. Let $f(x) = \sin x$, let $\boldsymbol{x} = [-\pi,\pi]$ and take $a \coloneqq x^c = 0$. Now, $S_f(\boldsymbol{x},a) = [0,1]$, but the sine is not non-decreasing on the whole interval $[-\pi,\pi]$. □

However, the situation is not so bad. Let us restrict our attention to the left endpoint $\underline{f}(\boldsymbol{x})$ only, and suppose that $S_f(\boldsymbol{x},a) > 0$. Even though the condition $0 \notin S_f(\boldsymbol{x},a)$ does not imply $\underline{f}(\boldsymbol{x}) = f(\underline{x})$, we are often able to significantly reduce the domain $\boldsymbol{x}$, where the minimal value is attained; cf. Ratz (2001).

Theorem 6.19. *Denote $\boldsymbol{s} \coloneqq S_f(\boldsymbol{x},a)$, and $p \coloneqq a + (\underline{x}-a)\underline{s}/\overline{s}$. If $\underline{s} > 0$, then $p \in [\underline{x},a)$ and*

$$\min_{x\in[\underline{x},p]} f(x) < \min_{x\in(p,\overline{x}]} f(x).$$

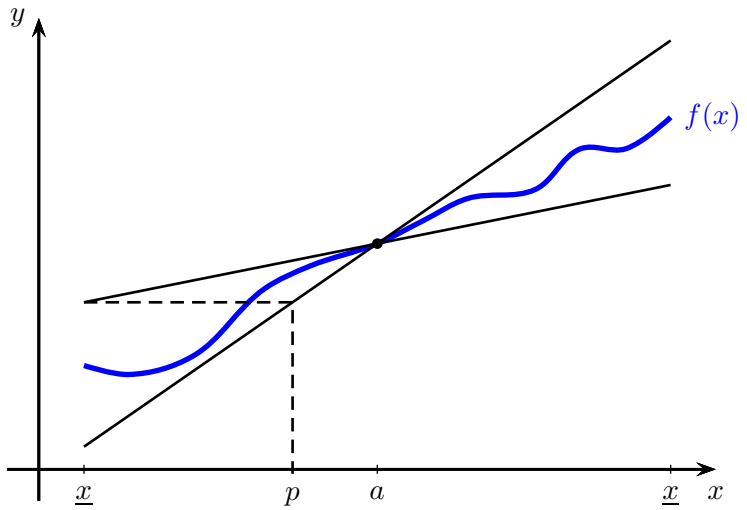


Figure 6.3: A geometrical interpretation of the value p from Theorem 6.19.

Proof. Obviously, $p \in [\underline{x}, a)$. Let $s^1 \in \boldsymbol{s}$ be such that $f(\underline{x}) = f(a) + s^1(\underline{x} - a)$. Let $\min_{x\in(p,\overline{x}]} f(x) = f(x^2)$, and $s^2 \in \boldsymbol{s}$ be such that $f(x^2) = f(a) + s^2(x^2 - a)$. Then we have

$$\begin{aligned}\min_{x\in[\underline{x},p]} f(x) &\le f(\underline{x}) = f(a) + s^1(\underline{x} - a) \le f(a) + \underline{s}(\underline{x} - a) = f(a) + \overline{s}(p - a)\\ &< f(a) + s^2(x^2 - a) = \min_{x\in(p,\overline{x}]} f(x).\end{aligned}$$

□

By this theorem, $\underline{f}(\boldsymbol{x})$ is achieved on $[\underline{x}, p] \subsetneqq \boldsymbol{x}$, that is, $\underline{f}(\boldsymbol{x}) = \underline{f}([\underline{x}, p])$. A geometrical interpretation of the value p is illustrated in Figure 6.3 by enclosing a function $f(x)$ with help of the mean value form.

Notes. Look in Krawczyk and Neumaier (1985) for more results on overestimation and for slopes in complex space. For an extension to a second order slope form see, e.g., Schichl and Neumaier (2004). Ratz (1998) provided algorithmic and implementation details, extension to multidimensional case, and applied slopes in nonsmooth global optimization.

6.3 Other methods

Centered form

We write $f(x)$ as

$$f(x) = f(x^c) + f(x) - f(x^c) = f(x^c) + f(y + x^c) - f(x^c) = f(x^c) + y^T h(y),$$

where $y \coloneqq x - x^c$ and $h\colon \mathbb{R}^n \to \mathbb{R}^n$ is an appropriate function. We get the enclosure

$$f(\boldsymbol{x}) \subseteq f(x^c) + \boldsymbol{y}^T \boldsymbol{h}(\boldsymbol{y}),$$

where $\boldsymbol{y} \coloneqq \boldsymbol{x} - x^c$ and $\boldsymbol{h}\colon \mathbb{IR}^n \to \mathbb{IR}^n$ is an interval extension of $h(y)$. We illustrate the centered form and finding of the function $h(y)$ by an example.

Example 6.20. Let $f(x) = 1 - 5x + \frac{1}{3}x^3$ and write

$$\begin{aligned}f(x) - f(x^c) &= -5(x - x^c) + \tfrac{1}{3}(x^3 - x^{c3}) = (x - x^c)(-5 + \tfrac{1}{3}(x^2 + xx^c + x^{c2}))\\ &= y(-5 + \tfrac{1}{3}((y + x^c)^2 + (y + x^c)x^c + x^{c2})) = y\,h(y).\end{aligned}$$

Thus

$$f(\boldsymbol{x}) \subseteq f(x^c) + \boldsymbol{y}^T \boldsymbol{h}(\boldsymbol{y}) = f(x^c) + \boldsymbol{y}(-5 + \tfrac{1}{3}((\boldsymbol{y} + x^c)^2 + (\boldsymbol{y} + x^c)x^c + x^{c2})).$$

For concreteness, let $\boldsymbol{x} = [2, 3]$. The natural interval extension of $f(x)$ yields $f(\boldsymbol{x}) \subseteq [-\frac{34}{3}, 0]$, and the optimal enclosure is $f(\boldsymbol{x}) = [-\frac{10}{3}\sqrt{5}+1, -5] \subseteq [-6.4546, -5]$. The centered form yields $f(\boldsymbol{x}) \subseteq [-\frac{91}{12}, -5] \subseteq [-7.5834, -5]$. □

6.4 Continuity and convergence

For intervals $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{IR}$, we define their *distance* as

$$\operatorname{dist}(\boldsymbol{a}, \boldsymbol{b}) \coloneqq \max\{|\underline{a} - \underline{b}|, |\overline{a} - \overline{b}|\}.$$

It is easy to see that it is a metric in $\mathbb{IR}$, that is, it satisfies the axioms

(1) $\operatorname{dist}(\boldsymbol{a}, \boldsymbol{b}) \geq 0$ with equality if and only if $\boldsymbol{a} = \boldsymbol{b}$,
(2) $\operatorname{dist}(\boldsymbol{a}, \boldsymbol{b}) = \operatorname{dist}(\boldsymbol{b}, \boldsymbol{a})$,
(3) $\operatorname{dist}(\boldsymbol{a}, \boldsymbol{b}) \leq \operatorname{dist}(\boldsymbol{a}, \boldsymbol{c}) + \operatorname{dist}(\boldsymbol{c}, \boldsymbol{b})$.

For interval vectors $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{IR}^n$, we extend the definition as follows

$$\operatorname{dist}(\boldsymbol{a}, \boldsymbol{b}) \coloneqq \max_{i=1,\ldots,n} \max\{|\underline{a}_i - \underline{b}_i|, |\overline{a}_i - \overline{b}_i|\}. \tag{6.8}$$

Basically, this is the standard Hausdorff distance of two sets in the case the sets are boxes. In general, for sets $\mathcal{A}, \mathcal{B} \subseteq \mathbb{R}^n$, their *Hausdorff distance* is defined

$$\operatorname{dist}(\mathcal{A}, \mathcal{B}) \coloneqq \max\left\{ \sup_{x \in \mathcal{A}} \operatorname{dist}(x, \mathcal{B}), \sup_{y \in \mathcal{B}} \operatorname{dist}(y, \mathcal{A}) \right\},$$

where

$$\operatorname{dist}(x, \mathcal{B}) \coloneqq \inf_{y \in \mathcal{B}} \|x - y\|$$

for a given vector norm; in (6.8), we used the maximum norm.

The distance function enables to define continuity of an interval function in a natural way, extending the standard definition for real functions. Notice that a certain kind of Lipschitz continuity was already addressed in Theorem 6.6.

Definition 6.21. An interval function $\boldsymbol{f}\colon \mathbb{IR}^n \to \mathbb{IR}$ is *continuous* if for every $\varepsilon > 0$ there is $\delta > 0$ such that for every $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{IR}^n$:

$$\operatorname{dist}(\boldsymbol{x}, \boldsymbol{y}) < \delta \quad \Rightarrow \quad \operatorname{dist}(\boldsymbol{f}(\boldsymbol{x}), \boldsymbol{f}(\boldsymbol{y})) < \varepsilon.$$

Observation 6.22. *The natural interval extension is a continuous function as long as there is no division by an interval containing zero.*

An interval extension $\boldsymbol{f}$ of a function $f\colon \mathbb{R}^n \to \mathbb{R}$ is *convergent of order* α on $\boldsymbol{x} \in \mathbb{IR}^n$ if there is a constant $K > 0$ such that

$$\operatorname{dist}(\boldsymbol{f}(\boldsymbol{y}), \square f(\boldsymbol{y})) \leq K \Big(\max_{i=1,\ldots,n} y_i^{\Delta} \Big)^{\alpha}$$

for every $\boldsymbol{y} \subseteq \boldsymbol{x}$. The natural interval extension (i.e., interval arithmetic) is convergent of order 1. This may be useful for wide intervals, however, when we process boxes near true solutions, the convergence is slow and the cluster effect of gathering many small boxes uses to happen.

Interval extensions of order 2 or higher have more suitable convergence properties for narrow intervals. As we observed in the previous sections, the mean value forms using derivatives or slopes have quadratic convergence.

6.5 Polynomials

Polynomials are frequently appearing functions and enclosing the range of polynomials on intervals is an important issue; the basic approaches are described in Stahl (1995).

Evaluating polynomials with the variable degree and the number of variables is an *NP*-hard problem (Gaganov, 1981, 1985). Nevertheless, it becomes tractable when the number of variables is fixed; a polynomial algorithm was proposed in Grigoryev and Vorobjov (1988).

Let $p(x) = a_n x^n + \cdots + a_1 x + a_0$ be a polynomial and $\boldsymbol{x} \in \mathbb{IR}$. The natural interval extension

$$\boldsymbol{p}(\boldsymbol{x}) \coloneqq a_n \boldsymbol{x}^n + \cdots + a_1 \boldsymbol{x} + a_0$$

gives too conservative enclosures. Therefore, we discuss alternative forms.

Horner form. Utilizing the Horner scheme

$$\boldsymbol{p}^H(\boldsymbol{x}) \coloneqq (\cdots(a_n\boldsymbol{x} + a_{n-1})\boldsymbol{x} + \cdots + a_1)\boldsymbol{x} + a_0,$$

we get much better results in general. Since it is a simple arithmetic expression, it is inclusion isotonic.

In some cases, such as the one mentioned below, the Horner scheme yields the exact range (Stahl, 1995). The intermediate evaluations of the Horner scheme are denoted as

$$\begin{aligned} p_n(x) &= a_n, \\ p_i(x) &= p_{i+1}(x)x + a_i, \quad i = n-1, \ldots, 0. \end{aligned}$$

Theorem 6.23 (Stahl, 1995)**.** *We have* $\boldsymbol{p}^H(\boldsymbol{x}) = \boldsymbol{p}(\boldsymbol{x})$ *provided either*

$$\underline{x} \geq 0 \quad \textit{and} \quad \underline{p}_i^H(\boldsymbol{x}) \geq \quad \forall i = 0, \ldots, n,$$

or

$$\underline{x} \leq 0 \quad \textit{and} \quad \inf\bigl((-1)^{n-i}\boldsymbol{p}_i^H(\boldsymbol{x})\bigr) \geq \quad \forall i = 0, \ldots, n.$$

From this point of view, it might be a good tightening technique for an interval $\boldsymbol{x}$ containing the zero in its interior (i.e., $\underline{x} < 0 < \overline{x}$) to split $\boldsymbol{x}$ in the zero and evaluate separately the positive and negative parts. Thus, we compute the enclosure $\boldsymbol{p}_i^H([\underline{x}, 0]) \cup \boldsymbol{p}_i^H([0, \overline{x}])$. Indeed, Stahl (1995) showed that the resulting overestimation decreases to less than one half provided $x^c = 0$.

Mean value form form. This form was first developed by Hansen (1969b) and works in the same manner as for the general functions. That is, the enclosure has the form of

$$p(\boldsymbol{x}) \subseteq \boldsymbol{p}^M(\boldsymbol{x}, a) \coloneqq p(a) + \boldsymbol{p}'(\boldsymbol{x})(\boldsymbol{x} - a),$$

where $a \in \boldsymbol{x}$, and $\boldsymbol{p}'(\boldsymbol{x})$ is some interval extension enclosing the derivative of $p(x)$ on $x \in \boldsymbol{x}$.

What is a good choice of the center a? The midpoint value $a \coloneqq x^c$ is a simple choice, but does not yield the tightest enclosure in general. Similarly as described in Theorem 6.10, we can calculate two centers $a_1, a_2 \in \boldsymbol{x}$. The first one aims to maximize the lower bound of the resulting enclosure, and the second one minimizes the upper bound. In the case of polynomials, the bounds computed by Theorem 6.10 are provably the best possible, that is,

$$\begin{aligned} \underline{p}^M(\boldsymbol{x}, a_1) &\geq \underline{p}^M(\boldsymbol{x}, a) \quad \forall a \in \boldsymbol{x}, \\ \overline{p}^M(\boldsymbol{x}, a_2) &\leq \overline{p}^M(\boldsymbol{x}, a) \quad \forall a \in \boldsymbol{x}. \end{aligned}$$

The resulting enclosure

$$\bigl[\underline{p}^M(\boldsymbol{x}, a_1),\, \overline{p}^M(\boldsymbol{x}, a_2)\bigr]$$

is called the bicentred form. Both the basic mean value form and the bicentred version are inclusion isotonic (Caprani and Madsen, 1980; Stahl, 1995).

The derivative in the mean value form can again be replaced by slopes. See Example 6.16. The slope form $\boldsymbol{p}^S(\boldsymbol{x})$, however, is not inclusion isotonic. To see it, consider the polynomial $p(x) = -3x^3 + 3x^2 + 8x$ with domains $\boldsymbol{x}_1 = [0, 1]$ and $\boldsymbol{x}_2 = [0, 2]$. Then $\boldsymbol{p}^S(\boldsymbol{x}_1) = [-0.75,\, 9.5]$, whereas $\boldsymbol{p}^S(\boldsymbol{x}_2) = [0,\, 16]$.

Bernstein form. Bernstein polynomials over the interval $\boldsymbol{x} \in \mathbb{IR}$ are defined

$$b_k(x) = \frac{1}{(2x^\Delta)^n}\binom{n}{k}(x - \underline{x})^k(\overline{x} - x)^{n-k}, \quad k = 0, \ldots, n.$$

They satisfy the following basic properties.

Observation 6.24. *For any $x \in \boldsymbol{x}$, we have*

$$b_k(x) \geq 0,$$
$$\sum_{k=0}^{n} b_k(x) = 1.$$

Proof. The former is obvious, and the latter follows from the binomial expansion:

$$\begin{aligned}\textstyle\sum_{k=0}^{n} b_k(x) &= \tfrac{1}{(2x^{\Delta})^n} \textstyle\sum_{k=0}^{n} \binom{n}{k} (x - \underline{x})^k (\overline{x} - x)^{n-k} \\ &= \tfrac{1}{(2x^{\Delta})^n} (x - \underline{x} + \overline{x} - x)^n = 1.\end{aligned}$$

□

Bernstein polynomials form a basis of the vector space of real polynomials of degree at most n. Therefore, $p(x)$ is a linear combination of these polynomials

$$p(x) = \sum_{k=0}^{n} \beta_k b_k(x),$$

where β_k, $k = 0, \ldots, n$ are the so-called Bernstein coefficients. Due to Observation 6.24, we directly have the important range enclosing property (Cargo and Shisha, 1966).

Theorem 6.25 (Cargo and Shisha, 1966)**.** *We have*

$$p(\boldsymbol{x}) \subseteq \square\{\beta_k,\ k = 0, \ldots, n\}.$$

Proof. For every $x \in \boldsymbol{x}$,

$$\begin{aligned}p(x) = \sum_{k=0}^{n} \beta_k b_k(x) &\leq \max_{k=0,\ldots,n} (\beta_k) \sum_{k=0}^{n} b_k(x) = \max_{k=0,\ldots,n} \beta_k, \\ p(x) = \sum_{k=0}^{n} \beta_k b_k(x) &\geq \min_{k=0,\ldots,n} (\beta_k) \sum_{k=0}^{n} b_k(x) = \min_{k=0,\ldots,n} \beta_k.\end{aligned}$$

□

This observation justifies introducing the Bernstein form of an enclosure as

$$\begin{aligned}\boldsymbol{p}(\boldsymbol{x})^B &\coloneqq \square\{\beta_k,\ k = 0, \ldots, n\} \\ &= \left[\min_{k=0,\ldots,n} \beta_k,\ \max_{k=0,\ldots,n} \beta_k\right].\end{aligned}$$

This form is inclusion isotonic (Hong and Stahl, 1995). The only problem that remains now to solve is how to efficiently calculate the Bernstein coefficients β_k, $k = 0, \ldots, n$. They have the explicit description

$$\beta_k = \sum_{i=0}^{k} \frac{\binom{k}{i}}{\binom{n}{i}} \frac{p^{(i)}(\underline{x})}{i!} (2x^{\Delta})^i. \tag{6.9}$$

As trivial special cases, we have

$$\begin{aligned}\beta_0 &= p(\underline{x}), \\ \beta_n &= \sum_{i=0}^{n} \frac{p^{(i)}(\underline{x})}{i!} (\overline{x} - \underline{x})^i = p(\overline{x}),\end{aligned}$$

where for the second one we used the Taylor expansion of $p(x)$ about the point $\underline{x}$. The formula (6.9) is applicable for calculating the Bernstein coefficients, but there are more efficient techniques based on two-dimensional recurrences (Ratschek and Rokne, 2003; Rokne, 1979; Stahl, 1995). See Ray and Nataraj (2012) for a matrix method for computing the Bernstein coefficients.

Bernstein polynomials in an interval setting are also known to be useful in computer graphics. They play an important role in approximation of functions via Bézier curves Sederberg and Buehler (1992); Sederberg and Farouki (1992); Ratschek and Rokne (2003), but they were also applied in a robust approach to surface intersection Hu et al. (1997), for instance.

Other forms. There were yet another forms studied. The *Taylor form* utilizes the fact that the polynomial $p(x)$ has a finite Taylor expansion. Nevertheless, the Taylor form turned out empirically not to be more convenient for interval evaluation.

The *interpolation form* approximates $p(x)$ by a quadratic function. By Taylor's theorem 1.24, for any $a \in \boldsymbol{x}$, there is $c \in \boldsymbol{x}$ such that

$$p(x) = p(a) + p'(a)(x-a) + \frac{1}{2}p''(c)(x-a)^2.$$

From this we have the enclosure

$$p(\boldsymbol{x}) \subseteq p(a) + p'(a)(x-a) + \frac{1}{2}\boldsymbol{p}''(\boldsymbol{x})(x-a)^2,$$

where $\boldsymbol{p}''(\boldsymbol{x})$ is an interval extension of the second derivative on $\boldsymbol{x}$.

Comparison. Firment (2017) compared the computational performance of various forms. The Horner scheme is the winner with respect to the running time, and even the enclosures are often not so bad. On the other hand, the Bernstein form gives the tightest enclosures on average, but on account of high computational cost. The bicentered mean value form, slope form and interpolation form are somewhere in the middle. They give a reasonable trade-off between time and tightness provided $\boldsymbol{x}$ is narrow. For wider intervals, the interpolation forms loses.

Interval polynomials. An interval polynomial is an interval function $\boldsymbol{p}(x)\colon \mathbb{R} \to \mathbb{IR}$

$$\boldsymbol{p}(x) = \boldsymbol{a}_n x^n + \cdots + \boldsymbol{a}_1 x + \boldsymbol{a}_0,$$

where $\boldsymbol{a}_0, \ldots, \boldsymbol{a}_n \in \mathbb{IR}$. Given $\boldsymbol{x}_n \in \mathbb{IR}$, the problem is to compute the range

$$\boldsymbol{p}(\boldsymbol{x}) = \cup_{x \in \boldsymbol{x}} \boldsymbol{p}(x)$$

or its tight enclosure. It is not hard to see that the range is always a compact interval. The problem of range computation is easily reduced to two or four problems of range computation of real-valued polynomials (Firment, 2017).

Theorem 6.26. *We have*

$$\boldsymbol{p}(\boldsymbol{x}) = \begin{cases} \big[\underline{g}_1(\boldsymbol{x}),\, \overline{g}_2(\boldsymbol{x})\big] & \text{if } \underline{x} \geq 0, \\ \big[\underline{g}_3(\boldsymbol{x}),\, \overline{g}_4(\boldsymbol{x})\big] & \text{if } \overline{x} \leq 0, \\ \big[\underline{g}_1([0,\overline{x}]),\, \overline{g}_2([0,\overline{x}])\big] \cup \big[\underline{g}_3([\underline{x},0]),\, \overline{g}_4([\underline{x},0])\big] & \text{if } \underline{x} < 0 < \overline{x}, \end{cases}$$

where

$$\begin{aligned} g_1(x) &= \textstyle\sum_{i=0}^n \underline{a}_i x^i, \\ g_2(x) &= \textstyle\sum_{i=0}^n \overline{a}_i x^i, \\ g_3(x) &= \textstyle\sum_{i=0}^n \big(a_i^c - (-1)^i a_i^\Delta\big) x^i, \\ g_4(x) &= \textstyle\sum_{i=0}^n \big(a_i^c + (-1)^i a_i^\Delta\big) x^i. \end{aligned}$$

Proof. Let $x \in \boldsymbol{x}$. If $x \geq 0$, then obviously

$$\boldsymbol{p}(x) = \big[\textstyle\sum_{i=0}^n \underline{a}_i x^i,\ \sum_{i=0}^n \overline{a}_i x^i\big].$$

If $x < 0$, then similarly

$$\begin{aligned} \boldsymbol{p}(x) &= \big[(a_n^c - (-1)^n a_n^\Delta)x^n + \cdots + \underline{a}_2 x^2 + \overline{a}_1 x + \underline{a}_0, \\ &\qquad (a_n^c + (-1)^n a_n^\Delta)x^n + \cdots + \overline{a}_2 x^2 + \underline{a}_1 x + \overline{a}_0\big] \\ &= \big[\, g_3(x),\, g_4(x)\,\big]. \end{aligned}$$

From these observations, the resulting formula follows; if $\underline{x} < 0 < \overline{x}$, we simply split the interval into the positive and negative parts $\boldsymbol{x} = [\underline{x}, 0] \cup [0, \overline{x}]$ and apply separately to both sub-intervals. □

Multivariate polynomials. The Bernstein form was extended to the multivariate case over the unit n-dimensional box by Garloff (1986). Efficient evaluation of Bernstein coefficients for sparse multivariate polynomials was presented by Smith (2009). Alternative generalizations consider the polynomials over the unit simplex in $\mathbb{R}^n$. Further extensions to multivariate rational functions were investigated by Narkawicz et al. (2012); inclusion isotonicity of this extended form was proved by Garloff and Hamadneh (2016).

Notes and further reading. Moore (1979) presents a finantial application of finding roots of (real and) interval polynomials – the computation of the internal rate-of-return of investments and its sensitivity analysis.

6.6 Symbolic and automatic differentiation

For the sake of simplicity, we consider a real function $f\colon \mathbb{R} \to \mathbb{R}$ and the first derivative, but it directly extends to multivariate functions and higher order derivatives, including gradients and Hessian matrices.

Symbolic differentiation

Given an arithmetic expression for $f(x)$ using standard arithmetic and basic functions, we easily compute $f'(x)$ recursively by using the well-known rules

$$\begin{aligned}(g(x) \pm h(x))' &= g'(x) \pm h'(x),\\ (g(x) \cdot h(x))' &= g'(x) \cdot h(x) + g(x) \cdot h'(x),\\ \left(\frac{g(x)}{h(x)}\right)' &= \frac{g'(x) \cdot h(x) - g(x) \cdot h'(x)}{h^2(x)},\end{aligned}$$

and the knowledge of derivatives of basic functions such as

$$\begin{aligned}c' &= 0, \quad (c \text{ is a constant})\\ x' &= 1,\\ \exp'(x) &= \exp'(x),\\ \sin'(x) &= \cos(x).\end{aligned}$$

Computing the expression for $f'(x)$ in this way is called *symbolic differentiation*. Its advantage is that we have an explicit expression for $f'(x)$ and can symbolically manipulate with it to equivalent forms. On the other hand, evaluation of the derivative is superfluously costly since we have to parse the expression twice – first to determine the expression for the derivative, and second, to evaluate the derivative.

Automatic differentiation

Automatic differentiation aims at computing the value of the derivative (but not the expression for it) more efficiently with respect to time. The idea is to parse the expression for $f(x)$ only once, and to recursively compute the pair $(f(x), f'(x))$. We adapt the differentiation arithmetic as follows

$$\begin{aligned}(g, g') \pm (h, h') &= (g \pm h,\ g' \pm h'),\\ (g, g') \cdot (h, h') &= (g \cdot h,\ g' \cdot h + g \cdot h'),\\ (g, g') \,/\, (h, h') &= (g \,/\, h,\ (g' \cdot h - g \cdot h') \cdot h^{-2}),\end{aligned}$$

and for basic functions, the pair $(f(x), f'(x))$ is analogously determined along the rules

$$\begin{aligned}c &\mapsto (c, 0),\\ x &\mapsto (x, 1),\\ \exp(x) &\mapsto (\exp(x), \exp(x)),\\ \sin(x) &\mapsto (\sin(x), \cos(x)).\end{aligned}$$

Example 6.27. Let

$$f(x) = \frac{3x+4}{x^2+1}, \quad x_0 \coloneqq 2.$$

Its symbolic derivative is

$$f'(x) = \frac{3(x^2+1) - (3x+4)2x}{(x^2+1)^2}, \qquad f'(x_0) = -1,$$

whereas the automatic derivative computation reads

$$(f(x_0), f'(x_0)) = \frac{(3,0)\cdot(x_0,1)+(4,0)}{(x_0^2,2x_0)+(1,0)} = \frac{(10,3)}{(5,4)} = (2,-1). \qquad \square$$

Interval domain

When computing an enclosure to $f'(\boldsymbol{x})$, we can adapt both the symbolic and automatic differentiation. Usually, the automatic one is implemented (e.g., in INTLAB). In contrast, the symbolic one has a possibility of simplifying the resulting expression and making the overestimation smaller. However, symbolic manipulation of expressions to make them more convenient for interval evaluation is a big challenge.

Example 6.28. Consider the function

$$f(x) = \frac{x+1}{x+2},$$

whose derivative reads

$$f'(x) = \frac{(x+2)-(x+1)}{(x+2)^2}.$$

The direct evaluation on the interval $\boldsymbol{x} = [2,3]$ yields the same enclosure on the derivative as the automatic differentiation

$$f'(\boldsymbol{x}) \subseteq \frac{(\boldsymbol{x}+2)-(\boldsymbol{x}+1)}{(\boldsymbol{x}+2)^2} = [0, 0.125].$$

However, simplifying the expression for the derivative to $f'(x) = (x+2)^{-2}$, we obtain the exact range of the derivative $f'(\boldsymbol{x}) = (\boldsymbol{x}+2)^{-2} = [0.04, 0.0625]$. $\square$

Notes and further reading. For more details on automatic differentiation see, e.g., Griewank and Corliss (1991); Kulisch (2001, 2013); Rall (1981). Backward (reverse) mode of automatic differentiation, enabling a more efficient evaluation when the number of variables is higher, is presented in Mayer (2017), for instance.

Chapter 7

Other topics

7.1 Generalized interval arithmetic

Interval arithmetic can be generalized in diverse ways. Already in Section 2.3, we extended it to zero containing interval division. Initial contribution in this topic is due to Kahan (1968b) (see also Laveuve (1975)), who extended intervals to comprise also the infinite ones, and introduced the so-called exterior intervals of type $[-\infty, a] \cup [b, \infty]$, which enabled to divide zero containing intervals. Kahan arithmetic was further adapted by Novoa and Ratz for use in nonlinear systems and optimization.

(interval.louisiana.edu/455/interval_arithmetic.ps)

7.1.1 Kaucher interval arithmetic

Kaucher (1980) generalized interval arithmetic by relaxing the condition $\underline{x} \leq \overline{x}$, $\boldsymbol{x} \in \mathbb{IR}$, and making the arithmetic algebraically complete (meaning to satisfy basic algebraic properties). Besides the algebraic advantages, Kaucher arithmetic also provides a new perspective of basic interval problems and brings simplifications in many results. Most of the results presented in this section come from Shary (2002).

Kaucher intervals are just pairs of real numbers $\boldsymbol{x} \coloneqq [\underline{x}, \overline{x}]$ without the condition $\underline{x} \leq \overline{x}$ and without the classical meaning. Proper intervals are those with $\underline{x} \leq \overline{x}$, while the others are called improper. Dualization operation interchanges the endpoints of intervals. Formally, dual: $[\underline{x}, \overline{x}] \to [\overline{x}, \underline{x}]$, thus making proper intervals from the improper ones.

Addition is defined as for the standard interval arithmetic

$$\boldsymbol{x} + \boldsymbol{y} \coloneqq [\underline{x} + \underline{y}, \overline{x} + \overline{y}].$$

Without the restrictive condition $\underline{x} \leq \overline{x}$, each Kaucher interval has its opposite interval (additive inverse)

$$\ominus \boldsymbol{x} = [-\underline{x}, -\overline{x}],$$

and that is why Kaucher arithmetic becomes an Abelian (commutative) group.

Multiplication is a little bit more complicated, one of several equivalent definitions reads

$$\boldsymbol{x}\boldsymbol{y} = \left[\max\left\{\underline{x}^+\underline{y}^+, \overline{x}^-\overline{y}^-\right\} - \max\left\{\overline{x}^+\underline{y}^-, \underline{x}^-\overline{y}^+\right\}, \max\left\{\overline{x}^+\overline{y}^+, \underline{x}^-\underline{y}^-\right\} - \max\left\{\underline{x}^+\overline{y}^-, \overline{x}^-\underline{y}^+\right\}\right],$$

Nevertheless, multiplication by a real number still reduces to the classical case

$$x\boldsymbol{y} = \begin{cases} [x\underline{y}, x\overline{y}] & \text{if } x \geq 0, \\ [x\overline{y}, x\underline{y}] & \text{otherwise.} \end{cases}$$

Multiplication is commutative, associative, and there is the identity element $[1, 1]$. Nevertheless, inverses need not exist. For instance, $[-2, 3]$ has no inverse, while $[1, 2]^{-1} = [1, \frac{1}{2}]$.

While proper intervals follow sub-distributivity law, improper intervals follow super-distributivity. Therefore, the classical enclosure interpretation of interval arithmetic makes no sense now. Nevertheless, Kaucher arithmetic has a different, but not less meaningless, interpretation.

Minimax interpretation. The remarkable property of Kaucher arithmetic is the minimax characterization of the resulting interval endpoints. Let $\circ$ stand for addition or multiplication. Then

$$\boldsymbol{x} \circ \boldsymbol{y} = \begin{cases} [\min_{x\in\boldsymbol{x}} \min_{y\in\boldsymbol{y}} x \circ y,\ \max_{x\in\boldsymbol{x}} \max_{y\in\boldsymbol{y}} x \circ y] & \text{if } \boldsymbol{x}, \boldsymbol{y} \text{ are proper,} \\ [\min_{x\in\boldsymbol{x}} \max_{y\in\boldsymbol{y}} x \circ y,\ \max_{x\in\boldsymbol{x}} \min_{y\in\boldsymbol{y}} x \circ y] & \text{if } \boldsymbol{x} \text{ is proper and } \boldsymbol{y} \text{ improper,} \\ [\max_{x\in\boldsymbol{x}} \min_{y\in\boldsymbol{y}} x \circ y,\ \min_{x\in\boldsymbol{x}} \max_{y\in\boldsymbol{y}} x \circ y] & \text{if } \boldsymbol{x} \text{ is improper and } \boldsymbol{y} \text{ proper,} \\ [\max_{x\in\boldsymbol{x}} \max_{y\in\boldsymbol{y}} x \circ y,\ \min_{x\in\boldsymbol{x}} \min_{y\in\boldsymbol{y}} x \circ y] & \text{if } \boldsymbol{x}, \boldsymbol{y} \text{ are improper.} \end{cases}$$

Thus, concerning the left endpoint, "min" is associated with proper and "max" with improper intervals.

The minimax property can be used to compute lower and upper bounds on the minimax problems. Let $\boldsymbol{x} \in \mathbb{IR}^m$, $\boldsymbol{y} \in \mathbb{IR}^n$ be proper intervals, and let $f(x, y)\colon \mathbb{R}^{m+n} \to \mathbb{R}$ be a real function given by an arithmetic expression. If the expression has at most one occurrence of the variables y_j, $j = 1, \ldots, n$, and to the first power only, then

$$\left[\min_{x\in\boldsymbol{x}} \max_{y\in\boldsymbol{y}} f(x, y),\ \max_{x\in\boldsymbol{x}} \min_{y\in\boldsymbol{y}} f(x, y)\right] \subseteq f(\boldsymbol{x}, \mathrm{dual}(\boldsymbol{y})).$$

Similarly, if the expression has at most one occurrence of the variables x_i, $i = 1, \ldots, m$, and to the first power only, then

$$\left[\max_{y\in\boldsymbol{y}} \min_{x\in\boldsymbol{x}} f(x, y),\ \min_{y\in\boldsymbol{y}} \max_{x\in\boldsymbol{x}} f(x, y)\right] \supseteq f(\boldsymbol{x}, \mathrm{dual}(\boldsymbol{y})).$$

Modal interpretation. A consequence of the minimax characterization is the modal interpretation of Kaucher arithmetic. Proper intervals can be interpreted as those associated with existential quantifiers, and improper intervals are interpretable as their duals associated with universal quantifiers. From this perspective, the result of an arithmetic operation consists of numbers fulfilling the quantification, sorted from universal to existential quantifiers. Let $\circ$ be an addition or multiplication, and for the sake of simplicity, suppose that $\boldsymbol{x} \circ \boldsymbol{y}$ is proper. Then for proper intervals $\boldsymbol{x}, \boldsymbol{y}$ we have

$$\boldsymbol{x} \circ \boldsymbol{y} = \{z \in \mathbb{R};\ \exists x \in \boldsymbol{x}\, \exists y \in \boldsymbol{y} : z = x \circ y\},$$

in accordance with the definition of standard interval arithmetic. For $\boldsymbol{x}$ improper and $\boldsymbol{y}$ proper, we have

$$\boldsymbol{x} \circ \boldsymbol{y} = \{z \in \mathbb{R};\ \forall x \in \mathrm{dual}(\boldsymbol{x})\, \exists y \in \boldsymbol{y} : z = x \circ y\},$$

and $\boldsymbol{x}$ proper and $\boldsymbol{y}$ improper, we have

$$\boldsymbol{x} \circ \boldsymbol{y} = \{z \in \mathbb{R};\ \forall y \in \mathrm{dual}(\boldsymbol{y})\, \exists x \in \boldsymbol{x} : z = x \circ y\}.$$

Example 7.1. The results of Kaucher multiplications:

$$[1, 2]\,[4, 3] = [4, 6],$$
$$[-2, 4]\,[3, -1] = 0.$$

□

Inner and outer enclosures. Kaucher arithmetic enables also to characterize the solutions of *AE* solution set Σ_{AE} from Section 3.10. We have $x \in \Sigma_{AE}$ if and only if

$$(\boldsymbol{A}^{\forall} + \mathrm{dual}(\boldsymbol{A}^{\exists}))x \subseteq \mathrm{dual}(\boldsymbol{b}^{\forall}) + \boldsymbol{b}^{\exists}. \tag{7.1}$$

Based on this result, it is easy to prove the following (cf. Shary (2002)).

Theorem 7.2 (Shary, 1995c)**.** *If* $(\boldsymbol{A}^{\forall} + \mathrm{dual}(\boldsymbol{A}^{\exists}))\boldsymbol{x} = \mathrm{dual}(\boldsymbol{b}^{\forall}) + \boldsymbol{b}^{\exists}$ *then* $\boldsymbol{x} \subseteq \Sigma_{AE}$.

In other words, any formal solution to the interval system $(\boldsymbol{A}^{\forall} + \mathrm{dual}(\boldsymbol{A}^{\exists}))x = \mathrm{dual}(\boldsymbol{b}^{\forall}) + \boldsymbol{b}^{\exists}$ gives an inner enclosure to the solution set. The opposite implication does not hold in general. For example, $[-1, 1]x = [-1, 2]$ has no formal solution, but the tolerable solution set is $[-1, 1]$.

Shary also proved that a formal solution maximal with respect to inclusion yields a maximal inner solution. This is particularly the case when the formal solution is unique. Notice that it is NP-hard to compute formal solutions, however, there are methods that work well in practice.

Based on (7.1), outer enclosures of Σ_{AE} can be developed, too, but we will not discuss it in more detail.

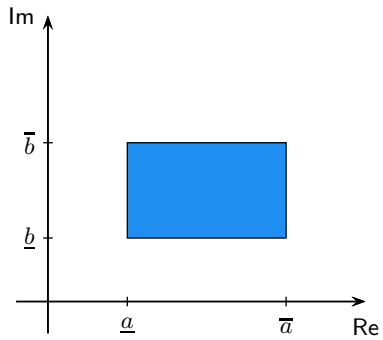


Figure 7.1: Rectangular form $\boldsymbol{a} + \boldsymbol{b}i$.

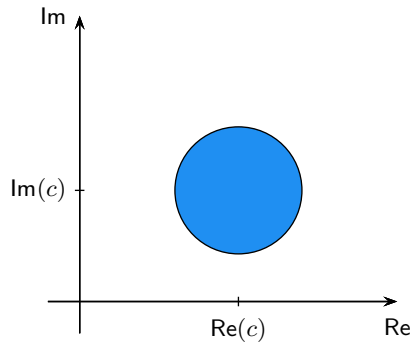


Figure 7.2: Circular form $\langle c, r\rangle$.

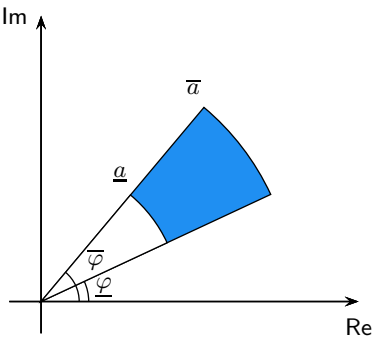


Figure 7.3: Polar form $\boldsymbol{a}\,\mathrm{cis}(\boldsymbol{\varphi})$.

7.1.2 Complex interval arithmetic

Complex intervals can basically be defined by three ways; see Figures 7.1–7.3.

- *rectangular* complex intervals $\boldsymbol{a} + \boldsymbol{b}i := \{a + ib;\ a \in \boldsymbol{a},\ b \in \boldsymbol{b}\}$, where $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{IR}$,
- *circular* complex intervals $\langle c, r\rangle := \{z \in \mathbb{C};\ |c - z| \le r\}$, where $c \in \mathbb{C}$ and $r \ge 0$,
- *polar* complex intervals $\boldsymbol{a}\,\mathrm{cis}(\boldsymbol{\varphi}) := \{a(\cos\varphi + i\sin\varphi);\ a \in \boldsymbol{a},\ \varphi \in \boldsymbol{\varphi}\}$, where $\boldsymbol{a}, \boldsymbol{\varphi} \in \mathbb{IR}$, $\underline{a} \ge 0$.

Once the complex interval form is specified, $\mathbb{IC}$ denotes the set of the corresponding complex intervals.

For all three concepts, we need to define complex interval arithmetic satisfying the fundamental enclosing property. That is, for all complex intervals $\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{v}}$ and any arithmetic operation $\circ \in \{+, -, \cdot, /\}$,

$$\{u \circ v;\ u \in \tilde{\boldsymbol{u}},\ v \in \tilde{\boldsymbol{v}}\} \subseteq \tilde{\boldsymbol{u}} \circ \tilde{\boldsymbol{v}}.$$

The rectangular form. The rectangular form was pioneered by Boche (1966); Alefeld (1968), and the corresponding complex interval arithmetic reads

$$\begin{aligned}
(\boldsymbol{a} + \boldsymbol{b}i) \pm (\boldsymbol{c} + \boldsymbol{d}i) &= \boldsymbol{a} \pm \boldsymbol{c} + (\boldsymbol{b} \pm \boldsymbol{d})i \\
(\boldsymbol{a} + \boldsymbol{b}i) \cdot (\boldsymbol{c} + \boldsymbol{d}i) &= \boldsymbol{ac} - \boldsymbol{bd} + (\boldsymbol{ad} + \boldsymbol{bc})i \\
(\boldsymbol{a} + \boldsymbol{b}i)/(\boldsymbol{c} + \boldsymbol{d}i) &= \frac{\boldsymbol{ac} + \boldsymbol{bd}}{\boldsymbol{c}^2 + \boldsymbol{d}^2} + \frac{\boldsymbol{bc} - \boldsymbol{ad}}{\boldsymbol{c}^2 + \boldsymbol{d}^2}i, \quad 0 \notin \boldsymbol{c}^2 + \boldsymbol{d}^2.
\end{aligned}$$

Obviously by the properties of real interval arithmetic, addition, subtraction, and multiplication by reals yield optimal results (Alefeld, 1968; Nickel, 1980). Multiplication produces the tightest interval enclosure, i.e., the interval hull; see Figure 7.4.

Observation 7.3. *For any complex intervals $\tilde{\boldsymbol{u}}$ and $\tilde{\boldsymbol{v}}$ in the rectangular form*

(1) $\{u \pm v;\ u \in \tilde{\boldsymbol{u}},\ v \in \tilde{\boldsymbol{v}}\} = \tilde{\boldsymbol{u}} \pm \tilde{\boldsymbol{v}}$.

(2) $\square\{u \cdot v;\ u \in \tilde{\boldsymbol{u}},\ v \in \tilde{\boldsymbol{v}}\} = \tilde{\boldsymbol{u}} \cdot \tilde{\boldsymbol{v}}$.

On the other hand, the other operations are overestimating the actual image (which is not a rectangular box); see Figure 7.5. In particular, the division does not produce the interval hull of the image in general; see Figure 7.5, where $2/(1+[-1,1]i) = [1,2]+[-2,2]i$ highly overestimates the imaginary part. Even in the simple example $\boldsymbol{a}+\boldsymbol{b}i = 1+0i$, $\boldsymbol{c}+\boldsymbol{d}i = [1,2]+0i$ the rectangular form overestimates a lot, producing the interval $[0.25, 2]$, while real interval arithmetic produces $[0.5, 1]$. Tighter results are obtained if the division is defined as $\tilde{\boldsymbol{u}}/\tilde{\boldsymbol{v}} = \tilde{\boldsymbol{u}} \cdot (1/\tilde{\boldsymbol{v}})$, where the inverse value value is computed exactly $1/\tilde{\boldsymbol{v}} = \square\{1/v;\ v \in \tilde{\boldsymbol{v}}\}$. However, the computation of the interval hull is expensive (Rokne and Lancaster, 1971).

Inclusion isotonicity (Boche, 1966) is still fulfilled here simply because of the inclusion isotonicity of real interval arithmetic.

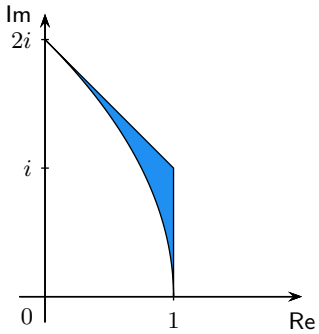


Figure 7.4: The exact image of $(1 + [0,1]i) \cdot (1 + [0,1]i)$. Complex interval arithmetic yields $[0,1] + [0,2]i$.

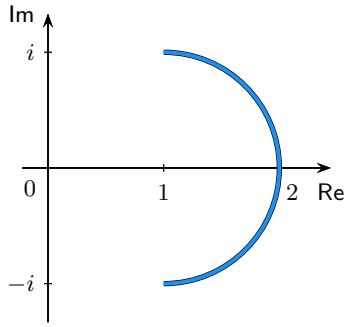


Figure 7.5: The exact image of $2/(1 + [-1,1]i)$. Complex interval arithmetic yields $[1,2] + [-2,2]i$.

The circular form. The circular form was first investigated by Henrici (1971); see also Gargantini and Henrici (1972); Hauenschild (1974). Its arithmetic is defined as follows. Let complex intervals $\tilde{\boldsymbol{u}} = \langle c_u, r_u \rangle$ and $\tilde{\boldsymbol{v}} = \langle c_v, r_v \rangle$,

$$\begin{aligned} \langle c_u, r_u \rangle \pm \langle c_v, r_v \rangle &= \langle c_u \pm c_v, r_u + r_v \rangle \\ \langle c_u, r_u \rangle \cdot \langle c_v, r_v \rangle &= \langle c_u c_v,\ |c_u| r_v + |c_v| r_u + r_u r_v \rangle \\ 1 \,/\, \langle c_u, r_u \rangle &= \big\langle \overline{c_u}/(|c_u|^2 - r_u^2),\ r_u/(|c_u|^2 - r_u^2) \big\rangle, \quad r_u < |c_u|. \end{aligned}$$

Similarly as above, addition, subtraction, multiplication by complex numbers, and inversion of a circle not containing the zero are operations yielding the true image (Alefeld and Herzberger, 1983; Mayer, 2017; Nickel, 1980). Multiplication of complex intervals $\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{v}}$ overestimates in general, but as we show below, the resulting circle is the smallest circle centered in the product of the centers of $\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{v}}$ and containing the true image.

Proposition 7.4. *For any complex intervals $\tilde{\boldsymbol{u}} = \langle c_u, r_u \rangle$ and $\tilde{\boldsymbol{v}} = \langle c_v, r_v \rangle$ in the circular form*

(1) $\{u \pm v;\ u \in \tilde{\boldsymbol{u}},\ v \in \tilde{\boldsymbol{v}}\} = \tilde{\boldsymbol{u}} \pm \tilde{\boldsymbol{v}}$,

(2) $\square\{u \cdot v;\ u \in \tilde{\boldsymbol{u}},\ v \in \tilde{\boldsymbol{v}}\} \subsetneqq \tilde{\boldsymbol{u}} \cdot \tilde{\boldsymbol{v}} = \cap\{\langle c_u c_v, r \rangle;\ uv \in \langle c_u c_v, r \rangle,\ \forall u \in \tilde{\boldsymbol{u}},\ \forall v \in \tilde{\boldsymbol{v}}\}$,

(3) $\square\{1/u;\ u \in \tilde{\boldsymbol{u}}\} \subsetneqq 1/\tilde{\boldsymbol{u}}$.

Proof.

(1) Obvious.

(2) "The inclusion." Let $u \in \tilde{\boldsymbol{u}}$ and $v \in \tilde{\boldsymbol{v}}$ in the forms $u = c_u + re^{i\alpha}$, $0 \le r \le r_u$, and $v = c_v + r'e^{i\beta}$, $0 \le r' \le r_v$. Then

$$uv = (c_u + re^{i\alpha})(c_v + r'e^{i\beta}) = c_u c_v + c_u r' e^{i\beta} + c_v r e^{i\alpha} + rr' e^{i(\alpha+\beta)})$$

So we have $uv \in \langle c_u c_v, s \rangle$ for a value s such that

$$s \le |c_u r' e^{i\beta} + c_v r e^{i\alpha} + rr' e^{i(\alpha+\beta)}| \le |c_u r' e^{i\beta}| + |c_v r e^{i\alpha}| + |rr' e^{i(\alpha+\beta)}| \le |c_u| r_v + |c_v| r_u + r_u r_v.$$

"The equation." Let $c_u = |c_u| e^{i\alpha}$ and $c_v = |c_v| e^{i\beta}$. Put $u := c_u + r_u e^{i\alpha} \in \tilde{\boldsymbol{u}}$ and $v := c_v + r_v e^{i\beta} \in \tilde{\boldsymbol{v}}$. Then

$$\begin{aligned} uv &= (c_u + r_u e^{i\alpha})(c_v + r_v e^{i\beta}) = (|c_u| + r_u) e^{i\alpha})(|c_v| + r_v) e^{i\beta} \\ &= (|c_u||c_v| + |c_u| r_v + r_u |c_v| + r_u r_v) e^{i(\alpha+\beta)} = c_u c_v + (|c_u| r_v + r_u |c_v| + r_u r_v) e^{i(\alpha+\beta)}, \end{aligned}$$

which lies on the boundary of the circle $\tilde{\boldsymbol{u}}\tilde{\boldsymbol{v}}$.

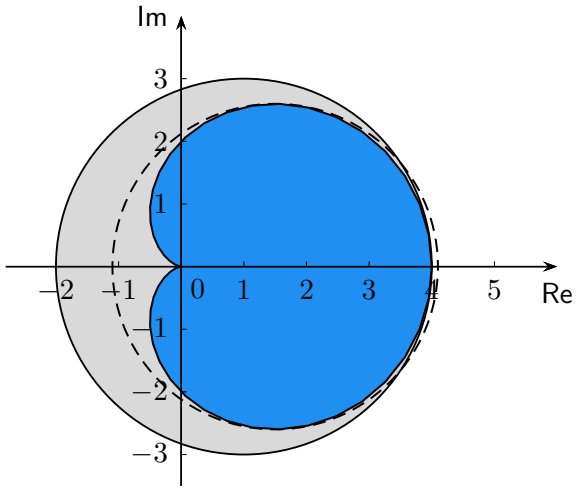


Figure 7.6: (Example 7.5) The exact image of $\langle 1,1\rangle \cdot \langle 1,1\rangle$ in blue (in the shape of a cardioid) and the result of complex interval arithmetic in gray. The smallest circle containing the true image is dashed and crosses the real axis in the right at point ≈ 4.0981.

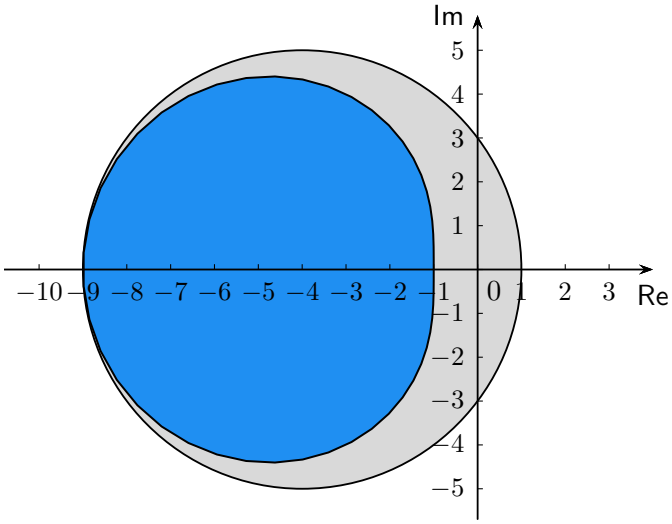


Figure 7.7: (Example 7.6) The exact image of $\langle -2,1\rangle \cdot \langle 2,1\rangle$ in blue does not contain the origin, but the result of complex interval arithmetic in gray contains the origin.

(3) Let $u \in \tilde{\boldsymbol{u}}$ in the form $u = c_u + re^{i\alpha}$, $0 \le r \le r_u$. We want to show that $|(|c_u|^2 - r_u^2)/u - \overline{c_u}| \le r_u$. Let us write

$$\frac{|c_u|^2 - r_u^2}{c_u + re^{i\alpha}} - \overline{c_u} = \frac{|c_u|^2 - r_u^2 - \overline{c_u}(c_u + re^{i\alpha})}{c_u + re^{i\alpha}} = \frac{-r_u^2 - \overline{c_u} re^{i\alpha}}{c_u + re^{i\alpha}}.$$

Its absolute value reads

$$\frac{|r_u^2 + \overline{c_u} re^{i\alpha}|}{|c_u + re^{i\alpha}|} \le \frac{-r_u^2 + |c_u| r_u}{|c_u| - r_u} = r_u. \qquad \square$$

The following example illustrates that multiplication overestimates in general.

Example 7.5 (Mayer, 2017)**.** Let $\tilde{\boldsymbol{u}} = \tilde{\boldsymbol{v}} = \langle 1,1\rangle$. Then $\tilde{\boldsymbol{u}}\tilde{\boldsymbol{v}} = \langle 1,3\rangle$, but the smallest enclosing circle is $\langle 1.5, 1.5\sqrt{3}\rangle$; see Figure 7.6. $\square$

Example 7.6 (Mayer, 2017)**.** Let $\tilde{\boldsymbol{u}} = \langle -2,1\rangle$ and $\tilde{\boldsymbol{v}} = \langle 2,1\rangle$. Then $\tilde{\boldsymbol{u}}\tilde{\boldsymbol{v}} = \langle -4,5\rangle$. So $0 \in \tilde{\boldsymbol{u}}\tilde{\boldsymbol{v}}$, even though $0 \notin \tilde{\boldsymbol{u}}$ and $0 \notin \tilde{\boldsymbol{v}}$; see Figure 7.7. $\square$

To compute optimal enclosing circles for multiplication and division is a nontrivial task, and the corresponding arithmetic turns out not to be inclusion isotone (Krier, 1973; Nickel, 1980). On the other hand, the overestimation produced by these operations is usually small. That is why circular arithmetic was implemented, e.g., in INTLAB (Rump, 1999a). One can show that the multiplication operation overestimates the optimal enclosing circle by no more than 50% of the length of its radius. Even more, by Krier (1973); Mayer (2017); Rump (1999b) we have for any complex intervals $\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{v}}$

$$\operatorname{rad}(\tilde{\boldsymbol{u}}\tilde{\boldsymbol{v}}) \le \frac{3}{2} \max\{\tfrac{1}{2}|z_1 - z_2|;\ z_1, z_2 \in \mathcal{S}\},$$

where $\mathcal{S} = \{u \cdot v;\ u \in \tilde{\boldsymbol{u}},\ v \in \tilde{\boldsymbol{v}}\}$.

Comparison. For both rectangular and circular forms, many algebraic properties of Theorem 2.5 remain valid (Alefeld and Herzberger, 1983; Djanybekov, 2006).

Theorem 7.7. *For rectangular or circular forms of complex intervals $\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{v}}$ and $\tilde{\boldsymbol{w}}$ we have*

(1) $\tilde{\boldsymbol{u}} + \tilde{\boldsymbol{v}} = \tilde{\boldsymbol{v}} + \tilde{\boldsymbol{u}}$,

(2) $(\tilde{\boldsymbol{u}} + \tilde{\boldsymbol{v}}) + \tilde{\boldsymbol{w}} = \tilde{\boldsymbol{v}} + (\tilde{\boldsymbol{u}} + \tilde{\boldsymbol{w}})$,

(3) $\tilde{\boldsymbol{u}}\tilde{\boldsymbol{v}} = \tilde{\boldsymbol{v}}\tilde{\boldsymbol{u}}$,

(4) $\tilde{\boldsymbol{u}}(\tilde{\boldsymbol{v}} + \tilde{\boldsymbol{w}}) \subseteq \tilde{\boldsymbol{u}}\tilde{\boldsymbol{v}} + \tilde{\boldsymbol{u}}\tilde{\boldsymbol{w}}$,

(5) $z(\tilde{\boldsymbol{u}} + \tilde{\boldsymbol{v}}) = z\tilde{\boldsymbol{v}} + z\tilde{\boldsymbol{u}}$, $z \in \mathbb{C}$,

(6) complex interval arithmetic is inclusion isotonic.

For circular forms of $\tilde{\boldsymbol{u}}, \tilde{\boldsymbol{v}}$ and $\tilde{\boldsymbol{w}}$ we have

(7) $(\tilde{\boldsymbol{u}}\tilde{\boldsymbol{v}})\tilde{\boldsymbol{w}} = \tilde{\boldsymbol{v}}(\tilde{\boldsymbol{u}}\tilde{\boldsymbol{w}})$.

Property (7) does not hold for rectangular forms in general. For example,

$$\begin{aligned} &[2,3]\big((1+i)(1-i)\big) = [2,3]2 = [4,6], \\ &\big([2,3](1+i)\big)(1-i) = ([2,3] + [2,3]i)(1-i) = [4,6] + [-1,1]i. \end{aligned}$$

As we have seen, none of the interval complex forms is preserved under all arithmetic operations. Indeed, Nickel (1980) proved that there is no such a form with finite description that is closed under arithmetic and contains real intervals as segments in the complex plane. Further, Kreinovich and von Gudenberg (2000) showed that the best approximation family of sets that is shift-, rotation-, and scale-invariant have piecewise linear or circular boundaries.

The polar form. In contrast to the previous forms, the polar form is closed under multiplication and division (and also exponentiation) as we have (Candau et al., 2006; Klatte and Ullrich, 1980)

$$\begin{aligned} \boldsymbol{a}\operatorname{cis}(\boldsymbol{\varphi}) \cdot \boldsymbol{b}\operatorname{cis}(\boldsymbol{\eta}) &= (\boldsymbol{ab})\operatorname{cis}(\boldsymbol{\varphi} + \boldsymbol{\eta}), \\ \boldsymbol{a}\operatorname{cis}(\boldsymbol{\varphi})/\boldsymbol{b}\operatorname{cis}(\boldsymbol{\eta}) &= (\boldsymbol{a}/\boldsymbol{b})\operatorname{cis}(\boldsymbol{\varphi} - \boldsymbol{\eta}), \quad 0 \notin \boldsymbol{b}. \end{aligned}$$

On the other hand, addition and subtraction is rather cumbersome. Nevertheless, Candau et al. (2006) proposed a method to enclose the resulting image in a minimal polar form with a reasonable cost. For the polar form, Wang et al. (2019) proposed an optimal affine approximation, which reduces the overestimation caused by dependencies. Other possibility is to use rectangular enclosure and addition / subtraction for rectangular forms.

Notes and further reading. An important application of complex interval arithmetic is the computation of complex roots of polynomials (Gargantini, 1976, 1978; Gargantini and Henrici, 1972; Hansen, 1969b; Krier, 1973; Nickel, 1975).

A generalization of the circular form to ellipsoids was considered by Kahan (1967); see references in (Nickel, 1980).

7.1.3 Affine arithmetic

7.1.4 Interval unions

Interval unions are finite sets of closed and disjoint intervals. They enable for a finer representation of enclosures and thus for more efficient algorithms. As an example of an interval union, consider

$$\mathcal{U} = [-2, -1] \cup [1, 2] \cup [3, 3].$$

The range of the function $f(x) = x^2$ on this domain is $f(\mathcal{U}) = [1, 4] \cup [9, 9]$ and its interval hull is $\Box f(\mathcal{U}) = [1, 9]$. To compare, the range of the function an the interval hull $\boldsymbol{x} = \Box\mathcal{U} = [-2, 3]$ is $f(\boldsymbol{x}) = [0, 9]$.

Arithmetic with interval union is analogous to interval arithmetic. One drawback of this representation is that the number of intervals produces during computation may increase exponentially; to handle this problem, various gap-filling strategies were proposed to keep the number of intervals reasonably large. Due to non-convexity of interval unions, we we also have to be more careful when using the mean value theorem.

Interval unions are particularly convenient when using extended interval arithmetic. In case of division, extended interval arithmetic may produce a union of two intervals. This motivates to process the union further. As a consequence, interval unions turn to be efficient in the interval Newton method (Section 8.1.1) and in solving interval systems of linear equations by the interval Gaussian elimination or Gauss–Seidel method.

Interval unions were introduced in Schichl et al. (2017), applied to solving interval systems of linear equations in Montanher et al. (2017) and overviewed in Domes et al. (2020).

Notes and further reading. There are several related concepts. *Multi-intervals* (Yakovlev, 1986) are defined as a union of closed but not necessarily disjoint intervals. They were implemented in a free software *UniCalc*; see Babichev et al. (1993) and download at

`http://www.rriai.org.ru/UniCalc/`

Another concept are *discontinuous intervals* introduced by Hyvönen (1992). This concept deals with a union of closed, open and half-open intervals. They were motivated by the propagation techniques in constraint programming (Chapter 8).

7.1.5 Dependent operations

7.2 Other kinds of enclosures

Interval enclosures are sometimes too coarse approximations and highly overestimates the true set. Thus, more subtle kinds of enclosures are in demand, too.

7.2.1 Ellipsoids

Ellipsoid enclosures for parameter identification problems with uncertain data are studied in Belforte et al. (1990); Fogel and Huang (1982); Pronzato and Walter (1994).

Ellipsoid enclosures of zonotopes were discussed by Černý (2012); Černý and Rada (2012). They adapted Goffin's polynomial algorithm to construct an enclosing Löwner–John ellipsoid with the property that if it is shrunk by the factor n (the dimension), we get an inscribed ellipse. This kind of ellipsoid was applied by Černý et al. (2013) to enclose the set of the least square solutions for overdetermined system of equations with interval right-hand side; cf. Section 3.8.1.

7.3 Integration

Enclosure of a real integral. Using interval computation for determining rigorous bounds of an integral appears already in Moore (1966). Let $f\colon \mathbb{R} \to \mathbb{R}$ be continuous and $[a, b] \in \mathbb{IR}$ be given. By the mean

value theorem there is $c \in [a,b]$ such that

$$\int_a^b f(x)\,\mathrm{d}x \subseteq f(c)(b-a).$$

This gives rise to the enclosure

$$\int_a^b f(x)\,\mathrm{d}x \subseteq f([a,b])(b-a).$$

A refinement makes to enclosure tighter. The points

$$x_k = a + \frac{k}{n}(b-a), \quad k = 0,\ldots,n$$

discretize the interval $[a,b]$ equidistantly into n sub-intervals. Now, the interval version of the Riemann sum

$$\boldsymbol{i}_n \coloneqq \frac{b-a}{n}\sum_{k=1}^{n} f([x_{k-1},x_k])$$

gives an enclosure of the integral

$$\int_a^b f(x)\,\mathrm{d}x \subseteq \boldsymbol{i}_n.$$

This simple idea is straightforwardly adapted for non-uniform discretization of the interval $[a,b]$.

Under some assumptions (e.g., when $f(x)$ is given by an arithmetic expression using only arithmetic operations) we have

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{n\to\infty} \boldsymbol{i}_n.$$

The rate of convergence of the interval extension of $f(x)$ conveys roughly to the rate of convergence of $\boldsymbol{i}_n$. Thus, if we use the natural interval extension having linear convergence (cf. Section 6.4), then by doubling n, the enclosure $\boldsymbol{i}_n$ approximately decreases the width by half.

Interval integral. Let $\boldsymbol{f}\colon \mathbb{IR} \to \mathbb{IR}$, where $\underline{f}$ and $\overline{f}$ are continuous, and let $[a,b] \in \mathbb{IR}$. We definie *the interval integral* as

$$\int_a^b \boldsymbol{f}(x)\,\mathrm{d}x = \bigcup\left\{\int_a^b f(x)\,\mathrm{d}x;\ f(x) \in \boldsymbol{f}(x) \text{ is continuous}\right\}.$$

It is easy to see (Moore, 1979) that

$$\int_a^b \boldsymbol{f}(x)\,\mathrm{d}x = \left[\int_a^b \underline{f}(x)\,\mathrm{d}x,\ \int_a^b \overline{f}(x)\,\mathrm{d}x\right].$$

It is also obvious that the interval integral is inclusion isotonic. That is, given $\boldsymbol{f},\boldsymbol{g}\colon \mathbb{IR} \to \mathbb{IR}$ such that $\boldsymbol{f}(x) \subseteq \boldsymbol{g}(x)$ for every $x \in [a,b]$, we have

$$\int_a^b \boldsymbol{f}(x)\,\mathrm{d}x \subseteq \int_a^b \boldsymbol{g}(x)\,\mathrm{d}x.$$

These properties turn out to be convenient when we enclose a real function $f(x)$ by using a higher order expansion (6.6). Then the integral of $f(x)$ is simply enclose by the integral of the corresponding interval polynomial.

Example 7.8 (Moore, 1979)**.** Let $f(x) = e^{-x^2}$ and $[a, b] = [0, 1]$. By using the enclosure (6.6), we get

$$\begin{aligned} f(x) &\in 1 - x^2 + \frac{1}{2}x^4 - \frac{1}{6}[e^{-1}, e^0]x^6 \\ &\subseteq 1 - x^2 + \frac{1}{2}x^4 - [0.0613, 0.1667]x^6 \end{aligned}$$

Therefore

$$\begin{aligned} \int_0^1 e^{-x^2}\,\mathrm{d}x &\in \int_0^1 1 - x^2 + \frac{1}{2}x^4 - [0.0613, 0.1667]x^6\,\mathrm{d}x \\ &= \left[\int_0^1 1 - x^2 + \frac{1}{2}x^4 - 0.1667x^6\,\mathrm{d}x,\ \int_0^1 1 - x^2 + \frac{1}{2}x^4 - 0.0613x^6\,\mathrm{d}x\right] \\ &\subseteq [0.7428, 0.7580]. \end{aligned}$$

□

7.4 Ordinary differential equation

Validated methods for initial values problems for ordinary differential equations (Nedialkov and Jackson, 2001)

$$y' = f(y), \quad y(t_0) = y_0,$$

compute bounds that are guaranteed to contain the solution. They are based on Taylor series (Makino and Berz, 2005) and other approaches.

Chapter 8

Constraint programming and global optimization

Constraint programming is a discipline that deals with solving a set of constraints. Constraints may involve equations, inequalities or other types of conditions. Global optimization is a slightly more general problem to find a global minimizer over a set of constraints. Thus, we have not only to find a feasible solution fulfilling the constraints, but also such of them that has a minimal value of the objective function.

Both disciplines are very hard, in fact undecidable in general. However, by using interval analysis, we can enclose all solutions or global minimizers into verified intervals. Naturally, time effort is usually very large as the dimension of the problem grows.

Complexity

Constraint solving. Some constraints can be solved efficiently. A system of linear equations $Ax = b$ is easy to solve by the Gaussian elimination, and a system of linear inequalities $Ax \leq b$ can be polynomially handled by means of linear programming.

Polynomial equations are considerably harder to solve. Let $p(x_1, \ldots, x_n)$ be a polynomial in variables $x_1, \ldots, x_n$. The equation $p(x_1, \ldots, x_n) = 0$ can be solved by Tarski's theorem (Tarski, 1951). This theorem says that the first-order theory of reals with $+, \times, =$ and $>$ allows quantifier elimination. Hence we can decide on solvability by elimination of the formula $\exists x_1 \ldots \exists x_n : p(x_1, \ldots, x_n)$, and the roots can be found by binary search to a given accuracy. A serious drawback is that the known methods for quantifier elimination have double exponential time complexity.

Even worse, constraint solving is undecidable in general. That is, there is provably no algorithm to solve it. Recall that the tenth Hilbert problem was to find an algorithm to determine whether a given polynomial Diophantine equation with integer coefficients has an integer solution. It was shown by Matiyasevich (1970) that this problem is undecidable. Undecidability holds also true for the continuous case as long as we admit equations with the sine function (Zhu, 2006). Clearly, a Diophantine equation $p(x_1, \ldots, x_n) = 0$ has an integer solution if and only if the equation

$$p(x_1, \ldots, x_n)^2 + \sum_{i=1}^{n} \sin^2(\pi x_i) = 0 \tag{8.1}$$

has a real solution.

Global optimization. Similar results hold for global optimization, too. Consider an optimization problem

$$\min \ f(x) \ \text{ subject to } \ c(x),$$

where $f(x)$ is the objective function and $c(x)$ the constraints in the form of polynomial equations, then $a \in \mathbb{R}$ is the optimal value if and only if

$$\exists x(a = f(x) \wedge c(x) \wedge \forall y(\neg c(x) \vee f(x) \geq f(y))).$$

Moreover, the optimal value a can be calculated by a suitable elimination and transformation to a disjunctive normal form. On the other hand, by minimizing the left-hand side function of (8.1) on $\mathbb{R}^n$ we can decide on solvability of (8.1) as the zero optimal value shows solvability, while the positive optimal value unsolvability. Thus, global optimization is undecidable in general, even with no constraints.

Range of a function. The above result implies that computing the range $f(\boldsymbol{x})$ of a real function $f\colon \mathbb{R}^n \to \mathbb{R}$ on an interval domain $\boldsymbol{x} \mathbb{IR}^n$ is undecidable as well. The global optimization problem

$$\min\ f(x)\ \text{ subject to }\ x \in \mathbb{R}^n$$

is transformed to

$$\min\ g(y)\ \text{ subject to }\ y \in (-\tfrac{\pi}{2}, \tfrac{\pi}{2})^n$$

by using the substitution $x_i := \tan(y_i)$, $i = 1, \ldots, n$. Defining $\tan(\pm\frac{\pi}{2}) := \frac{1}{2}$, for example, we obtain closed intervals and the optimization problem reads

$$\min\ g(y)\ \text{ subject to }\ y \in \boldsymbol{y} := [-\tfrac{\pi}{2}, \tfrac{\pi}{2}]^n.$$

The optimal value is then simply the left endpoint $\underline{f(\boldsymbol{y})}$.

8.1 Systems of nonlinear equations

First, we consider a basic problem of solving n nonlinear equations with respect to n variables. Let $\boldsymbol{x} \in \mathbb{IR}^n$ and let $f\colon \mathbb{R}^n \to \mathbb{R}^n$ be differentiable in each component on an open superset of $\boldsymbol{x}$. Our aim is to find all roots of f on $\boldsymbol{x}$, that is, all $x \in \boldsymbol{x}$ such that $f(x) = 0$. Since we cannot expect determining the roots exactly, we focus on calculation of their tight enclosures.

A bisection method. A naive approach is a *bisection method*. We check if $0 \in f_i(\boldsymbol{x})$ for all $i = 1, \ldots, n$. If it is not the case, there is no root of f in $\boldsymbol{x}$ and we stop. Otherwise, we split the box $\boldsymbol{x}$ into two or more sub-boxes, and repeat the procedure recursively. This process produces sequences of nested sub-boxes containing the possible roots, however, on account of huge time complexity. Computing the range $f_i(\boldsymbol{x})$ is difficult, so the condition $0 \in f_i(\boldsymbol{x})$ is usually replaced by the weaker one $0 \in \boldsymbol{f}_i(\boldsymbol{x})$, where $\boldsymbol{f}_i(\boldsymbol{x})$ is an interval extension of $f_i(\boldsymbol{x})$.

The above branch & prune approach approach is the basis of many interval methods; in Section 8.2, we discuss it more in detail. However, in order to avoid the tremendous time complexity, we try not to split boxes too frequently. One way to do it is to iteratively shrink the enclosing boxes by an appropriate method that do not cut off any root.

A fixed point approach. Our problem is equivalent to the problem of finding a fixed point of a function, which is a more suitable problem for designing iterative methods. Define $g(x) := x + Cf(x)$, where $C \in \mathbb{R}^{n\times n}$ is nonsingular. Now, $f(x) = 0$ if and only if $g(x) = x$. Thus, we reduced the problem of finding roots of f to finding fixed points of g. By Brouwer's Theorem 1.26, we have the following sufficient condition on the existence of a fixed point.

Theorem 8.1. *If $g(\boldsymbol{x}) \subseteq \boldsymbol{x}$, then g has a fixed point in $\boldsymbol{x}$.*

Interval operators. Iterative methods for finding all fixed points of $g(x)$ on $\boldsymbol{x}$ are usually based on an interval operator $\mathcal{O}\colon \mathbb{IR}^n \to \mathbb{IR}^n$ that must have the property that each fixed point in $\boldsymbol{x}$ is included also in $\mathcal{O}(\boldsymbol{x})$ (cf. the linear case in Section 3.4). That is, we cannot miss any fixed point in $\boldsymbol{x}$. As a consequence, if $\mathcal{O}(\boldsymbol{x}) \cap \boldsymbol{x} = \emptyset$, there is no fixed point in $\boldsymbol{x}$.

The first idea that comes into mind is to set $\mathcal{O}(\boldsymbol{x}) := g(\boldsymbol{x})$. Such an operator, however, would not have many nice properties. In the sequel, we describe two widely used iterative methods for solving nonlinear equations, namely the interval Newton and the Krawczyk method.

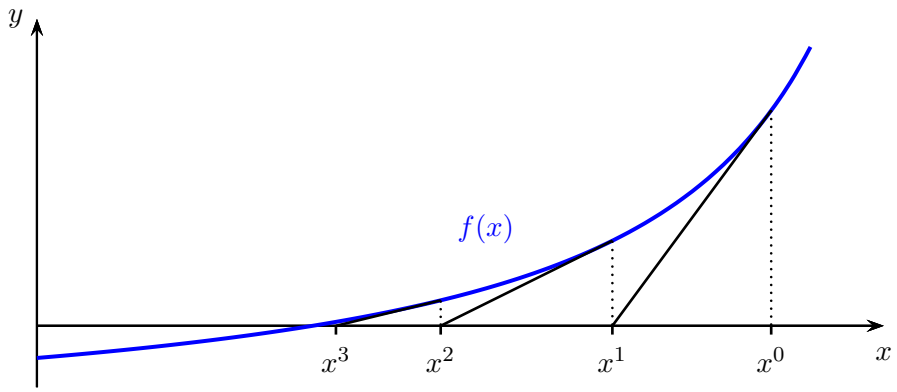


Figure 8.1: An illustration of the classical (real) Newton method.

8.1.1 Interval Newton method

Newton method is a traditional method of numerical analysis to approximately find the roots of a function. It outperforms Newton method in many points: It finds guaranteed enclosures of roots, it has better convergence and time properties, and it can rigorously prove that a box contains a unique root.

One-dimensional case

The real Newton method – a reminder. The Newton method is the classical method for finding a root of a function. Let $f\colon \mathbb{R} \to \mathbb{R}$ be differentiable and let $x^0 \in \mathbb{R}$ be an initial point. Construct a tangent line to $f(x)$ at point x^0. The line has description $\ell(x) = f(x^0) + f'(x^0)(x - x^0)$. Now, determine the crossing point of the line and the x-axis by solving the equation $\ell(x) = 0$. This yields the solution $x^1 := x^0 - f(x^0)/f'(x^0)$. Point x^1 is the starting point of the next iteration and we produce a sequence of points $x^0, x^1, x^2, \ldots$. It converges under certain assumptions. The basic idea of the Newton method is illustrated in Figure 8.1.

The interval Newton method. Let $x^0 \in \boldsymbol{x}$ be any point and x^* a root. By The mean value theorem, $f(x^*) - f(x^0) = f'(c)(x^* - x^0)$ for some $c \in \boldsymbol{x}$. Since $f(x^*) = 0$, we have $-f(x^0) = f'(c)(x^* - x^0)$. Suppose that $f'(c) \neq 0$ (otherwise x^0 is a root), whence $x^1 = x^0 - f'(c)^{-1} f(x^0)$. Therefore,

$$x^* \in x^0 - f'(\boldsymbol{x})^{-1} f(x^0).$$

holds for any root x^*. This justifies the interval Newton operator

$$N(\boldsymbol{x}, x^0) := x^0 - f'(\boldsymbol{x})^{-1} f(x^0).$$

The interval Newton method then works as follows: Put $k := 0$, $\boldsymbol{x}^0 := \boldsymbol{x}$, and repeat the iterations

1: choose $x^k \in \boldsymbol{x}^k$,
2: $\boldsymbol{x}^{k+1} := \boldsymbol{x}^k \cap N(\boldsymbol{x}^k, x^k)$,
3: $k := k + 1$.

The stopping criteria are discussed in the multi-dimensional case, as well as important theoretical and algorithmic properties. Now, we mention applicability of extended interval arithmetic, analyse the order of convergence and present a few of examples.

Extended interval arithmetic. Under some assumptions, we may incorporate extended interval arithmetic (see Section 2.3) to the interval Newton method. Suppose that x^* is a root and x^0 is not a root. As above, we derive that $f(x^*) - f(x^0) = f'(c)(x^* - x^0)$ for some $c \in \boldsymbol{x}$. Since $f(x^*) \neq f(x^0)$, the derivative $f'(c)$ is not zero. Thus, even though we calculate an enclosure of $f'(\boldsymbol{x})$ and it contains the zero, so we can omit the zero from the interval and compute $N(\boldsymbol{x}^k, x^k) := x^k - f'(\boldsymbol{x}^k)^{-1} f(x^k)$ by means of extended interval arithmetic.

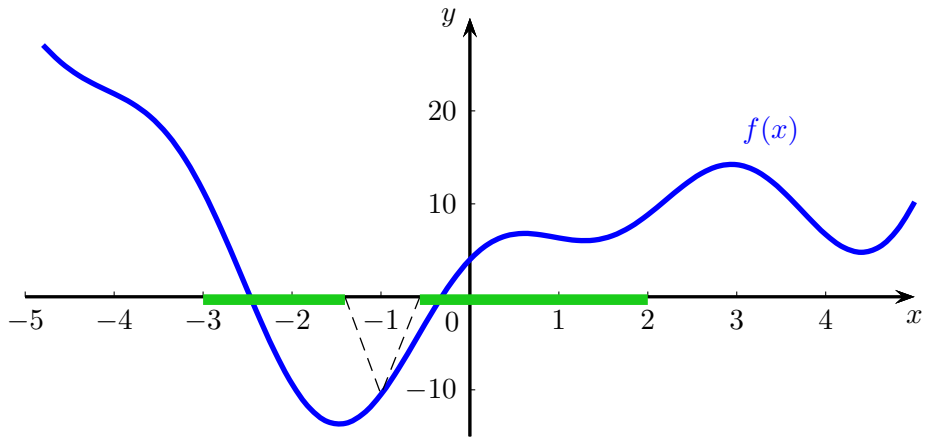


Figure 8.2: (Example 8.2) The interval Newton method with extended interval arithmetic.

Example 8.2. Let

$$f(x) = 10\sin(x) + 5\cos(2x) + x^2 - 1, \quad \boldsymbol{x} = [-3, 2],$$

and choose $x^0 := -1$. First, we compute $f'(\boldsymbol{x}) \subseteq [-25.9, 24]$. Since $0 \in f'(\boldsymbol{x})$, we cannot proceed by using standard interval arithmetic. However, using extended interval arithmetic, one iteration of the interval Newton method draws

$$N(\boldsymbol{x}, x^0) = x^0 - \frac{f(x^0)}{f'(\boldsymbol{x})} \subseteq -1 - \frac{[-10.4955, -10.4954]}{[-25.9, 24]} \subseteq (-\infty, -1.4052] \cup [-0.5626, \infty).$$

Now,

$$N(\boldsymbol{x}, x^0) \cap \boldsymbol{x} = [-3, -1.4052] \cup [-0.5626, 2].$$

This allows us to split the interval domain into two sub-intervals, $\boldsymbol{x}' := [-3, -1.4052]$ and $\boldsymbol{x}'' := [-0.5626, 2]$, and to seek for the roots of $f(x)$ recursively therein. Anyway, in the actual step, we are able to remove the set $(-1.4052, -0.5626)$ from the domain $\boldsymbol{x}$; see Figure 8.2. □

Convergence. Suppose that $0 \notin \boldsymbol{f}'(\boldsymbol{x}^k)$ for the interval extension $\boldsymbol{f}'$ of f' used. The interval Newton iteration reads

$$\boldsymbol{x}^{k+1} := \left(x^k - f(x^k)/\boldsymbol{f}'(\boldsymbol{x}^k)\right) \cap \boldsymbol{x}^k.$$

It is easy to see the linear convergence $\operatorname{rad}(\boldsymbol{x}^{k+1}) \leq \frac{1}{2}\operatorname{rad}(\boldsymbol{x}^k)$ provided $x^k \in \boldsymbol{x}^k$ is chosen as the center $x^k = \operatorname{mid}(\boldsymbol{x}^k)$. The assumption $0 \notin f'(\boldsymbol{x})$ implies that $\boldsymbol{f}'(\boldsymbol{x}^k)^{-1} f(x^k)$ lies either in $[0, \infty)$ or in $(-\infty, 0]$, so $N(\boldsymbol{x}^k, x^k) := x^k - \boldsymbol{f}'(\boldsymbol{x}^k)^{-1} f(x^k)$ cuts off at least one half of $\boldsymbol{x}^k$ when making intersection $\boldsymbol{x}^k \cap N(\boldsymbol{x}^k, x^k)$.

Another linear convergence result is the following; the speed of convergence depends on the range of derivatives.

Theorem 8.3. *Denote* $\boldsymbol{d} := \boldsymbol{f}'(\boldsymbol{x}^k)$. *If* $\boldsymbol{d} > 0$, *then*

$$\operatorname{rad}(\boldsymbol{x}^{k+1}) \leq \left(1 - \underline{d}/\overline{d}\right) \operatorname{rad}(\boldsymbol{x}^k).$$

Proof. Denote $\boldsymbol{y} := \boldsymbol{x}^{k+1}$ and $\boldsymbol{x} := \boldsymbol{x}^k$. Without loss of generality, assume that $f(x^k) \geq 0$. If $f(x^k) \leq 2x^\Delta \underline{d}$, then

$$\begin{aligned} 2y^\Delta &= \overline{y} - \underline{y} = x^k - f(x^k)/\overline{d} - \underline{y} \\ &\leq x^k - f(x^k)/\overline{d} - x^k + f(x^k)/\underline{d} = f(x^k)(1 - \underline{d}/\overline{d})/\underline{d} \\ &\leq 2x^\Delta(1 - \underline{d}/\overline{d}). \end{aligned}$$

Otherwise, if $f(x^k) > 2x^\Delta \underline{d}$, then

$$\begin{aligned} 2y^\Delta &= \overline{y} - \underline{y} = x^k - f(x^k)/\overline{d} - \underline{y} = (x^k - \underline{y}) - f(x^k)/\overline{d} \\ &\leq 2x^\Delta - 2x^\Delta \underline{d}/\overline{d} = 2x^\Delta(1 - \underline{d}/\overline{d}). \end{aligned}$$

□

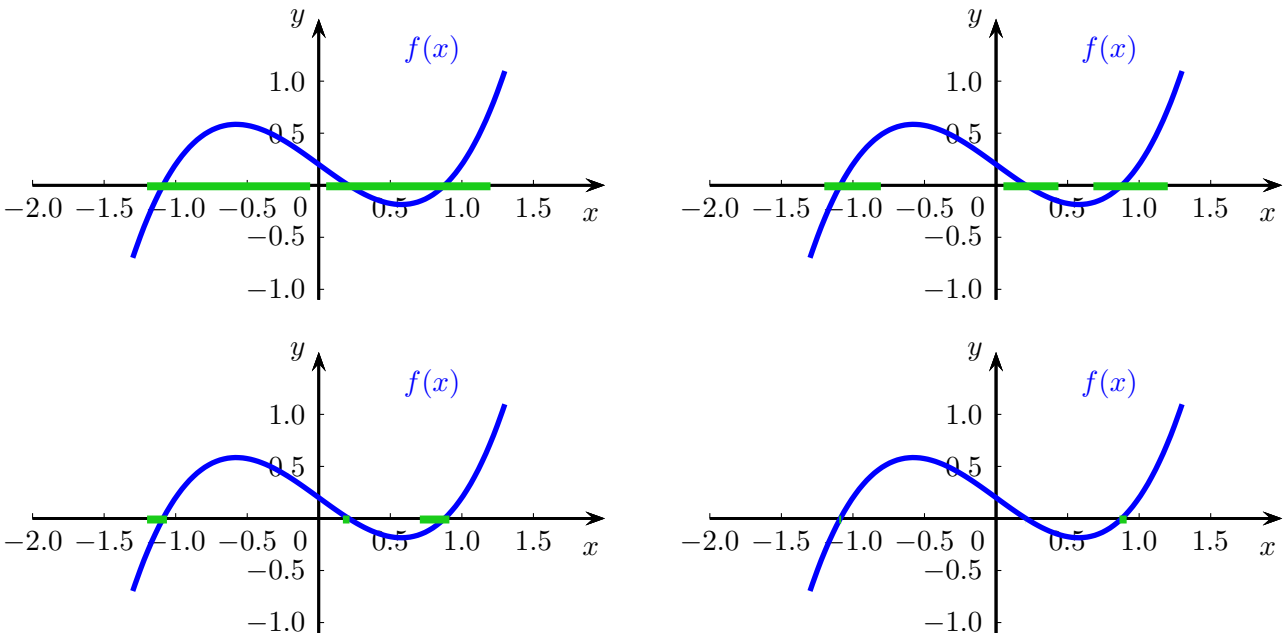


Figure 8.3: (Example 8.5) Four iterations of the interval Newton method with extended interval arithmetic.

Below, we also show the quadratic convergence of the interval Newton method (the same as for the classical Newton method).

Theorem 8.4. *Suppose that the derivative $f'(x)$ with some $\boldsymbol{u} \in \mathbb{IR}$ satisfies Property L (page 118), its enclosure is calculated by the natural extension and does not contain zero. Then there is $\gamma > 0$ such that*

$$\operatorname{rad}(\boldsymbol{x}^{k+1}) \le \gamma \cdot \operatorname{rad}(\boldsymbol{x}^k)^2.$$

Proof. Using Theorem 6.6 and the notation and the statement of Theorem 8.3,

$$\begin{aligned}\operatorname{rad}(\boldsymbol{x}^{k+1}) &\le \left(1 - \underline{d}/\overline{d}\right) \operatorname{rad}(\boldsymbol{x}^k) = \left(\frac{\overline{d} - \underline{d}}{\overline{d}}\right) \operatorname{rad}(\boldsymbol{x}^k) \\ &\le \frac{2\operatorname{rad}(\boldsymbol{f}'(\boldsymbol{x}^k))}{\underline{f'}(\boldsymbol{x}^0)} \operatorname{rad}(\boldsymbol{x}^k) \le \frac{2\tilde{\gamma}}{\underline{f'}(\boldsymbol{x}^0)} \operatorname{rad}(\boldsymbol{x}^k)^2.\end{aligned}$$

for certain $\tilde{\gamma}$. Thus, we can put $\gamma := 2/\underline{f'}(\boldsymbol{x}^0)$. □

Example 8.5 (Floudas and Pardalos, 2009, Moore: Interval Analysis: Systems of Nonlinear Equations, 1721–1727)**.** Consider function $f(x) = x^3 - x + 0.2$, illustrated in Figure 8.3. It is known that all roots of a polynomial $x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$ lie in the interval with the center at the zero and the radius $\sum_{i=0}^{n-1} |a_i|$. Thus, take the initial enclosure $\boldsymbol{x}^0 := [-1.2, 1.2]$.

After one iteration, we have $N(\boldsymbol{x}^0) = (-\infty, -0.06024] \cup [0.05154, \infty)$, so we split $\boldsymbol{x}^0$ into two sub-intervals $\boldsymbol{x}^1 := [-1.2, -0.06024]$ and $\boldsymbol{x}^2 := [0.05154, 1.2]$. In the four next iterations, $\boldsymbol{x}^1$ is contracted to $[-1.0880568, -1.0880142]$, which provable contains a unique root.

Since $0 \in N(\boldsymbol{x}^2)$, the interval $\boldsymbol{x}^2$ is split into sub-intervals $\boldsymbol{x}^3 := [0.05154, 0.44359]$ and $\boldsymbol{x}^4 := [0.68020, 1.2]$. The former is contracted to $[0.2091487, 0.2091490]$ in three iterations, and the latter is contracted to $[0.8781166, 0.8797997]$ in three iterations as well. Both contain unique roots. The precision of the resulting enclosures is between 10^{-3} and 10^{-7}, but only one additional iteration applied to each interval increases the accuracy to one between 10^{-7} and 10^{-15}. □

Example 8.6. Consider the function by Moore (1993)

$$f(x) = x^2 + \sin(x^{-3}).$$

All 318 roots of in the interval $[0.1, 1]$ were found with accuracy 10^{-10} in 1538 seconds of that CPU time. The left most root is contained in $[0.10003280626, 0.10003280628]$. □

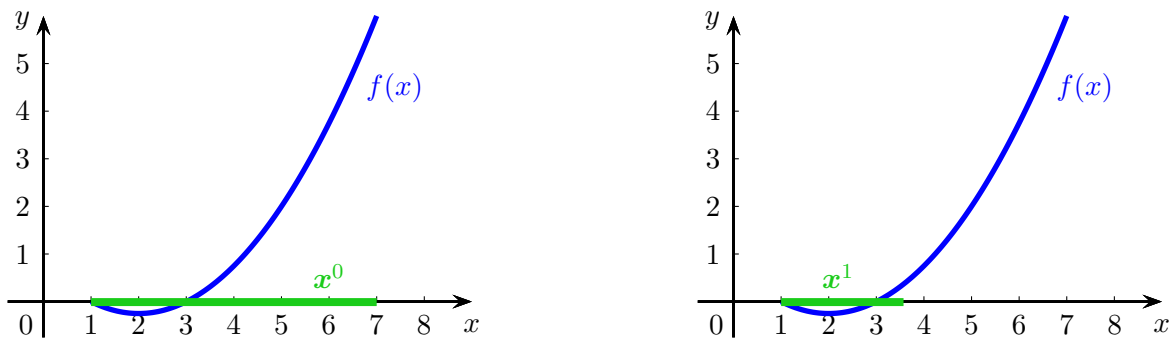


Figure 8.4: (Example 8.7) One step of the Interval Newton method by using slopes instead of derivatives.

Example 8.7. To make the interval Newton operator more efficient, we can use slopes instead of derivatives (see Section 6.2). The theory remains valid, only the uniqueness property of Theorem 8.10 may not hold. For illustration, consider a similar function as in Example 6.15

$$f(x) = \tfrac{1}{4}x^2 - x + \tfrac{3}{4}$$

and the initial domain $\boldsymbol{x}^0 = [1, 7]$. For $x^0 := x^c = 4$, the interval slope is $S_f(\boldsymbol{x}^0, x^0) = [\frac{1}{4}, \frac{7}{4}]$, and we calculate $N(\boldsymbol{x}^0, x^0) = [1, \frac{25}{7}]$. The initial interval $\boldsymbol{x}^0 = [1, 7]$ thus was significantly reduced to $\boldsymbol{x}^1 = [1, \frac{25}{7}]$; see Figure 8.4. On the other hand, there are two roots in $\boldsymbol{x}^0$ (the values of 1 and 3), even though we have $N(\boldsymbol{x}^0, x^0) \subseteq \boldsymbol{x}^0$. This shows that employing slopes yields tighter enclosures, but we lost evidence of uniqueness as stated in Theorem 8.10.

Further iterations will not improve the enclosure $\boldsymbol{x}^1$ when we use the midpoint value as the center of linearization. However, if we use an appropriate value (e.g., $x^k := (3 + \overline{x}^k)/2$), then the iterations will converge to $[1, 3]$, which is the interval hull of the roots. □

Multi-dimensional case

An interval linearization. The underlying idea of the Interval Newton (and the Krawczyk method presented in Section 8.1.2) is to linearize the hard nonlinear functions and reduce the problem to interval linear equations.

Let $x^0 \in \boldsymbol{x}$. Proceeding as in Section 6.1.2, we can write

$$f(x) \in f(x^0) + \nabla f(\boldsymbol{x})(x - x^0), \quad x \in \boldsymbol{x}.$$

Thus, we enclose a nonlinear function $f(x)$ by an interval linear function. Now, if $x^* \in \boldsymbol{x}$ is a root of f, then

$$0 \in f(x^0) + \nabla f(\boldsymbol{x})(x^* - x^0). \tag{8.2}$$

That is, $x^* - x^0$ belongs to the solution set of interval linear equations with the interval constraint matrix $\nabla f(\boldsymbol{x})$ and the real right-hand side $-f(x^0)$. This is the essence of the Interval Newton method.

The interval Newton method. We introduce the interval Newton operator as an operator yielding the following interval vector

$$N(\boldsymbol{x}, x^0) := \square\{x \in \mathbb{R}^n;\ \exists A \in \nabla f(\boldsymbol{x}) : A(x - x^0) = -f(x^0)\}.$$

Since it is hard to determine its value exactly, we will be content with a tight enclosure. Practically, we proceed as follows. Consider the system of interval linear equations $\boldsymbol{A}y = b$, where $\boldsymbol{A} := \nabla f(\boldsymbol{x})$ and $b := -f(x^0)$. Let $\boldsymbol{y} \in \mathbb{IR}^n$ be an enclosure of the solution set. Then we set $N(\boldsymbol{x}, x^0) := x^0 + \boldsymbol{y}$. Due to the properties of interval methods, we immediately have:

Theorem 8.8. *If $x^*, x^0 \in \boldsymbol{x}$ and $f(x^*) = 0$, then $x^* \in N(\boldsymbol{x}, x^0)$.*

Thus, the interval Newton operator misses no root of $f(x)$ on $\boldsymbol{x}$. We emphasize this property in the following.

Corollary 8.9. *All roots of f in $\boldsymbol{x}$ are contained in $\boldsymbol{x} \cap N(\boldsymbol{x}, x^0)$. In particular, if $\boldsymbol{x} \cap N(, \boldsymbol{x}, x^0) = \emptyset$, then there is no root in $\boldsymbol{x}$.*

We can say even more. If $N(\boldsymbol{x}^k, x^k) \subseteq \boldsymbol{x}^k$, then there is provably a unique root (Nickel, 1971). This (unique) existence property is one of the most important issues in interval analysis.

Theorem 8.10 (Nickel, 1971)**.** *If $\emptyset \neq N(\boldsymbol{x}, x^0) \subseteq \boldsymbol{x}$, then there is a unique root in $\boldsymbol{x}$ and $\nabla f(\boldsymbol{x})$ is regular.*

Proof. "Regularity." Since $\emptyset \neq N(\boldsymbol{x}, x^0) \subseteq \boldsymbol{x}$, the interval Jacobian $\nabla f(\boldsymbol{x})$ must be regular by Theorem 3.51.

"Existence." By the Mean value theorem, for each $x \in \boldsymbol{x}$ and $i \in \{1, \dots, n\}$, there is $c_i(x) \in \mathbb{R}^n$ such that $f_i(x) - f_i(x^0) = \nabla f_i(c_i(x))^T(x - x^0)$. Let $A(x) \in \mathbb{R}^{n \times n}$ be the matrix with rows $\nabla f_i(c_i(x))^T$, $i = 1, \dots, n$ (Notice that in general there is no $c(x) \in \boldsymbol{x}$ such that $A(x) = \nabla f(c(x))$, so we had to proceed for the particular functions separately.). Then we have

$$f(x) - f(x^0) = A(x)(x - x^0).$$

Now, consider the function

$$p(x) \coloneqq x - A(x)^{-1} f(x) = x - A(x)^{-1}(f(x) - f(x^0)) - A(x)^{-1} f(x^0) = x^0 - A(x)^{-1} f(x^0).$$

Since $p(x) \in N(\boldsymbol{x}, x^0) \subseteq \boldsymbol{x}$ $\forall x \in \boldsymbol{x}$, we have $p(\boldsymbol{x}) \subseteq \boldsymbol{x}$, and by Brouwer's fixed-point Theorem 8.1, there is a fixed point $x^* \in \boldsymbol{x}$ of $p(x)$, which gives

$$x^* = p(x^*) = x^* - A(x^*)^{-1} f(x^*)$$

and therefore $f(x^*) = 0$.

"Uniqueness." Suppose to the contrary that there are two different roots x_1 and x_2 in $\boldsymbol{x}$. By the Mean value theorem, $f(x_1) - f(x_2) = A(x_1 - x_2)$ for some $A \in \nabla f(\boldsymbol{x})$; notice that this does not mean that $A = \nabla f(x)$ for some $x \in \boldsymbol{x}$. Since $f(x_1) = f(x_2) = 0$, we get $A(x_1 - x_2) = 0$ and by the nonsingularity of A, the roots are identical. □

The algorithm. The above results justify the interval Newton method. It is an iterative method described below. At the beginning, we put $k \coloneqq 0$ and $\boldsymbol{x}^0 \coloneqq \boldsymbol{x}$, and the kth iteration is

1: choose $x^k \in \boldsymbol{x}^k$,
2: calculate $N(\boldsymbol{x}^k, x^k)$,
3: $\boldsymbol{x}^{k+1} \coloneqq \boldsymbol{x}^k \cap N(\boldsymbol{x}^k, x^k)$,
4: $k \coloneqq k + 1$.

The typical stopping criteria of the iterations are some (or a mixture) of the following conditions:

- $\sum_{i=1}^n \operatorname{rad}(\boldsymbol{x}_i^k) < \varepsilon$ for a given threshold $\varepsilon > 0$,
- $\|f(\boldsymbol{x}^k)\| < \varepsilon'$ for a given threshold $\varepsilon' > 0$ and vector norm $\|\cdot\|$,
- $0 \in f(\boldsymbol{x}^k)$, the Jacobian $\nabla f(\boldsymbol{x}^k)$ is regular and $N(\boldsymbol{x}^k, x^k) \subseteq \operatorname{int} \boldsymbol{x}^k$.

As we mentioned earlier, in step 2 we calculate $N(\boldsymbol{x}^k, x^k)$ simply by solving the interval linear system

$$\nabla f(\boldsymbol{x}^k)\big(N(\boldsymbol{x}^k, x^k) - x^k\big) = -f(x^k). \tag{8.3}$$

We may employ any method from Sections 3.2–3.5.5, but usually the preconditioned Gauss–Seidel or Krawczyk method is employed since they are iterative and give a sufficiently tight enclosure with small effort. When using these iteration methods, however, we must be careful about Theorem 8.10 since the

iteration methods typically utilize $\boldsymbol{x}^k$ as the starting enclosure. So they compute directly $N(\boldsymbol{x}^k, x^k) \cap \boldsymbol{x}^k$ instead of $N(\boldsymbol{x}^k, x^k)$ itself, which is good from the performance viewpoint, but we cannot check for the inclusion $N(\boldsymbol{x}^k, x^k) \subseteq \boldsymbol{x}^k$.

In order that the interval Jacobian $\nabla f(\boldsymbol{x})$ is as narrow as possible, we are free to employ the successive mean value form as described in Section 6.1.2. Instead of derivatives, we can compute the interval Jacobian $\nabla f(\boldsymbol{x})$ by using slopes (Section 6.2); see the illustration in Example 8.7. More details are in Kearfott (1996a); Neumaier (1990); Rump (1994) and the second order slope form was employed by Schichl and Neumaier (2004). As we observed in Example 8.7, however, the uniqueness property of Theorem 8.10 may not hold.

The performance of the method depends on the choice of $x^k \in \boldsymbol{x}^k$. See Alefeld and Herzberger (1983); Chernous'ko (1968) for a thorough discussion on the selection of this point and for observing that the choice $x^k := \operatorname{mid}(\boldsymbol{x}^k)$ is an optimal choice with respect to some robustness criteria. For this commonly used choice, we then simply write $N(\boldsymbol{x}^k)$ instead of $N(\boldsymbol{x}^k, \operatorname{mid}(\boldsymbol{x}^k))$.

Splitting boxes. The interval Newton operator need not shrink the input box $\boldsymbol{x}$. This usually happens when $\boldsymbol{x}$ is too wide and there are multiple roots inside, the interval Jacobian matrix is overestimated or some other conditions are not satisfied. In this case, we have to split the actual interval vector $\boldsymbol{x}$ into smaller sub-boxes. The most straightforward approach to split the box along the widest edge is not always the best choice. An appropriate may dramatically improve the overall time complexity. One possibility is to split along that ith coordinate for which

$$x_i^{\Delta} \sum_{j=1}^{n} \operatorname{mag}(\nabla f(\boldsymbol{x}))_{ji}$$

is maximal (Hansen and Walster, 2004).

Convergence. In the multidimensional case, quadratic convergence is observed under some general assumptions (Alefeld and Herzberger, 1983). However, as the following example shows, the iterations need not converge in every case even if the interval Jacobian $\nabla f(\boldsymbol{x})$ is regular; cf. Mayer (2017).

Example 8.11 (Schwandt, 1981)**.** Let

$$f(x) = (-x_1^2 + x_2^2 - 1,\ x_1^2 - x_2)^T, \quad \boldsymbol{x} = ([1.1, 1.9], [1.1, 1.9])^T.$$

Function $f(x)$ has the unique root x^* in $\boldsymbol{x}$

$$x^* = \left(\sqrt{\tfrac{1+\sqrt{5}}{2}}, \tfrac{1+\sqrt{5}}{2}\right)^T.$$

Even though $\nabla f(\boldsymbol{x})$ is regular, the interval Newton method does not shrink the initian box since $\boldsymbol{x} \subseteq N(\boldsymbol{x}, x^c)$ when computing $N(\boldsymbol{x}, x^c)$ by the interval Gaussian elimination. □

For other asymptotic properties see, e.g., Alefeld et al. (1998); Mayer (2017).

Remark 8.12. As we mentioned at the beginning of Section 8.1, the classical methods for finding the roots of $f(x)$ are often based on iterations converging to a fixed point of $g(x) = x + Cf(x)$. In particular, the iterations of the Newton method have the form

$$x^{k+1} := x^k - \nabla f(x^k)^{-1} f(x^k), \quad k = 0, 1, \ldots$$

However, the direct interval extension

$$\boldsymbol{x}^{k+1} := \boldsymbol{x}^k - \nabla f(\boldsymbol{x}^k)^{-1} f(\boldsymbol{x}^k)$$

will almost never work since the intervals tend to blow up; cf. Rump (2010). This shows the typical behavior of interval methods – the direct intervalization usually fails, and one has to proceed in a more sophisticated way.

Notes and further reading. The pioneering works on the interval Newton method are Sunaga (1958); Moore (1966); Chernous'ko (1968). There are variants of the interval Newton method known (Ratschek and Rokne 2003), including methods for non-smooth functions (Kearfott, 2002).

8.1.2 Krawczyk method

The Krawczyk iterations were introduced independently by Kahan (1968b) and Krawczyk (1969), and it is a variation of the interval Newton method. Instead of the Newton operator, we use the Krawczyk operator

$$K(\boldsymbol{x}, x^0, C) \coloneqq x^0 - Cf(x^0) + (I_n - C\nabla f(\boldsymbol{x}))(\boldsymbol{x} - x^0),$$

where $x^0 \in \boldsymbol{x}$ and $C \in \mathbb{R}^{n\times n}$ is a regular matrix, usually $C \approx \nabla f(x^0)^{-1}$. We will write simply $K(\boldsymbol{x})$ for $K(\boldsymbol{x}, x^0, C)$ provided x^0 and C are fixed.

Theorem 8.13. *Any root of $f(x)$ in $\boldsymbol{x}$ is included in $K(\boldsymbol{x})$.*

Proof. Consider the system of interval linear equations (8.2) and write it as

$$\boldsymbol{A}y = \boldsymbol{b}, \quad \text{where} \quad \boldsymbol{A} \coloneqq \nabla f(\boldsymbol{x}),\ \boldsymbol{b} \coloneqq -f(x^0),\ y \coloneqq x - x^0.$$

Applying the Krawczyk method from Section 3.4.2 on the preconditioned system $(C\boldsymbol{A})y = C\boldsymbol{b}$, we obtain the interval operator

$$K(\boldsymbol{y}) \coloneqq C\boldsymbol{b} + (I_n - C\boldsymbol{A})\boldsymbol{y} = -Cf(x^0) + (I_n - C\nabla f(\boldsymbol{x}))\boldsymbol{y}.$$

Substitution $y \coloneqq x - x^0$ then concludes the proof. □

Krawczyk method can be seen as the interval Newton method, in which we solve the interval linear system (8.3) by the Kraczyk method for linear systems and run only one iteration. That is, each iteration we update the interval Jacobian matrix $\nabla f(\boldsymbol{x})$.

By utilizing fixed point theory, we can state a condition for solvability. Under slightly stronger assumption, we can achieve uniqueness of a solution.

Theorem 8.14. *If $K(\boldsymbol{x}) \subseteq \boldsymbol{x}$, then there is a root in $\boldsymbol{x}$.*

Proof. We know that the roots of $f(x)$ are the fixed points of $g(x) \coloneqq x - Cf(x)$. By the mean value theorem, $g(x) \in g(x^0) + \nabla g(\boldsymbol{x})(x - x^0)$ for each $x \in \boldsymbol{x}$, whence

$$g(\boldsymbol{x}) \subseteq g(x^0) + \nabla g(\boldsymbol{x})(\boldsymbol{x} - x^0) = x^0 - Cf(x^0) + (I_n - C\nabla f(\boldsymbol{x}))(\boldsymbol{x} - x^0) = K(\boldsymbol{x}) \subseteq \boldsymbol{x}.$$

By Theorem 8.1, there is a fixed point of $g(x)$ in $\boldsymbol{x}$, which corresponds to a root of $f(x)$. □

Theorem 8.15. *If $K(\boldsymbol{x}) \subseteq \operatorname{int}\boldsymbol{x}$, then there is a unique root in $\boldsymbol{x}$ and $\nabla f(\boldsymbol{x})$ is regular.*

Proof. The inclusion $K(\boldsymbol{x}) \subseteq \operatorname{int}\boldsymbol{x}$ reads

$$-Cf(x^0) + (I_n - C\nabla f(\boldsymbol{x}))(\boldsymbol{x} - x^0) \subseteq \operatorname{int}(\boldsymbol{x} - x^0)$$

By Theorem 3.41, applied for $\boldsymbol{b} \coloneqq -f(x^0)$, $\boldsymbol{A} \coloneqq \nabla f(\boldsymbol{x})$ and $\boldsymbol{x} \coloneqq \boldsymbol{x} - x^0$, we have that $\nabla f(\boldsymbol{x})$ is regular. Regularity now implies uniqueness in the same manner as in the proof of Theorem 8.10. □

A nice property of the Krawczyk method is that the interval operation is applicable even though $\nabla f(\boldsymbol{x})$ is irregular. The radius of $K(\boldsymbol{x})$ is minimal for $x^0 \coloneqq x^c$, and then the smallest value of $K(\boldsymbol{x})$ with respect to inclusion is achieved when preconditioned by $C = \nabla f(x^c)^{-1}$; see Mayer (2017); Neumaier (1990). We prove the former.

Theorem 8.16. *We have* $\operatorname{rad}(K(\boldsymbol{x}, x^c, C)) \le \operatorname{rad}(K(\boldsymbol{x}, x^0, C))$.

Proof. Using the property $\operatorname{mag}(\boldsymbol{A})\operatorname{rad}(\boldsymbol{B}) \le \operatorname{rad}(\boldsymbol{AB})$ (Table B.3), we can write

$$\begin{aligned}
\operatorname{rad}(K(\boldsymbol{x}, x^c, C)) &= \operatorname{mag}(I_n - C\nabla f(\boldsymbol{x}))\operatorname{rad}(\boldsymbol{x} - x^c) \\
&= \operatorname{mag}(I_n - C\nabla f(\boldsymbol{x}))\operatorname{rad}(\boldsymbol{x} - x^0) \\
&\le \operatorname{rad}\big((I_n - C\nabla f(\boldsymbol{x}))(\boldsymbol{x} - x^0)\big) \\
&= \operatorname{rad}(K(\boldsymbol{x}, x^0, C)).
\end{aligned}$$

□

Convergence. If $\beta \coloneqq \rho(\operatorname{mag}(I_n - C\nabla f(\boldsymbol{x}))) < 1$, and we choose the midpoints as the centers x^k-s, then the Krawczyk method converges with a sequence of nested intervals the radii of which converge linearly to zero with a factor β. Notice that by the proof of Theorem 3.41, the condition $\beta < 1$ is satisfied when the strong Krawczyk test $K(\boldsymbol{x}) \subseteq \operatorname{int} \boldsymbol{x}$ succeeds.

Theorem 8.17. *Suppose $\beta < 1$ and we choose $x^0 \coloneqq x^c$. Then there is a vector norm (not depending on $\boldsymbol{x}$) such that*

$$\|K(\boldsymbol{x})^\Delta\| \leq \beta \|x^\Delta\|.$$

Proof. Assume that the matrix $M \coloneqq \operatorname{mag}(I_n - C\nabla f(\boldsymbol{x}))$ is componentwisely positive. If it is not the case, we can replace the zero entries by sufficiently small $\varepsilon > 0$. Now, let $v > 0$ be the Perron vector corresponding to the eigenvalue β of the matrix M^T. We will utilize the 1-norm with the scaling given by v, that is, $\|x\| \coloneqq \|\operatorname{diag}(v)x\|_1$. Then the induced matrix norm is $\|M\| = \|\operatorname{diag}(v)M \operatorname{diag}(v)^{-1}\|_1 = \beta$. Eventually, we have

$$\begin{aligned} \|K(\boldsymbol{x})^\Delta\| &= \|(I_n - C\nabla f(\boldsymbol{x}))(\boldsymbol{x} - x^c)\| = \|M[-x^\Delta, x^\Delta]\| \\ &\leq \|M\| \cdot \|x^\Delta\| = \beta \|x^\Delta\|. \end{aligned}$$

□

If $\beta < \frac{1}{2}$, then we have two stronger results: First, the convergence is superlinear (Alefeld and Herzberger, 1983). Second, the linear convergence is true for any selection of the centers $x^k \in \boldsymbol{x}^k$ (Mayer, 2017; Neumaier, 1990).

Under certain assumptions, the Krawczyk method converges quadratically; see (Mayer, 2017, Thm. 6.3.5) or Moore et al. (2009).

Notes. Mayer (2017) discusses several variants of the Krawczyk method. In particular, he addresses several combinations of the Krawczyk and the Gauss–Seidel methods.

Remark 8.18 (Another linearizations)**.** The principle of the interval Newton method is to linearize the nonlinear constraints $f(x) = 0$ to obtain an interval linear system of equations $\boldsymbol{A}x = \boldsymbol{b}$, the solution set of which encloses the original solutions. Nevertheless, such a linearization can be performed in diverse ways. For instance, the nonlinear system

$$\begin{aligned} x_1^2 x_2 + x_2 \sin(x_2) - \exp(x_1) + 1 &= 0, \\ 3x_1^5 - x_2 \cos(x_1) - 7 &= 0 \end{aligned}$$

can be written as

$$\begin{pmatrix} x_1 x_2 & \sin(x_2) \\ 3x_1^4 & -\cos(x_1) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \exp(x_1) - 1 \\ 7 \end{pmatrix}.$$

Evaluating the ranges of particular entries of the constraint matrix and the right hand side vector, we obtain an interval linear system of type $\boldsymbol{A}x = \boldsymbol{b}$.

Notes and further reading (Section 8.1). There are basically two approaches to solving nonlinear systems by interval computation. The first one (including the interval Newton method) is based on evaluation of the interval Jacobian matrix $\nabla f(\boldsymbol{x})$ and solving the corresponding interval linear system of equations. The second approach (including the Krawczyk method) is based on an iterative method and updates the interval Jacobian matrix during each iteration. Another popular method belonging to the second class is that by Hansen and Sengupta (1981), who apply Gauss–Seidel iterations and show similar properties as for the Krawczyk method (cf. Mayer (2017)). Goldsztejn (2007) then considers the Jacobi iterations.

For more information about the interval Newton or Krawczyk method, see, e.g., Floudas and Pardalos (2009); Neumaier (1990); Rump (2010). Reinforcement learning approaches to this problem were incorporated by Goualard and Jermann (2008). Interval methods for finding real and complex roots of polynomials are discussed in Alefeld and Herzberger (1983). Some test problems are recorded in Kearfott (1987).

An existence of a solution can also be proved based on other topological theorems than Brouwer's fixed-point theorem, for instance, on the theorems of Kantorovich, Miranda or Borsuk (Alefeld et al., 2004; Franek et al., 2018).

It is a basic problem of numerical mathematics to solve a system of nonlinear equations and it finds applications everywhere. As an example, Moore (1979) finds rigorously the optimal solutions of a nonlinear programming problem by utilizing optimality conditions and solving the associated system of equations.

8.2 Constraint satisfaction problem

Constraint programming is a discipline in which we are solving a set of constraints. It became popular especially in the discrete case, when the domains of variables are finite sets. Herein, we consider a general constraint programming problem, but the main applicability is for continuous domains. Surveys on this topic are provided, e.g. in Benhamou and Granvilliers (2006); Hyvönen (1992); Jaulin et al. (2001).

The setup. We are given a set of constraints $c_k(x)$, $k = 1, \ldots, K$ with variables $x_1, \ldots, x_n$, and the aim is to find a verified enclosure to the solution set Σ. Let $\boldsymbol{x}^0 \in \mathbb{IR}^n$ be an initial box for domains of variables, in which we are seeking for solutions. Constraints may involve:

- equations $f_j(x) = 0$, $j = 1, \ldots, j^*$,
- inequalities $g_\ell(x) \leq 0$, $\ell = 1, \ldots, \ell^*$,
- other constraints, e.g., inequations $h_1(x) \neq h_2(x)$, integer conditions $x_k \in \mathbb{Z}$ for some k, using quantifications and logical operators $(\exists y \in \boldsymbol{y} : f(x) = y)$, various orderings (e.g., lexicographic) etc.

We suppose that the functions characterizing the constraints are real-valued, however, the methods that we will describe apply to interval-valued functions, too.

Branch & prune algorithm. The basic algorithms to compute verified enclosures of all solutions follow a branch & prune scheme. Let $\boldsymbol{x} \in \mathbb{IR}^n$ be given. One of the three possibilities may happen; we find out that:

1. There is no solution in $\boldsymbol{x}$, that is, $\boldsymbol{x} \cap \Sigma = \emptyset$.
2. The box $\boldsymbol{x}$ consists only of solutions, that is, $\boldsymbol{x} \subseteq \Sigma$.
3. We cannot decide, i.e., the box $\boldsymbol{x}$ possibly contains both solutions and not feasible solutions.

If the first condition holds true, then we remove $\boldsymbol{x}$. Some simple tests for the first case are as follows:

- If $0 \notin f_i(\boldsymbol{x})$, then the ith equation holds for no $x \in \boldsymbol{x}$.
- If $g_j(\boldsymbol{x}) > 0$, then the jth inequality cannot hold.

Since determining the range of a function on intervals is a difficult problem, too, it suffices to utilize an appropriate interval extension. Not surprisingly, the more tighter enclosure on the range, the more stronger condition we get. Other constraints than equations and inequalities require special treatment.

If we know that the second condition is true, then we output $\boldsymbol{x}$ to be a box containing no infeasible point. For equations, this case is not likely to happen. However, if all constraints are inequalities, then $g_\ell(\boldsymbol{x}) \leq 0$ for each $\ell = 1, \ldots, \ell^*$ implies that this is the case.

If we cannot decide for the first or second condition, we split the box $\boldsymbol{x}$ into two or more sub-boxes and recursively repeat the process. The termination condition is that the box is sufficiently small, for instance, $\max_{i=1,\ldots,n} x_i^\Delta < \varepsilon$ for a given threshold $\varepsilon > 0$.

Algorithm 8.1 displays a scheme of the branch & prune algorithm. We keep three lists of boxes, $\mathcal{L}$, $\mathcal{V}$ and $\mathcal{S}$. The first one is the list of working boxes, whereas the remaining ones go for output. Boxes in $\mathcal{L}$ and $\mathcal{V}$ possibly contain solutions and boxes in $\mathcal{S}$ contain only solutions. A threshold parameter $\varepsilon > 0$ the accuracy of approximation of the solution set by boxes. As the outpu, we get an inner approximation of Σ in the form $\cup\mathcal{S}$ and an outer approximation of Σ in the form $\cup(\mathcal{V} \cup \mathcal{S})$. Covering the solution set by the boxes from $\mathcal{S}$ and $\mathcal{V}$ is called a *subpaving* (Jaulin et al., 2001); see Figure 8.5. Sometimes, we keep also the set of removed boxes.

Algorithm 8.1 Branch & prune algorithm

1: $\mathcal{L} := \{\boldsymbol{x}^0\}$, $\mathcal{V} := \emptyset$, $\mathcal{S} := \emptyset$,
2: **while** $\mathcal{L} \neq \emptyset$ **do**
3: choose $\boldsymbol{x} \in \mathcal{L}$ and remove $\boldsymbol{x}$ from $\mathcal{L}$,
4: contract $\boldsymbol{x}$,
5: **if** $\boldsymbol{x}$ contains only solutions **then**
6: put $\boldsymbol{x}$ into $\mathcal{S}$,
7: **else if** $\boldsymbol{x}$ contains possibly a solution **then**
8: **if** $\max_i x_i^{\Delta} > \varepsilon$ **then**
9: split $\boldsymbol{x}$ into sub-boxes and put them into $\mathcal{L}$,
10: **else**
11: put $\boldsymbol{x}$ into $\mathcal{V}$,
12: **end if**
13: **end if**
14: **end while**
15: **return** Sets $\mathcal{V}, \mathcal{S}$ such that $\cup\mathcal{S} \subseteq \Sigma \subseteq \cup(\mathcal{V} \cup \mathcal{S})$.

(a) The exact solution set. (b) Subpaving approximation.

Figure 8.5: (Example 8.19) The exact solution set and its subpaving.

Efficiency of the branch & prune method depends on several factors. The choice how to split the box is important, since we can split along various coordinates and to two or more sub-boxes, not necessarily uniformly. More importantly, the overall performance is highly improved by consistency contractors (step 4); they reduce boxes without omitting any feasible solution, and we discuss them in the following sections.

Example 8.19. Consider the constraints

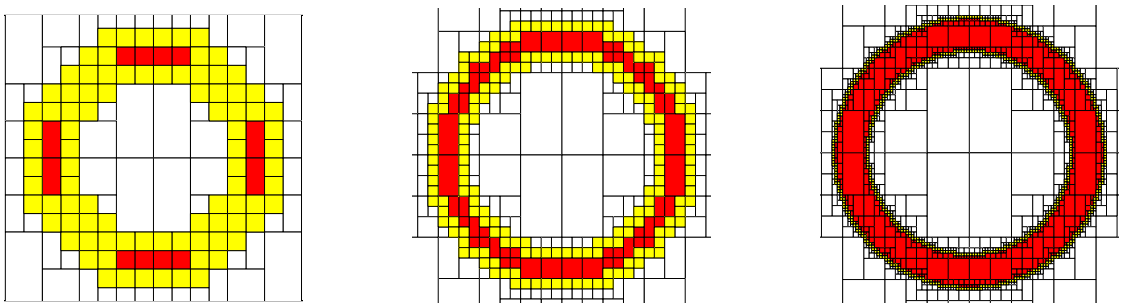
$$x^2 + y^2 \leq 16, \quad x^2 + y^2 \geq 9$$

with the initial domain $x, y \in [-5, 5]$ and precision $\varepsilon = 0.25$. Figure 8.5 illustrates the solution set and its subpaving obtained by Algorithm 8.1. The red boxes are contained in the solution set Σ (boxes in $\mathcal{S}$), the yellow are not decidable or distinguishable (boxes in $\mathcal{V}$), and the white ones contain no solution. Figure 8.6 shows subpavings for three different precisions; the running time (computed in 2014) is for the simple branch & prune method with no improvements and contractions discussed in the following sections. □

Example 8.20. Consider the constraints

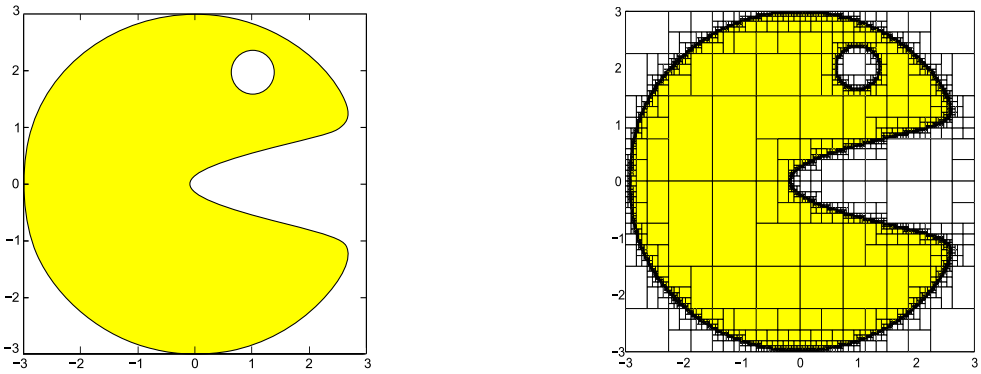
$$(x-1)^2 + (y-2)^2 \leq \tfrac{1}{7},$$
$$(x^2 + y^2 - 9)(\tfrac{1}{3}x - y^2) \geq \tfrac{1}{2}.$$

The solution set and its subpaving are illustrated in Figure 8.7. The inner part of the solution set is now in yellow and the outer part in white. □



(a) $\varepsilon = 1.0$, time: 0.952 s. (b) $\varepsilon = 0.5$, time: 2.224 s. (c) $\varepsilon = 0.125$, time: 9.966 s.

Figure 8.6: (Example 8.19) Subpavings for different precisions.



(a) The exact solution set. (b) Subpaving approximation.

Figure 8.7: (Example 8.20) Exact solution set and its subpaving.

Notes and further reading. Instead of boxes, we can cover the solution set by other geometrical objects, for instance, by polyhedral sets (Araya et al., 2012; Lebbah et al., 2005). They may be more efficient by means of their number and accuracy, but one has to effectively evaluate the range of a function, prune, branch and contract over polyhedral sets.

A more general constraint programming equipped with quantifiers was treated, e.g., in Ratschan (2006), and implemented in RSolver by Ratschan et al. (2004).

8.2.1 Consistency and contractors

The branching in the branch & prune method, as described in Algorithm 8.1, may produce a lot of boxes to be processed, which in turn cause high time complexity. To enhance efficiency of the method, we have to postpone the splitting as far as possible and apply strong contractor techniques first. *Contractors* are methods that reduce the domains of variables without omitting any feasible solution. A framework for contractor programming was set up in Chabert and Jaulin (2009); Jaulin et al. (2001); Lebbah and Lhomme (2002).

But how much we can or is possible to shrink the interval domains? The limit state we can achieve is called *consistency*. There are several types of consistencies (Collavizza et al., 1999).

2B-consistency. A set of constraints $c_k(x)$, $k = 1, \ldots, K$, on a box $\boldsymbol{x}^0 \in \mathbb{IR}^n$ is *2B-consistent* if for each $k \in \{1, \ldots, K\}$ and each $i \in \{1, \ldots, n\}$ there are some $x, x' \in \boldsymbol{x}^0$ such that $x_i = \underline{x}_i^0$, $x'_i = \overline{x}_i^0$, and conditions $c_k(x)$ and $c_k(x')$ are valid.

2B-consistency says that, looking at the constraints separately, we cannot shrink any domain without omitting a feasible point.

Example 8.21. Consider a constraint $x_1 = x_2^2$ over $\boldsymbol{x}^0 = ([1, 4], [-2, 2])^T$. This setting is 2B-consistent even though, for instance, for $x_2 = 0$ there is no corresponding $x_1 \in [1, 4]$ to solve the constraint. Thus, 2B-consistency slightly differs from arc-consistency studied in discrete constraint programming. □

3B-consistency. Checking constraints separately for a consistency may be very restrictive. 3B-consistency tries to overcome this drawback by interconnecting the particular constraints in the following manner. A set of constraints $c_k(x)$, $k = 1, \ldots, K$, on a box $\boldsymbol{x}^0 \in \mathbb{IR}^n$ is *3B-consistent* if for each $i \in \{1, \ldots, n\}$ there are some $x, x' \in \boldsymbol{x}^0$ such that $x_i = \underline{x}_i^0$, $x_i' = \overline{x}_i^0$, and conditions $c_k(x)$ and $c_k(x')$, $k = 1, \ldots, K$, are valid. Notice that after fixing a variable to the lower or upper limit, the program may become 2B-inconsistent, but we care only about solvability.

In other words, a system is 3B-consistent if each facet of the box $\boldsymbol{x}$ contains at least one solution. It is easy to see that 3B-consistency implies 2B-consistency.

Example 8.22. Consider a system $x + y = 0,\ x - y = 0$ on $x, y \in [-1, 1]$. We cannot shrink any input interval if we process the constraints separately. Thus the domains are 2B-consistent. However, the unique solution is $(x, y) = (0, 0)$, so the initial domains are not 3B-consistent. □

Consistency and contractors. Contractors can be viewed as techniques that try to approach the desired kind of consistency. In the following sections, we describe several contractors. Most of them, including constraint programming, box and hull consistency contractors approach 2B-consistency since they process the constraints separately. In practice, it is a good habit to combine the contractors to achieve the highest efficiency (Granvilliers, 2001).

8.2.2 Constraint propagation

Constraint propagation is a simple but efficient contractor, that is, it shrinks the interval domains with omitting no solution (Hickey et al., 1998). It processes the constraints individually.

The basic idea is to express the particular variables by means of the others and by evaluation by interval arithmetic a new interval domain is obtained. For example, consider variables x, y, z with interval domains $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{IR}$. The interval $\boldsymbol{x}$ can be shrunk as follows, depending on the form of the constraint:

1. If the constraint reads $x + y = z$, then we contract $\boldsymbol{x} \coloneqq \boldsymbol{x} \cap (\boldsymbol{z} - \boldsymbol{y})$.
2. If the constraint reads $xy = z$, then we contract $\boldsymbol{x} \coloneqq \boldsymbol{x} \cap (\boldsymbol{z}/\boldsymbol{y})$.
3. If the constraint reads $x^2 = z$, then we contract $\boldsymbol{x} \coloneqq \boldsymbol{x} \cap [\sqrt{\mathrm{mig}(\boldsymbol{z})}, \sqrt{\mathrm{mag}(\boldsymbol{z})}]$.

We handle other operations (subtraction, division, ...) and basic functions (exp, sin, ...) similarly. If the function f is more complex, which is the typical case, then we recursively propagate the contractions to the sub-expressions.

From this viewpoint, the Gauss–Seidel method for interval linear equations (Section 3.4.1) can be viewed as a constraint propagation contracting the diagonal entries only.

Example 8.23. Consider the constraint

$$x + yz = 7, \quad x \in [0, 3],\ y \in [3, 5],\ z \in [2, 4].$$

The tightening of the domains of variables proceeds as follows:

- Express x
$$x = 7 - yz \in 7 - [3, 5][2, 4] = [-13, 1].$$
Thus, the domain of x is $[0, 3] \cap [-13, 1] = [0, 1]$.

- Express y
$$y = (7 - x)/z \in (7 - [0, 1])/[2, 4] = [1.5,\ 3.5].$$
Thus, the domain of y is $[3, 5] \cap [1.5,\ 3.5] = [3,\ 3.5]$.

- Express z
$$z = (7 - x)/y \in (7 - [0, 1])/[3,\ 3.5] = [\tfrac{12}{7}, \tfrac{7}{3}].$$
Thus, the domain of z is $[2, 4] \cap [\frac{12}{7}, \frac{7}{3}] = [2, \frac{7}{3}]$.

Table 8.1: (Example 8.24)Domain reductions of particular iterations.

iteration	$\boldsymbol{x}$	$\boldsymbol{y}$	$\boldsymbol{z}$
1	$[4, 4.2485]$	$[3.4991, 4]$	$[2.6243, 3]$
2	$[4, 4.1106]$	$[3.6165, 4]$	$[2.7124, 3]$
3	$[4, 4.0831]$	$[3.6409, 4]$	$[2.7306, 3]$
4	$[4, 4.0775]$	$[3.6458, 4]$	$[2.7344, 3]$
5	$[4, 4.0764]$	$[3.6469, 4]$	$[2.7351, 3]$
$\vdots$			
∞	$[4, 4.0761]$	$[3.6471, 4]$	$[2.7353, 3]$

Even though the domains have shrunken, no further propagation makes sense because each variable appears just once in the expression. The reason is simply that in the domain reduction for z we removed values in $\boldsymbol{z}$ that are not in the image of $\boldsymbol{x}$, so the shrunken domain for z makes no progress in reducing the domain of x (Trombettoni et al., 2010). □

Example 8.24. Consider the constraint

$$e^x - xyz = 10, \quad x \in \boldsymbol{x} = [4, 5],\ y \in \boldsymbol{y} = [3, 4],\ z \in \boldsymbol{z} = [2, 3].$$

- Express the first occurrence of x

$$x = \log(10 + xyz) \in \log(10 + \boldsymbol{xyz}) \cap \boldsymbol{x} = [3.5263,\ 4.2485] \cap \boldsymbol{x} = [4,\ 4.2485].$$

- Express y

$$y = (e^x - 10)/(xz) \in (e^{\boldsymbol{x}} - 10)/(\boldsymbol{xz}) \cap \boldsymbol{y} = [3.4991,\ 7.5001] \cap \boldsymbol{y} = [3.4991,\ 4].$$

- Express z

$$z = (e^x - 10)/(xy) \in (e^{\boldsymbol{x}} - 10)/(\boldsymbol{xy}) \cap \boldsymbol{z} = [2.6243,\ 4.2868] \cap \boldsymbol{z} = [2.6243,\ 3].$$

- Express the second occurrence of x

$$x = (e^x - 10)/(yz) \in (e^{\boldsymbol{x}} - 10)/(\boldsymbol{yz}) \cap \boldsymbol{x} = [3.7165,\ 6.5339] \cap \boldsymbol{x} = [4,\ 4.2485].$$

Now, we need not stop as double the appearance of x in the expression causes that the propagation is not perfect and other iterations may further reduce the domain. In our example, the iterative propagation converges to the domains shown in Table 8.1. □

The above example shows that constraint propagation cannot effectively handle dependencies. Indeed, this is a difficult problem, and the methods presented later can deal with them only partially.

In principle, the domain of parameters may be the whole space $\mathbb{R}$. The following example show that we can process even unbounded intervals.

Example 8.25. Let

$$x^2 + y^2 = 1, \quad x, y \in \mathbb{R}.$$

By a simple propagation,

$$x \in \sqrt{1 - \boldsymbol{y}^2} \subseteq \sqrt{1 - [0, \infty)} = [-1, 1],$$

whence also $y \in [-1, 1]$. Thus, both domains were reduced to $[-1, 1]$ in one step. □

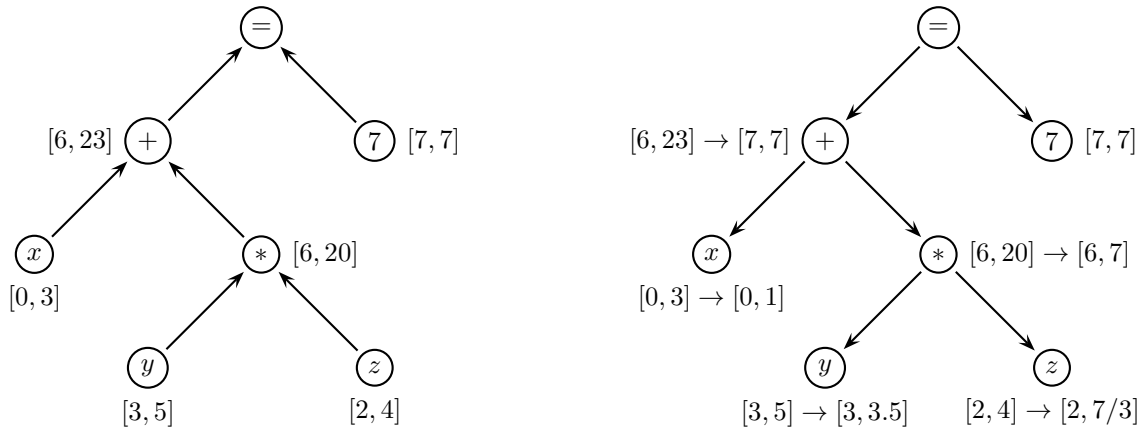


Figure 8.8: (Example 8.23) Upward and downward propagation.

The expression tree. The constraint propagation is more useful to describe (and implement) by using the expression tree. Figure 8.8 shows the tree for the constraint from Example 8.23. Each iteration can be implemented as an upward and downward evaluation and domain reduction. In the upward phase, we evaluate the expression by interval arithmetic. In the downward phase, we reduce the domains of variables on the intermediate expressions.

8.2.3 Other contractors

First we describe box and hull consistency contractors according to Hansen and Walster (2004). Suppose that the constraints have the form of equations $f_j(x) = 0$, $j = 1, \dots, J$, and the initial domain is $\boldsymbol{x}^0 \in \mathbb{IR}^n$. In principle, any inequality $g_\ell(x) \le 0$ can be transformed to an equation by writing $g_\ell(x) + [0, \infty] = 0$.

Both contractors work coordinate by coordinate, so they can be easily parallelized.

Box consistency contractor. Box consistency was pioneered by Benhamou et al. (1994). Let $i \in \{1, \dots, n\}$, $j \in \{1, \dots, j^*\}$ and define an interval-valued function

$$\boldsymbol{h}(y) \coloneqq f_j(\boldsymbol{x}_1^0, \dots, \boldsymbol{x}_{i-1}^0, y, \boldsymbol{x}_{i+1}^0, \dots, \boldsymbol{x}_n^0).$$

We say that the domain $\boldsymbol{x}_i^0$ of variable x_i is *box consistent* with the jth constraint if

$$\boldsymbol{x}_i^0 = \square\{x_i \in \boldsymbol{x}_i^0;\, 0 \in \boldsymbol{h}(x_i)\}.$$

Provided $\boldsymbol{h}(y)$ overestimate the true range, box consistency is weaker than 2B-consistency.

To narrow the interval domain $\boldsymbol{x}_i^0$ and approach the box consistency, we iteratively cutt off the left part of the domain, and similarly for the right part of the domain. Define $\boldsymbol{y}^0 \coloneqq [\underline{x}_i^0, \underline{x}_i^0 + \alpha \operatorname{rad}(\boldsymbol{x}_i^0)]$, where $0 \le \alpha \le 2$ is a parameter. A recommended choice of the parameter is to put $\alpha = \frac{1}{2}$ and dynamically adjust according to the computation. The smaller α, the more effective is the consistency, but the smaller the cuts are.

Now, we call the interval Newton method to shrink the interval $\boldsymbol{y}^0$. The interval Newton operator is

$$N(\boldsymbol{y}) \coloneqq y - \boldsymbol{h}'(\boldsymbol{y})^{-1} \boldsymbol{h}(y),$$

where $y \in \boldsymbol{y}$ is arbitrarily chosen, and the shrinking step reads

$$\boldsymbol{y}^1 \coloneqq \boldsymbol{y}^0 \cap N(\boldsymbol{y}^0).$$

If we reduce $\boldsymbol{y}^0$ to a sub-interval $\boldsymbol{y}^*$, then the ith domain is narrowed to $[\underline{y}^*, \overline{x}_i^0]$. If we reduce $\boldsymbol{y}^0$ to the empty set, then the ith domain is shrunken to $[\underline{x}_i^0 + \alpha \operatorname{rad}(x_i^0), \overline{x}_i^0]$.

Hull consistency contractor. As in the box consistency, we go through particular equations and for each we fix all variables but one and define a function

$$\boldsymbol{h}(y) \coloneqq f_j(\boldsymbol{x}_1^0, \ldots, \boldsymbol{x}_{i-1}^0, y, \boldsymbol{x}_{i+1}^0, \ldots, \boldsymbol{x}_n^0)$$

on $\boldsymbol{y} \coloneqq \boldsymbol{x}_i^0$. Now, the idea is to express $\boldsymbol{h}(y)$ as $\boldsymbol{h}(y) = \boldsymbol{h}_1(y) - \boldsymbol{h}_2(y)$, where $\boldsymbol{h}_1(y)$ is an easily invertible function. Then for any root y, we can write $y = \boldsymbol{h}_1^{-1}(\boldsymbol{h}_2(y))$, which gives rise to the operator

$$C(\boldsymbol{y}) \coloneqq \boldsymbol{h}_1^{-1}(\boldsymbol{h}_2(\boldsymbol{y})),$$

which does not omit any root in $\boldsymbol{y}$. Hence we have a contracting

$$\boldsymbol{y}^1 \coloneqq C(\boldsymbol{y}) \cap \boldsymbol{y}.$$

The expansion $\boldsymbol{h}(y) = \boldsymbol{h}_1(y) - \boldsymbol{h}_2(y)$ is not unique and there are usually many ways to do it. It is hard to choose the right one, however, when we find a suitable expansion, then the contractor works very well.

Example 8.26. Consider the constraint $x^4 + x - 2 = 0$ on $x \in [-100, 100]$. The good choice is $h_1(x) = x^4$ and $h_2(x) = 2 - x$, since we derive $x^4 = 2 - x \in [-98, 100]$. Due to non-negativity of the fourth power, we get $x^4 \in [0, 100]$, and so $x \in [-\sqrt{10}, \sqrt{10}] \subseteq [-3.17, 3.17]$. □

Example 8.27. Let $x^2(2 + \sin(\pi x)) \le 2$ be a constraint with $x \in [2, \infty]$. Derive $x^2 \le \frac{2}{2+\sin(\pi x)} \in \left[\frac{2}{3}, 2\right]$, whence $x \in [-\sqrt{2}, \sqrt{2}]$. Therefore there is no solution on $x \in [2, \infty]$. □

Interval extensions. When constructing a contractor, we are free to employ any interval extension of the functions involved.

For instance, consider a constraint $f(x) \le c$ with a function $f\colon \mathbb{R} \to \mathbb{R}$ and $c \in \mathbb{R}$. Let us have a mean value form

$$f(x) \in f(a) + \boldsymbol{s}(\boldsymbol{x} - a),$$

where $a \in \boldsymbol{x}$ and $\boldsymbol{s} \in \mathbb{IR}$ is a slope (Section 6.2). Suppose $\underline{s} > 0$; the case $\overline{s} < 0$ is dealt similarly. Relaxing the constraint to

$$f(a) + s(x - a) \le c, \quad s \in \boldsymbol{s},\ x \in \boldsymbol{x},$$

we easily express the variable x

$$x \le \frac{c - f(a)}{s} + a, \quad s \in \boldsymbol{s},\ x \in \boldsymbol{x},$$

from which we derive a new upper bound on x

$$x \le \begin{cases} \frac{c-f(a)}{\overline{s}} + a & \text{if } c \le f(a), \\ \frac{c-f(a)}{\underline{s}} + a & \text{if } c > f(a). \end{cases}.$$

This approach was utilized in Ratz (2001) for contracting the domain in global optimization when an upper bound c on the optimal value is known; see Section 8.3.

Another utilization of the mean value forms will be discussed in Section 9.4.2.

Linear combination. The above-mentioned methods process particular constraints separately, which is often too restrictive. As illustrated by Example 8.28, a combination of constraints may be very productive.

Example 8.28 (Example 8.22 continued)**.** Consider a system $x + y = 0,\ x - y = 0$ on $x, y \in [-1, 1]$. Since it is 2B-consistent, we cannot narrow any interval when considering the equations individually. However, summing up the equations, we get $2x = 0$, from which we immediately have the solution $x = y = 0$. □

In general, it is a nontrivial problem to find a right linear combination. People use to precondition the system $f(x) = 0$ by the matrix $C \approx \nabla f(x^c)^{-1}$, since the Jacobian matrix characterizes the tangents at x^c and preconditioning by its inverse transforms the tangents to be more-or-less parallel to the coordinate system.

Remark 8.29. Consistency techniques enable to improve enclosures on the range of functions over intervals, the problem studied in Chapter 6. Let $f\colon \mathbb{R}^n \to \mathbb{R}$, $\boldsymbol{x} \in \mathbb{IR}^n$ and $\boldsymbol{f}(\boldsymbol{x})$ an enclosure to the range $f(\boldsymbol{x})$. Next, let $\boldsymbol{a} \subseteq f(\boldsymbol{x})$ be an inner approximation to the range of f. It is determined, for example, by evaluating f at several points, and $\underline{a}$ is taken as their minimum and $\overline{a}$ as their maximum. To improve the lower bound on the range, we consider the inequality $f(x) \leq \underline{a}$ on $x \in \boldsymbol{x}$. Consistency techniques may narrow the domain intervals in $\boldsymbol{x}$, which in turn leads to tighter evaluation of an enclosure. We proceed analogously for the upper bound by employing $f(x) \geq \overline{a}$ with $x \in \boldsymbol{x}$.

Software. We list some of the free software for solving constraints. Some of them return a subpaving of the solution set and enable its visualization.

- *Alias* (by Jean-Pierre Merlet, COPRIN team),
  A C++ library for system solving, with Maple interface,
  `http://www-sop.inria.fr/coprin/logiciels/ALIAS/ALIAS-C++/ALIAS-C++.html`
- *Quimper* (by Gill Chabert and Luc Jaulin),
  written in an interval C++ library IBEX,
  a language for interval modelling and handling constraints,
  `http://www.ibex-lib.org/`
- *RealPaver* (by L. Granvilliers and F. Benhamou),
  a C++ package for modeling and solving nonlinear and nonconvex constraint satisfaction problems,
  `http://pagesperso.lina.univ-nantes.fr/~granvilliers-l/realpaver/`
- *RSolver* (by Stefan Ratschan),
  solver for quantified constraints over the real numbers,
  implemented in the programming language OCaml,
  `http://rsolver.sourceforge.net/`

Notes and further reading. For the sake of propagation, it might be useful to consider some generalized concepts of intervals, e.g., interval unions or discontinuous intervals; cf. Section 7.1.4. The motivation is clear – consider the constraint

$$y = x^2, \quad x \in \boldsymbol{x} = [-5, 5],\ y \in \boldsymbol{y} = [4, 9].$$

A simple propagation reduces the domain $\boldsymbol{x}$ to $[-3, 3]$, but, in essence, one can reduce the domain further to the union of two intervals $[-3, -2] \cup [2, 3]$. Unbounded, open and half-open intervals could be incorporated, too. For example, the constraint

$$y = \log(x), \quad x \in \boldsymbol{x} = (-\infty, \infty),\ y \in \boldsymbol{y} = (-\infty, \infty)$$

allows the domain $\boldsymbol{x}$ to contract to $(0, \infty)$.

Among the many applications of constraint solving, let us mention workspace analysis in robotics (Oetomo et al., 2009). The typical problems are to determine the range of end-effector workspace, inverse kinematics, detection of singularities, or design of kinematic machines (Viegas et al., 2017).

8.3 Global optimization

Global optimization aims to find a global optimum (w.l.o.g. a global minimum) of an objective function f on a set described by constraints. Most of the traditional optimization methods converge to a local minimum (Bazaraa et al., 2006; Nocedal and Wright, 2006), which need not be the global one in general.

Algorithm 8.2 Branch & bound algorithm for global optimization

1: $\mathcal{L} \coloneqq \{\boldsymbol{x}^0\}$, $\mathcal{G} \coloneqq \emptyset$, $c^* \coloneqq \infty$,
2: **while** $\mathcal{L} \neq \emptyset$ **do**
3: choose $\boldsymbol{x} \in \mathcal{L}$ and remove $\boldsymbol{x}$ from $\mathcal{L}$,
4: contract $\boldsymbol{x}$,
5: find a feasible point $x \in \boldsymbol{x}$ and update c^*,
6: **if** $\max_i x_i^\Delta > \varepsilon$ **then**
7: split $\boldsymbol{x}$ into sub-boxes and put them into $\mathcal{L}$,
8: **else**
9: put $\boldsymbol{x}$ in $\mathcal{G}$,
10: **end if**
11: **end while**
12: **return** List of boxes $\mathcal{G}$ enclosing all global minimizers and c^* bounding f^* from above.

Stochastic methods try to overcome this drawback by incorporating various heuristics, among others. However, there is no guarantee that they produce a global minimum. In contrast, deterministic global optimization methods are designed to achieve the global minima. Notice that there are problems when computing the global minima is essential; see Floudas (2000); Hiriart-Urruty (2013); Pintér (1996) among others.

Global optimization by interval analysis approach finds verified enclosures for all global minimizers and for the optimal value. On the other hand, due to the curse of dimensionality, the procedure might be time consuming.

The problem. Consider an optimization problem in the form

$$\min \ f(x)$$

subject to

$$\begin{aligned} h_j(x) &= 0, \quad j = 1, \dots, j^*, \\ g_\ell(x) &\leq 0, \quad \ell = 1, \dots, \ell^*, \\ x &\in \boldsymbol{x}^0, \end{aligned}$$

where $\boldsymbol{x}^0 \in \mathbb{IR}^n$ is an initial box. In principle, it is also possible that $\boldsymbol{x}^0$ is the whole space $\mathbb{R}^n$. We will denote by f^* the optimal value. We call the optimization problem *unconstrained* if there are no constraints except for an initial box; thus $j^* = \ell^* = 0$.

The algorithm. The basic branch & bound framework is depicted in Algorithm 8.2. Therein, $\mathcal{L}$ is a list of actual boxes possibly containing the minimizers, and $\varepsilon > 0$ is a threshold for a minimal size of boxes. The resulting list of boxes $\mathcal{G}$ encloses all global minimuzers. The value c^* is a dynamically improving upper bound of the optimal value f^*; the lower bound of f^* is obtained as the lower bound of the objective function on $\mathcal{G}$. We terminate the procedure when $\mathcal{L} = \emptyset$, but basically we can stop it any time and c^* will still provide an upper bound on f^* and $\mathcal{L} \cup \mathcal{G}$ an enclosure of the global minimizers.

In the following sections, we concentrate on particular steps of the algorithm.

Contractions

The idea of contractions in step 4 of Algorithm 8.2 is to shrinken the current box $\boldsymbol{x}$ such that no global minimum is removed (but feasible solutions can and often are removed). In principle, it can remove the whole box $\boldsymbol{x}$, in which case step 4 closes the current iteration of the while loop.

Basically, we can perform consistency contractions on the equations and inequalities constraints as described in Section 8.2.1. Even though this contraction does not utilizes the objective function, it removes the infeasible parts of the initial domain.

Next, we can add the constraint $f(x) \le c^*$ and handle it in the same manner as the other constraints. Thus, for instance, if we evaluate an enclosure $\boldsymbol{f}(\boldsymbol{x})$ to the range $f(\boldsymbol{x})$ and find out that

$$\underline{f}(\boldsymbol{x}) > c^*, \tag{8.4}$$

then $\boldsymbol{x}$ is removed. Nevertheless, the objective function has often different properties than the constraint functions, so using different techniques makes sense. For instance, Martínez et al. (2004) employed the centered form of $f(x)$ to contract $\boldsymbol{x}$ according to the constraint $f(x) \le c^*$.

Unconstrained case. Suppose that $\boldsymbol{x}$ lies in the interior of the feasible set, which is the case when $j^* = 0$ and $\overline{g}_\ell(\boldsymbol{x}) \le 0\ \forall \ell$. There are two optimality conditions for x to be a local minimum: f has zero partial derivatives at x, and f is locally convex in a neighbourhood of x. Since the gradient of a global minimizer must be zero, we apply the interval Newton method or consistency contractors to the additional constraint $\nabla f(x) = 0$. This is one of the most powerful tools in the unconstrained case.

Yet another technique, called the nonconvexity test, can be performed. If f is differentiable but convex in no point in $\boldsymbol{x}$, then there is no local minimum in $\boldsymbol{x}$ and the box is removed. The function f is nonconvex in $\boldsymbol{x}$ if the interval Hessian $\nabla^2 f(\boldsymbol{x})$ contains no positive semidefinite matrix. Positive semidefiniteness of interval matrices was dealt with in Section 5.2.

Constrained case. In the constrained case, we can employ Fritz–John or Karush–Kuhn–Tucker optimality conditions (Hansen and Walster, 2004). The former, which needs no constraint qualification, say that for x to be a local minimum, there must exist $u_0 \in \mathbb{R}$, $u \in \mathbb{R}^{j^*}$ and $v \in \mathbb{R}^{\ell^*}$ such that

$$\begin{aligned} u_0 \nabla f(x) + u^T \nabla h(x) + v^T \nabla g(x) = 0, \\ v \ge 0,\ h(x) = 0,\ g(x) \le 0,\ v^T g(x) = 0, \\ (u_0, u, v) \ne 0. \end{aligned}$$

In particular, the nonlinear system

$$u_0 \nabla f(x) + u^T \nabla h(x) + v^T \nabla g(x) = 0,\ h(x) = 0,\ v_\ell g_\ell(x) = 0\ \forall \ell,\ \|(u_0, u, v)\| = 1$$

has a solution. It represents a system of $n + j^* + \ell^* + 1$ equations with respect to $n + j^* + \ell^* + 1$ variables x, u, v and u_0. The interval Newton method is then used to shrink the box $\boldsymbol{x}$.

Karush–Kuhn–Tucker optimality conditions are stronger, but require additional constraint qualification. We will present the version assuming that $j^* = 0$, function $g_\ell(x)$ is convex for each ℓ and the so-called *Slater condition*

$$\exists x^0 \in \mathbb{R}^n : g(x^0) < 0$$

is satisfied. Then for x to be a local minimum, there must exist $v \in \mathbb{R}^{\ell^*}$ such that

$$\begin{aligned} \nabla f(x) + v^T \nabla g(x) = 0, \\ v \ge 0,\ g(x) \le 0,\ v^T g(x) = 0. \end{aligned}$$

Again, as a consequence, the square nonlinear system of size $n + \ell^*$

$$\nabla f(x) + v^T \nabla g(x) = 0,\ v_\ell g_\ell(x) = 0\ \forall \ell$$

must have a solution.

Feasibility test. The purpose of a feasibility test is to find a feasible solution x and update the upper bound $c^* \ge f^*$ on the optimal value by putting $c^* := \min\{c^*, f(x)\}$. In the unconstrained case, we take $x := x^c$. In the case when there are no equations in the constraints, i.e., $j^* = 0$, then we take $x := x^c$ provided $g(x^c) \le 0$. In the general case, finding a feasible point is a more challenging issue.

A rigorous method to verify feasibility is due to Hansen and Walster (2004); Kearfott (1998). Suppose $j^* < n$, which is a typical situation. We fix $n - j^*$ components $x_i := x_i^c$ so that the equation system

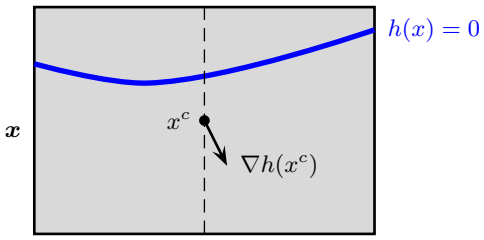


Figure 8.9: Feasibility test with one equation.

consists of j^* equations and j^* variables. We call the interval Newton method which hopefully finds a tiny box $\boldsymbol{y} \subseteq \boldsymbol{x}$ that is guaranteed to contain a feasible point. We then put $c^* := \min\{c^*, \overline{f}(\boldsymbol{y})\}$.

Which components should be fixed? In the case of one equation, $j^* = 1$, Figure 8.9 indicates that it is convenient to fix those components that correspond to the smallest entries of $|\nabla h_1(x^c)|$. If $j^* > 1$, then the recommended heuristic is the following: Transform the Jacobian matrix $\nabla h(x^c)$ to a row echelon form by using complete pivoting and fix the components corresponding to the right most columns. Recall that the complete pivoting takes the largest (by absolute value) element in the matrix as the pivot by considering not only row but also column interchanges.

Handling the boxes

Box selection. There are various strategies for a box selection:

- the oldest one;
- the one with the largest edge, i.e., for which $\max_i x_i^\Delta$ is maximal;
- the one with the smallest upper bound on the objective function, i.e., for which $\underline{f}(\boldsymbol{x})$ is minimal.

List ordering. The box ordering in the list $\mathcal{L}$ and a suitable data structures influence the overall performance of the branch & bound method. Ratz (1994, 2001) recommends to store the boxes in nondecreasing order with respect to $\underline{f}(\boldsymbol{x})$, and secondary in decreasing order with respect to the age of the boxes. Whenever the upper bound c^* is updated, the tail of the list is removed by (8.4). When an appropriate data structure is used, the tail removal is done very efficiently.

Division directions. In step 7 of Algorithm 8.2, the box $\boldsymbol{x}$ is split into two or more sub-boxes. The selection of a number of sub-boxes, the splitting coordinate and the position of splits highly influence the overall performance of the algorithm. Nevertheless, it is not obvious which way is the best one. A simple and popular rule is the circular choice of coordinate directions (Moore, 1979) combined with the bisection, which splits the box into two parts uniformly. However, this need not be always a good strategy.

Example 8.30. Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be defined as $f(x) = x_1^4 + 1000x_2 - 2$ on $\boldsymbol{x} = [-1, 1]^2$. It is easy to see that division along the first coordinate x_1 has very small effect on the resulting image compared to the division along x_2. □

A priori rules. A coordinate $k \in \{1, \ldots, n\}$ to split along can be chosen according the following rules (Floudas and Pardalos, 2009, Csendes: Interval Analysis: Subdivision Directions in Interval Branch and Bound Methods, 1717–1720), (Csendes and Ratz, 1997)

(a) Take the widest edge of $\boldsymbol{x}$, that is,

$$k := \arg \max_{i=1,\ldots,n} x_i^\Delta.$$

(b) (Walster, 1992) Choose a coordinate in which f varies mostly. We approximate it by the value

$$k := \arg \max_{i=1,\ldots,n} \operatorname{rad}(f'_{x_i}(\boldsymbol{x}))x_i^\Delta.$$

(c) (Ratz, 1992) It is similar to the previous one, but uses

$$k \coloneqq \arg\max_{i=1,\dots,n} \operatorname{mag}(f'_{x_i}(\boldsymbol{x}))x_i^{\Delta}.$$

instead.

The numerical tests carried out by Csendes and Ratz (1997) indicate that rule (c) is a good choice. According to the experiments by (Floudas and Pardalos, 2009, Ratschek & Rokne: Interval Global Optimization, 1739–1757), there is no best strategy for splitting, and it is reasonable to combine several of them. Anyway, box selection and splitting strategy influence the overall performance of the algorithm as well as the theoretical convergence properties.

A posteriori rules. Besides the above a priori rules, there are also the so-called a posteriori rules. Provided the dimension is low, one can think of checking all the splitting coordinates and choosing the most promising one. Denote $\boldsymbol{x} = \boldsymbol{x}^{1(i)} \cup \boldsymbol{x}^{2(i)}$ the splitting of $\boldsymbol{x}$ along the i-th coordinate. Out of the diverse heuristics that come into head, Csendes et al. (2000) showed a good behavior of these rules:

- $k \coloneqq \arg\min_{i=1,\dots,n} \operatorname{rad}(\boldsymbol{f}(\boldsymbol{x}^{1(i)}) \cap \boldsymbol{f}(\boldsymbol{x}^{2(i)}))$,
- $k \coloneqq \arg\min_{i=1,\dots,n} \max\{\underline{f}(\boldsymbol{x}^{1(i)}), \underline{f}(\boldsymbol{x}^{2(i)})\}$.

On average, the rules work well, but there are some examples of a strange behavior.

Example 8.31. Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be defined as $f(x) = 10x_1^2 + x_2 - 5$ on $\boldsymbol{x} = [-1,1]^2$. Division along the first coordinate has no effect on the resulting value at the first level, that is, $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x}^{1(1)}) = \boldsymbol{f}(\boldsymbol{x}^{1(2)})$. This would force us to divide still along the second coordinate, which is obviously not the best strategy. □

Despite these examples, practical experiences show that the overall average performance is more influenced by the box selection rule than by the division directions. The position and the number of splits are the least important, so the simple bisection is frequently used.

Other approaches were discussed in Csendes (2001); Nataraj and Sheela (2002); Ratschan (2002); Ratschek and Rokne (1992); Trombettoni et al. (2011).

Convergence

The early works on convergence are Ratschek (1985); Moore and Ratschek (1988). Later, it was studied by Csendes and Ratz (1997); Kearfott (2011), among others.

To state some convergence results, consider the simple unconstrained case

$$\min f(x) \quad \text{subject to} \quad x \in \boldsymbol{x}^0.$$

Suppose that $f(x)$ is continuously differentiable on an open superset of $\boldsymbol{x}^0$. Suppose that we can evaluate it on intervals by using an inclusion isotonic interval extension $\boldsymbol{f}(\boldsymbol{x})$ satisfying

$$\operatorname{rad}(\boldsymbol{f}(\boldsymbol{x}^i)) \to 0 \text{ for any sequence } \boldsymbol{x}^i \in \mathbb{IR}^n,\ i = 1, \dots, \text{ such that } \|\operatorname{rad}(\boldsymbol{x}^i)\| \to 0. \tag{8.5}$$

This is, of course, satisfied for the natural interval extension, the mean value and the slope forms.

Using rule (a), we can state strong convergence properties. The upper bound c^* as well as the lower bound $\min_{\boldsymbol{x}\in\mathcal{L}} \underline{f}(\boldsymbol{x})$ converge to the optimal value. If a box $\boldsymbol{x}$ does not contain any global minimizer, then it is eliminated in a finite number of steps provided the threshold $\varepsilon > 0$ is small enough. Otherwise, under general assumptions, $\mathcal{L}$ converge to the optimal solution set.

To prove it, denote by $\mathcal{G}^*$ the set of all global minima, denote by $\mathcal{L}_k$ the list of boxes in iteration k, and denote by $\mathcal{Z}_k$ the union of boxes in $\mathcal{L}_k$. We will also utilize the assumption that in each iteration k, the upper bound $c_k^* \geq f^*$ is updated as

$$c_{k+1}^* = \min\{c_k^*, f(x^k)\}, \tag{8.6}$$

where $x^k \in \boldsymbol{x}^k$ is arbitrary.

Theorem 8.32 (Ratschek, 1985)**.** *If assumption (8.5) is valid and division rule (a) is used, then* $c_k^* \to f^*$ *and* $\min_{\boldsymbol{x}\in\mathcal{L}_k} \underline{f}(\boldsymbol{x}) \to f^*$ *as* $k \to \infty$.

Proof. Let $\min_{\boldsymbol{x}\in\mathcal{L}_k} \underline{f}(\boldsymbol{x})$ be attained for a box $\boldsymbol{x}^k$, and let $x^k \in \boldsymbol{x}^k$ be the value of this box used to update the upper bound c_k^*. Then $c_k^* \le f(x^k)$. By rule (a), $\lim_{k\to\infty} \boldsymbol{x}^k = \lim_{k\to\infty} x^k$, from which by using (8.5) we derive

$$f^* \le \lim_{k\to\infty} c_k^* \le \lim_{k\to\infty} f(x^k) = \lim_{k\to\infty} \underline{f}(\boldsymbol{x}^k) \le f^*.$$

Therefore, the above inequalities hold as equations. □

Theorem 8.33 (Moore and Ratschek, 1988)**.** *If assumptions (8.5) and (8.6) are valid and division rule (a) is used, then* $\operatorname{dist}(\mathcal{Z}_k, \mathcal{G}^*) \to 0$ *as* $k \to \infty$.

Proof. Denote $\mathcal{Z} := \cap_{k=1}^{\infty} \mathcal{Z}_k$. We will show that $\mathcal{G}^* = \mathcal{Z}$. The inclusion $\mathcal{G}^* \subseteq \mathcal{Z}$ follows from the fact that $\mathcal{Z}^* \subseteq \mathcal{Z}_k$ for every $k = 1, \dots$ due to the inclusion property. To prove the converse inclusion, let $x \in \mathcal{Z}_k$ for every $k = 1, \dots$, and we want to show that $x \in \mathcal{G}^*$. By the assumption, there are the corresponding boxes $\boldsymbol{x}^k$, $k = 1, \dots$, such that $x \in \boldsymbol{x}^k \in \mathcal{L}_k$. Since $(\boldsymbol{x}^k)^\Delta \to_{k\to\infty} 0$, we have $\boldsymbol{f}(\boldsymbol{x}^k) \to_{k\to\infty} f(x)$.

We want to show that $x \in \mathcal{G}^*$, that is, $f(x) = f^*$. Suppose to the contrary that $f(x) > f^*$. From the proof of Theorem 8.32 we know that there is a sequence of upper bounds $c_k^* \ge f^*$, $k = 1, \dots$, such that $c_k^* \to_{k\to\infty} f^*$. That is, there must be k such that $\underline{f}(\boldsymbol{x}^k) > c_k^*$. This means that $\boldsymbol{x}^k$ must have been removed at iteration k; a contradiction. □

For rules (b)–(c), we can state the same convergence properties as long as the derivative $f'(x)$ satisfies (8.5), and $\boldsymbol{f}(\boldsymbol{x})$ is as tight as the mean value form.

Example 8.34 (Csendes and Ratz, 1997)**.** This example shows that the algorithm need not converge in general when the division direction is not chosen appropriately. Let

$$f(x) = x_1 + x_2^2, \quad x \in \boldsymbol{x} = [0,1]^2.$$

The function has a unique minimizer in $\boldsymbol{x}$, which is the point $(0,0)^T$. Using rule (b) for the division direction, we see that always $k = 2$ is chosen since $\operatorname{rad}(f'_{x_1}(\boldsymbol{x})) = 0 < \operatorname{rad}(f'_{x_2}(\boldsymbol{x}))$. Thus, all boxes are divided along the second coordinate. As a consequence, the algorithm yields a list of boxes in the form $[0,1] \times \boldsymbol{x}_2^k$, and therefore it does not converge to the unique minimizer. □

8.3.1 Lower bounds on the optimum

We have already noticed that in the contraction part of the branch & bound scheme it is important to effectively calculate a tight enclosure $\boldsymbol{f}(\boldsymbol{x})$ to the range of f over a box $\boldsymbol{x}$. In particular, the lower bound $\underline{f}(\boldsymbol{x})$ is of interest. Thus, besides the techniques studied in Chapter 6, we concentrate more on the lower bound now.

Lipschitz continuity approach. A lower bound can be easily obtained by utilizing Lipschitz continuity of f. The function f is Lipschitz continuous on $\boldsymbol{x}$ if there is $L > 0$ such that $|f(x) - f(y)| \le L\|x - y\|$ for every $x, y \in \boldsymbol{x}$. Putting $y := x^c$ we get

$$f(x) \ge f(x^c) - L\|x - x^c\| \ge f(x^c) - L\|x^\Delta\|.$$

Thus, we have

$$\underline{f}(\boldsymbol{x}) \ge f(x^c) - L\|x^\Delta\|.$$

Nevertheless, it need not be easy to determine or approximate the Lipschitz constant L from above.

Lower bounds on the objective function can also be found by a linearization (Tawarmalani and Sahinidis, 2005), or by a convex underestimation, which is described in the sequel.

Convex underestimators

Instead of computing a lower bound of $f(x)$ on a box $\boldsymbol{x}$, we can also estimate $f(x)$ from below by a more simple function $g(x)$. To this end, we usually seek for a linear or, more generally, a convex function.

A convex underestimator of $f(x)$ on $\boldsymbol{x}$ is a function $g(x)$ that is convex and $g(x) \le f(x)$ for every $x \in \boldsymbol{x}$. Determining the minimum of $g(x)$ on $\boldsymbol{x}$ is then an easy task, and the minimum serves as a lower bound to $\underline{f}(\boldsymbol{x})$.

Convex underestimator for bilinear terms. A convex underestimator can be computed along the following lines. If f is a sum of simpler functions, then we process the individual terms separately. Linear terms can be left out. Bilinear terms are approximated from below by the following linearization.

Proposition 8.35 (McCormick, 1976; Al-Khayyal and Falk, 1983)**.** *For every $y \in \boldsymbol{y} \in \mathbb{IR}$ and $z \in \boldsymbol{z} \in \mathbb{IR}$ we have*

$$yz \ge \max\{\underline{y}z + \underline{z}y - \underline{y}\underline{z},\ \overline{y}z + \overline{z}y - \overline{y}\overline{z}\}. \tag{8.7}$$

Moreover, it is the tightest convex underestimator of yz.

Proof. Write

$$yz \ge yz - (y - \underline{y})(z - \underline{z}) = \underline{y}z + \underline{z}y - \underline{y}\underline{z},$$
$$yz \ge yz - (\overline{y} - y)(\overline{z} - z) = \overline{y}z + \overline{z}y - \overline{y}\overline{z}.$$

It is easy to see that the first linear function $\underline{y}z + \underline{z}y - \underline{y}\underline{z}$ touches the bilinear function yz at the left and the bottom boundary of the rectangle $\boldsymbol{y} \times \boldsymbol{z}$, that is, at the segments $(\underline{y}, \underline{z}) - (\underline{y}, \overline{z})$ and $(\underline{y}, \underline{z}) - (\overline{y}, \underline{z})$. Similarly, the second linear function $\overline{y}z + \overline{z}y - \overline{y}\overline{z}$ touches the bilinear function yz at the right and the top boundary of the rectangle $\boldsymbol{y} \times \boldsymbol{z}$. The tightest convex underestimator of yz can be expressed as a pointwise maximum of all linear underestimators of yz. However, each linear underestimator of yz must underestimate it on both triangles $(\underline{y}, \underline{z}) - (\underline{y}, \overline{z}) - (\overline{y}, \underline{z})$ and $(\overline{y}, \overline{z}) - (\underline{y}, \overline{z}) - (\overline{y}, \underline{z})$. The first linear function in (8.7) touches the vertices of the first triangle and similarly for the second one. Therefore, each linear underestimator has to lie below their maximum. □

Thus, the bilinear term yz can be linearized by any of the two linear functions. By taking their maximum, we have the tightest approximation.

Convex underestimator for general terms. To underestimate the general terms, suppose that f is twice continuously differentiable and put $g(x) \coloneqq f(x) + \sum_{i=1}^{n} \alpha_i (x_i - \underline{x}_i)(x_i - \overline{x}_i)$, where α_i, $i = 1, \dots, n$ are nonnegative parameters to determine. Then $g(x) \le f(x)$ on $\boldsymbol{x}$. Compactly, we write

$$g(x) = f(x) + (x - \underline{x})^T \operatorname{diag}(\alpha)(x - \overline{x}).$$

It is known that $g(x)$ is convex as long as its Hessian $\nabla^2 g(x) = \nabla^2 f(x) + 2\operatorname{diag}(\alpha)$ is positive semidefinite for every $x \in \boldsymbol{x}$. Thus, we have to determine the vector of parameters α such that the interval matrix $\nabla^2 f(\boldsymbol{x}) + 2\operatorname{diag}(\alpha)$ is positive semidefinite. To simplify it, all parameters are supposed to be equal, $\beta \coloneqq \alpha_1 = \cdots = \alpha_n$. The interval matrix $\nabla^2 f(\boldsymbol{x}) + 2\beta I_n$ is positive semidefinite if and only if the smallest eigenvalue is non-negative, or equivalently, if and only if the smallest eigenvalue of $\nabla^2 f(\boldsymbol{x})$ is greater than or equal to -2β. So we can put

$$\beta \coloneqq -\frac{1}{2}\underline{\lambda}_n(\nabla^2 f(\boldsymbol{x})).$$

Eigenvalues of (symmetric) interval matrices were discussed in Section 5.1, so we can utilize some of the methods to compute either the eigenvalue exactly or its lower bound.

Example 8.36. Consider the function from Hladík (2016a); Skjäl et al. (2012)

$$f(x_1, x_2) = (2x_1 + x_2 - 3)^2 + (x_1 x_2 - 1)^2$$

on the box $\boldsymbol{x} = [0, 4]^2$. Figure 8.10 from Hladík (2016a) illustrates the function and its convex underestimator computed by the described method with $\alpha = (21, 24)^T$. □

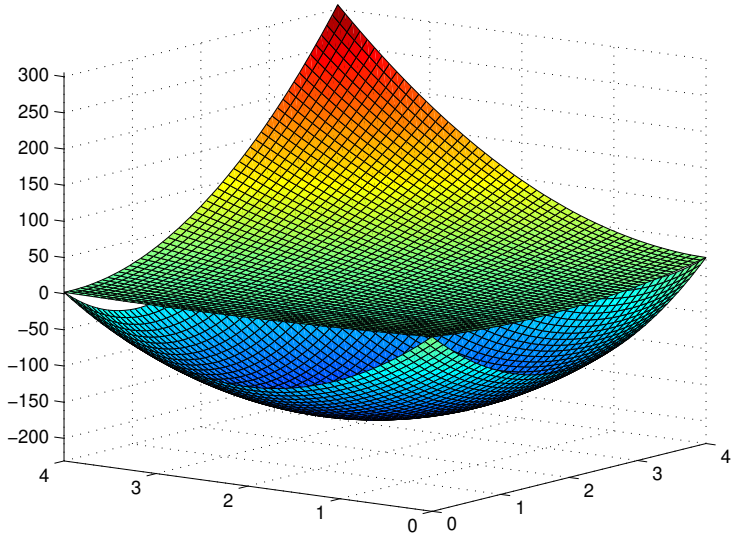


Figure 8.10: (Example 8.36) The function $f(x)$ and its convex underestimator.

The above described method for computing a convex underestimator of a general function was used in the αBB algorithm (Akrotirianakis and Floudas, 2004; Androulakis et al., 1995; Floudas and Pardalos, 2009), which is a deterministic global optimization algorithm based on a branch & bound framework.

There are also other techniques to construct convex enclosures. One of them, a relaxation to interval linear functions, will be discussed in Section 9.4.2.

Notes and further reading (Section 8.3).

The first approaches to deterministic global optimization include Moore (1966); Skelboe (1974), but a more systematic study of global optimization using interval analysis dates about to Hansen (1980). The books devoted to the interval approach to global optimization are, e.g., Floudas (2000); Floudas and Pardalos (2009); Hansen and Walster (2004); Hendrix and Gazdag-Tóth (2010); Kearfott (1996a); Neumaier (2004); Ratschek and Rokne (2007), and recent survey papers include Kearfott (2011). Eight global optimization solvers were numerically compared in Neumaier et al. (2005). Due to the nature of the interval approach to global optimization, a parallel computation is very appropriate. One of the early parallelization implementations is due to Henriksen and Madsen (1992); Madsen (1991).

Test data. A large set of constrained global optimization problems is collected in books Floudas and Pardalos (1990); Floudas et al. (1999). Concerning journal papers, Ratz (2001) includes some smooth and also non-smooth unconstrained examples. Internet databases of examples and benchmarks include

- Floudas et al. (1999) is supplemented by the web pages of test problems in `GAMS` modeling language
  `http://titan.princeton.edu/TestProblems/`
- *COCONUT benchmarks*
  more than 1000 benchmarks for global optimization and constraint satisfaction problems in `AMPL`, `GAMS` and `DAG` format
  `http://www.mat.univie.ac.at/~neum/glopt/coconut/Benchmark/Benchmark.html`
- *Neumaier's web page*
  many global optimization links
  `http://www.mat.univie.ac.at/~neum/glopt/test.html`
- *COPRIN examples*
  benchmark data for constraint satisfaction, optimization and robot-related problems
  `http://www-sop.inria.fr/teams/coprin/logiciels/ALIAS/Benches/benches.html`

Software. Rigorous global optimization software include

- *GlobSol* (by R.B. Kearfott)
  written in Fortran 95, open-source, there exist conversions from AMPL and GAMS representations,
  `https://interval.louisiana.edu/kearfott.html#software`
- *COCONUT Environment*
  open-source C++ classes
  `http://www.mat.univie.ac.at/~coconut/coconut-environment/`
- *GLOBAL* (by T. Csendes)
  for Matlab / Intlab, free for academic purposes
  `http://www.inf.u-szeged.hu/~csendes/linkek_en.html`
- *IBEX* (by G. Chabert, B. Neveu, J. Ninin and others),
  an open-source interval C++ library,
  mainly for system solving and global optimization,
  `http://www.ibex-lib.org/`
- *PROFIL / BIAS* (by O. Knüppel et al.)
  free C++ class
  `http://www.ti3.tu-harburg.de/Software/PROFILEnglisch.html`

For other global optimization software links, see also

- *C.A. Floudas* (`http://titan.princeton.edu/tools/`)
- *A. Neumaier* (`http://www.mat.univie.ac.at/~neum/glopt.html`)

Global optimization solvers using interval computation, but not in a numerically rigorous way (do not handle round-off errors)

- *BARON (Branch-And-Reduce Optimization Navigator, by N.V. Sahinidis et al.)*
  commercial system, available under GAMS and AIMMS
  `http://archimedes.cheme.cmu.edu/?q=baron`
- *ANTIGONE (by R. Misener and C.A. Floudas)*
  deterministic global optimization of mixed-integer nonlinear programs,
  available through GAMS and Princeton University
  `http://ares.tamu.edu/ANTIGONE/`
- *Couenne (Convex Over and Under ENvelopes for Nonlinear Estimatio, by P. Belotti)*
  open source branch & bound algorithm for solving mixed-integer nonlinear programming problems,
  `http://www.coin-or.org/Couenne/`

Notice that the approaches described by Tawarmalani and Sahinidis (2005) and implemented in the global optimization package BARON received the 2006 Beale-Orchard-Hays Prize awarded by Mathematical Programming Society. As written in the laudatio, "... BARON also incorporates techniques from automatic differentiation, interval arithmetic, and other areas to yield an automatic, modular, and relatively efficient solver for the very difficult area of global optimization."

Chapter 9

Interval linear programming

This chapter is devoted to optimization problems with interval coefficients. We consider only linear objective function and linear constraints, which is called linear programming; the nonlinear models are out of the main focus of this book.

The history of interval linear programming goes back to 1970's; see Beeck (1978); Krawczyk (1975); Machost (1970). Since then, it has been a steady increase in research and development of the discipline; see survey papers Hladík (2012a); Rohn (2006b).

Linear programming. Consider a linear programming (LP) problem

$$\min\ c^T x \ \text{ subject to } \ Ax = b,\ x \geq 0, \tag{9.1}$$

where $c \in \mathbb{R}^n$ is the objective vector, $A \in \mathbb{R}^{m\times n}$ is the constraint matrix and $b \in \mathbb{R}^m$ the corresponding right-hand side. An LP problem can also be expressed in the form

$$\min\ c^T x \ \text{ subject to } \ Ax \leq b \tag{9.2}$$

or in the form

$$\min\ c^T x \ \text{ subject to } \ Ax \leq b,\ x \geq 0. \tag{9.3}$$

All these forms are equivalent, so usually only one of them is considered as a canonical form. As we will see later, this is not the case when the coefficients are interval-valued. The particular forms are no mnore equivalent and we have to study them individually. However, for the sake of simplicity of exposition, we consider (9.1) the basic form, and the distinctions are pointed out.

Interval linear programming. Let $\boldsymbol{A} \in \mathbb{IR}^{m\times n}$, $\boldsymbol{b} \in \mathbb{IR}^m$ and $\boldsymbol{c} \in \mathbb{IR}^n$. An *interval LP problem* is the family of LP problems (9.1) with $A \in \boldsymbol{A}$, $b \in \boldsymbol{b}$ and $c \in \boldsymbol{c}$. Analogously for (9.2) and (9.3). In short, we write it as

$$\min\ \boldsymbol{c}^T x \ \text{ subject to } \ \boldsymbol{A}x = \boldsymbol{b},\ x \geq 0. \tag{9.4}$$

By a *realization*, we mean a particular selection of the interval values, that is, an arbitrary LP problem with input data $A \in \boldsymbol{A}$, $b \in \boldsymbol{b}$ and $c \in \boldsymbol{c}$.

In contrast to the real LP, there is no natural definition of an optimal value or an optimal solution. Indeed, we do not define what is an optimal value and an optimal solution.

We do not solve interval LP problems at all!

The purpose of interval linear programming is to study effects of variations of the coefficients in their interval domains, and we are interested how the variations influence the optimal value and optimal solutions. And not only that; we also want to know if there are some characteristic stable under the data variations. We ask questions such as: Is there a basis optimal for each realization of intervals? Does there exists an optimum for each realization? Is the LP problem bounded for each realization? etc.

In the next sections, we will address some of these questions. We will heavily make use of the theory and methods of interval linear systems (Chapters 3–4).

9.1 Optimal value range

The optimal value function belongs to the most studied problems in interval LP. Denote by $f(A,b,c)$ the optimal value of (9.1). By convention, we define $f(A,b,c) = -\infty$ if the problem is unbounded and $f(A,b,c) = \infty$ if the problem is infeasible.

In accordance with the definition (2.4), the range of optimal values is defined as

$$f(\boldsymbol{A},\boldsymbol{b},\boldsymbol{c}) = \{f(A,b,c);\ (A,b,c) \in (\boldsymbol{A},\boldsymbol{b},\boldsymbol{c})\}.$$

The best case and *the worst case* optimal values are the, respectively, the minimum and maximum of the range,

$$\begin{aligned} \underline{f} &\coloneqq \min\ f(A,b,c)\ \text{ subject to }\ (A,b,c) \in (\boldsymbol{A},\boldsymbol{b},\boldsymbol{c}),\\ \overline{f} &\coloneqq \max\ f(A,b,c)\ \text{ subject to }\ (A,b,c) \in (\boldsymbol{A},\boldsymbol{b},\boldsymbol{c}). \end{aligned}$$

The optimal value $f(A,b,c)$ of any realization then satisfies $f(A,b,c) \in [\underline{f},\overline{f}]$.[1)]

The best case optimal value $\underline{f}$ is efficiently computable by solving one suitable linear program. Its formulation was stated in the early papers Beeck (1978); Machost (1970); Rohn (1976).

Theorem 9.1. *We have*

$$\underline{f} = \min\ \underline{c}^T x \quad \textit{subject to}\quad \underline{A}x \le \overline{b},\ -\overline{A}x \le -\underline{b},\ x \ge 0. \tag{9.5}$$

Proof. Using Corollary 3.7, we can write

$$\begin{aligned} \underline{f} &= \min\ c^T x\ \text{ subject to }\ Ax = b,\ x \ge 0,\ (A,b,c) \in (\boldsymbol{A},\boldsymbol{b},\boldsymbol{c})\\ &= \min\ \underline{c}^T x\ \text{ subject to }\ Ax = b,\ x \ge 0,\ (A,b) \in (\boldsymbol{A},\boldsymbol{b})\\ &= \min\ \underline{c}^T x\ \text{ subject to }\ \underline{A}x \le \overline{b},\ -\overline{A}x \le -\underline{b},\ x \ge 0. \end{aligned}$$

□

In contrast to $\underline{f}$, computation the worst case optimal value $\overline{f}$ is more difficult. Below, we present a reduction to 2^m LP realizations, so the complexity grows exponentially w.r.t. the number of equations.

Theorem 9.2 (Rohn, 1984)**.** *We have*

$$\overline{f} = \max_{s\in\{\pm1\}^m}\ f(A_{s,e}, b_s, \overline{c}). \tag{9.6}$$

Proof. "$\ge$" This inequality is obvious since the right-hand side takes the maximum of optimal values of 2^m instances of the ILP problem.

"$\le$" Let $A \in \boldsymbol{A}$, $b \in \boldsymbol{b}$ and $c \in \boldsymbol{c}$ be arbitrary. It is sufficient to show that $f(A,b,c) \le \alpha$, where

$$\alpha \coloneqq \max_{s\in\{\pm1\}^m} f(A_{s,e}, b_s, \overline{c}).$$

If $f(A,b,c) = -\infty$, then the inequality obviously holds. If $f(A,b,c) = \infty$, then the corresponding LP problem is infeasible. Thus, the interval system $\boldsymbol{A}x = \boldsymbol{b}$, $x \ge 0$ is not strongly solvable. By Theorem 4.18, there is $s \in \{\pm1\}^m$ such that the system

$$A_{s,e}x = b_s,\quad x \ge 0 \tag{9.7}$$

is unsolvable, whence $\alpha = \infty$.

It remains to analyse the case with $f(A,b,c)$ finite. Let y^* be an optimal solution of the dual problem

$$\max\ b^T y\ \text{ subject to }\ A^T y \le c. \tag{9.8}$$

[1)] Herein, we allow unbounded intervals such as $[2,\infty]$ or $[-\infty,5]$, but we believe no confusion emerge.

By the duality theorem (Theorem 1.32), we have $f(A,b,c) = b^T y^*$. Denote $s^* \coloneqq \operatorname{sgn}(y^*)$. Since by Proposition 2.11

$$b^T y^* \le (b^c)^T y^* + (b^\Delta)^T |y^*| = (b^c + \operatorname{diag}(s^*) b^\Delta)^T y^* = b_{s^*}^T y^*,$$
$$(A_{s^*,e})^T y^* = (A^c - \operatorname{diag}(s^*) A^\Delta)^T y^* = (A^c)^T y^* - (A^\Delta)^T |y^*| \le A^T y^* \le c \le \overline{c},$$

the optimal value of (9.8) is less then or equal to the optimal value of the LP problem

$$\max\ b_{s^*}^T y \ \text{ subject to } \ (A_{s^*,e})^T y \le \overline{c}.$$

This is because y^* remains feasible and has not less objective value. Moreover, by the duality theorem, the optimal value of this LP problem is $f(A_{s^*,e}, b_{s^*}, \overline{c})$. Therefore

$$f(A,b,c) = b^T y^* \le f(A_{s^*,e}, b_{s^*}, \overline{c}) \le \alpha. \qquad \square$$

The exponential number of instances in the above reduction is justified by intractability of computation $\overline{f}$. *NP*-hardness was proved in Rohn (1995, 1997, 2006b) and strong *NP*-hardness in (Gabrel et al., 2008, 2010).

Theorem 9.3. *It is co-NP-hard to check if $\overline{f} < \infty$. This is true even for problems with A and c real.*

Proof. We first observe that condition $\overline{f} < \infty$ is equivalent to strong feasibility of the constraints; the realizations addressed in Theorem 9.2 are the same as the realizations from Theorem 4.18. Now, by Theorem 4.26, checking strong solvability of the interval linear system $Ax = \boldsymbol{b}$, $x \ge 0$ is *co-NP*-hard. Therefore, deciding whether the interval LP problem

$$\min\ 0^T x \ \text{ subject to } \ Ax = \boldsymbol{b},\ x \ge 0$$

has all optimal values finite (and zero) is *co-NP*-hard, too. $\square$

The value $\underline{f}$ need not be attained for a particular realization provided $\underline{f} = -\infty$; see Example 9.4 below. However, if $\underline{f} > -\infty$, then it is always attained. In this case, it is easy to construct a realization $(A,b,c) \in (\boldsymbol{A},\boldsymbol{b},\boldsymbol{c})$ such that $f(A,b,c) = \underline{f}$. Basically, one can proceed as in the proof of Proposition 3.3.

In contrast, the worst case optimal value $\overline{f}$ is always attained. By Theorem 9.2, it is attained for the realization $(A,b,c) = (A_{s,e}, b_s, \overline{c})$, where $s \in \{\pm 1\}^m$.

Example 9.4 (Rohn, 2006b)**.** Consider the interval LP problem

$$f(a) = \min\ -x \ \text{ subject to } \ ax = 1,\ x \ge 0,$$

where $a \in \boldsymbol{a} = [0,1]$. For $a > 0$, the optimum is $x = 1/a$ and the optimal value $f(a) = -1/a$. For $a = 0$, the problem is infeasible. Thus, $\underline{f} = -\infty$, $\overline{f} = \infty$ and the set of optimal values is $f(\boldsymbol{a}) = (-\infty, -1] \cup \{\infty\}$. While $\overline{f}$ is attained for $a = 0$, the value $\underline{f}$ is attained for no $a \in \boldsymbol{a}$. $\square$

Example 9.5 (modification of the Bereanu example, see Beeck, 1978)**.** Consider the interval LP problem

$$f(a) = \min\ -x \ \text{ subject to } \ x \le 1,\ ax \le 0,\ x \ge 0.$$

where $a \in \boldsymbol{a} = [-1,1]$. For $a > 0$, the optimum is $x = 0$ and the optimal value $f(a) = 0$. For $a \le 0$, the optimum is $x = 1$ and the optimal value $f(a) = -1$. Thus, $\underline{f} = -1$, $\overline{f} = 0$ and the set of optimal values is $f(\boldsymbol{a}) = \{-1\} \cup \{0\}$. Therefore, there can be gaps in the set $f(\boldsymbol{a})$ and not all values in $f(\boldsymbol{a})$ are attained for particular realizations, even when both the best and worst case optimal values are finite. $\square$

The above example also illustrates that the optimal value function $f(a)$ need not be continuous. Continuity can be achieved under certain assumptions, e.g., in the basis stable case; see Section 9.3. Another, quite general, condition was stated in Mostafaee et al. (2016).

Inequality constrained LP form. Consider an interval LP problem in the form (9.2),

$$\min\ \boldsymbol{c}^T x\ \text{ subject to }\ \boldsymbol{A}x \le \boldsymbol{b}.$$

Now, the situation is the opposite. The best case optimal value $\underline{f}$ is hard to compute, but the worst case optimal value $\overline{f}$ is calculated by solving a certain LP problem.

Theorem 9.6. *We have*

$$\underline{f} = \min\ (c^c)^T x - (c^\Delta)^T|x| \ \ \textit{subject to}\ \ A^c x - A^\Delta|x| \le \overline{b} \tag{9.9}$$

$$= \min_{s\in\{\pm1\}^n}\ f\big(A_{e,s}, \overline{b}, c_{-s}\big), \tag{9.10}$$

$$\overline{f} = \min\ \overline{c}^T x^1 - \underline{c}^T x^2 \ \ \textit{subject to}\ \ \overline{A}x^1 - \underline{A}x^2 \le \underline{b},\ x^1, x^2 \ge 0. \tag{9.11}$$

Proof. "Case $\underline{f}$." Using Gerlach's Theorem 4.1, we can write

$$\begin{aligned}\underline{f} &= \min\ c^T x\ \text{ subject to }\ Ax \le b,\ (A,b,c) \in (\boldsymbol{A},\boldsymbol{b},\boldsymbol{c})\\ &= \min\ \Big(\min_{c\in\boldsymbol{c}} c^T x\Big)\ \text{ subject to }\ Ax \le b,\ (A,b) \in (\boldsymbol{A},\boldsymbol{b})\\ &= \min\ (c^c)^T x - (c^\Delta)^T|x|\ \text{ subject to }\ A^c x - A^\Delta|x| \le \overline{b}.\end{aligned}$$

The formula (9.10) is derived by using the orthant-by-orthant decomposition approach. For $s \in \{\pm1\}^n$, each vector lying in the corresponding orthant $\operatorname{diag}(s)x \ge 0$ satisfies $|x| = \operatorname{diag}(s)x$. Thus, we can write

$$\begin{aligned}\underline{f} &= \min\ (c^c)^T x - (c^\Delta)^T|x|\ \text{ subject to }\ A^c x - A^\Delta|x| \le \overline{b}\\ &= \min_{s\in\{\pm1\}^n}\ \min\ (c^c)^T x - (c^\Delta)^T|x|\ \text{ subject to }\ A^c x - A^\Delta|x| \le \overline{b},\ \operatorname{diag}(s)x \ge 0\\ &= \min_{s\in\{\pm1\}^n}\ \min\ (c^c)^T x - (c^\Delta)^T \operatorname{diag}(s)x\ \text{ subject to }\ A^c x - A^\Delta \operatorname{diag}(s)x \le \overline{b},\ \operatorname{diag}(s)x \ge 0\\ &= \min_{s\in\{\pm1\}^n}\ \min\ c_{-s}^T x\ \text{ subject to }\ A_{e,s}x \le \overline{b},\ \operatorname{diag}(s)x \ge 0\\ &= \min_{s\in\{\pm1\}^n}\ f\big(A_{e,s}, \overline{b}, c_{-s}\big).\end{aligned}$$

Notice that we can leave out the constraint $\operatorname{diag}(s)x \ge 0$ in the above derivation for the same reason as mentioned Remark 3.10.

"Case $\overline{f}$." If the interval system $\boldsymbol{A}x \le \boldsymbol{b}$ is not strongly feasible, then (9.11) is valid simply because both sides are equal to $+\infty$; the constraints in (9.11) are those that appear in Theorem 4.28.

Hence we can assume now that the interval system $\boldsymbol{A}x \le \boldsymbol{b}$ is strongly feasible. By the duality theory in linear programming, we can replace the LP problem by its dual. Proceeding then similarly as in Theorem 9.1,

$$\begin{aligned}\overline{f} &= \max\ f(A,b,c)\ \text{ subject to }\ (A,b,c) \in (\boldsymbol{A},\boldsymbol{b},\boldsymbol{c})\\ &= \max\ b^T y\ \text{ subject to }\ A^T y = c,\ y \le 0,\ (A,b,c) \in (\boldsymbol{A},\boldsymbol{b},\boldsymbol{c})\\ &= \max\ \underline{b}^T y\ \text{ subject to }\ A^T y = c,\ y \le 0,\ (A,c) \in (\boldsymbol{A},\boldsymbol{c})\\ &= \max\ \underline{b}^T y\ \text{ subject to }\ \overline{A}^T y \le \overline{c},\ -\underline{A}^T y \le -\underline{c},\ y \le 0\\ &= \min\ \overline{c}^T x^1 - \underline{c}^T x^2\ \text{ subject to }\ \overline{A}x^1 - \underline{A}x^2 \le \underline{b},\ x^1, x^2 \ge 0.\end{aligned}$$

□

The value $\underline{f}$ can be computed by solving the nonconvex and nonsmooth optimization problem (9.9). The alternative formula (9.10) reduces the problem to 2^n LP realizations, so it is exponential w.r.t. the number of variables. Since the constraints in (9.9) characterize the set of weak solutions of $\boldsymbol{A}x \le \boldsymbol{b}$, we immediately get from Theorem 4.4 the *NP*-hardness of computing $\underline{f}$. The result is true even for problems with A and c real (Gabrel et al., 2008, 2010; Rohn, 1995, 1997).

Theorem 9.7. *It is NP-hard to check if* $\underline{f} < \infty$.

Inequality constrained LP form with nonnegativity. Eventually, we consider an interval LP problem in the form (9.3),

$$\min\ \boldsymbol{c}^T x\ \text{ subject to }\ \boldsymbol{A}x \le \boldsymbol{b},\ x \ge 0.$$

From the point of view of computing the range of optimal values, this is the most easy case. By direct inspection, both the best and worst case optimal values are attained for obvious realizations.

Theorem 9.8 (Vajda, 1961)**.** *We have* $\underline{f} = f(\underline{A}, \overline{b}, \underline{c})$ *and* $\overline{f} = f(\overline{A}, \underline{b}, \overline{c})$.

Notes and further reading. One can think of a general interval LP problem involving equations and inequalities, and involving both free and nonnegative variables. That is, an LP problem constrained by a system (4.3). The formulae for the optimal value range of such a general model were derived in Král (2020). The general case was also presented by Chinneck and Ramadan (2000) in a different fashion.

Since $\underline{f}$ and $\overline{f}$ are hard to compute in general, it makes sense to approximate these values. Various lower and upper bounds were proposed in Hladík (2014c); Mohammadi and Gentili (2019).

If we are focused on feasible instances of an interval LP problem (typically in transportation problems), then we wish to find the worst case finite optimal value instead of the worst case optimal value (which might be $+\infty$ due to infeasibility). This issue was addressed in Hladík (2018b) and for the interval transportation problem in Garajová et al. (2020); Garajová and Rada (2021).

Duality in interval linear programming was investigated by Novotná et al. (2020), among others.

9.2 The optimal solution set

Denote by $\mathcal{S}(A, b, c)$ the set of optimal solutions of an LP realization $(A, b, c) \in (\boldsymbol{A}, \boldsymbol{b}, \boldsymbol{c})$. Then *the optimal solution set* is defined

$$\mathcal{S} \coloneqq \bigcup_{A \in \boldsymbol{A},\, b \in \boldsymbol{b},\, c \in \boldsymbol{c}} \mathcal{S}(A, b, c).$$

That is, the optimal solution set comprises all optima of all realizations. Many questions related to $\mathcal{S}$ are hard, and handling this set is one of the most challenging issues in interval linear programming.

In general, $\mathcal{S}$ can be nonconvex and even disconnected.

Example 9.9 (Example 9.5 continued)**.** Consider again the interval LP problem

$$f(a) = \min\ -x\ \text{ subject to }\ x \le 1,\ ax \le 0,\ x \ge 0.$$

where $a \in \boldsymbol{a} = [-1, 1]$. For $a > 0$, the optimum is $x = 0$. For $a \le 0$, the optimum is $x = 1$. Therefore, the optimal solutions set consists of two isolated points, $\mathcal{S} = \{0\} \cup \{1\}$. □

The optimal solution set $\mathcal{S}$ has the form of a convex polyhedron and is easy to describe only in special situations. One of such situations is basis stability, discussed in Section 9.3.

In the following two paragraphs, we present two approaches to characterize $\mathcal{S}$. Both of them also provide a method to find an enclosure. Computing the interval hull $\square\mathcal{S}$ exactly, however, is a strongly *NP*-hard problem, even of intervals are situated in the vector $\boldsymbol{b}$ only (Garajová and Hladík, 2019). It is also *NP*-hard to check if a given point $x \in \mathbb{R}^n$ belongs to $\mathcal{S}$ (Rada et al., 2019).

Characterization by duality. In order to characterize the optimal solution set, we make use of the optimality conditions in linear programming. Throughout this section, we consider the interval LP problem in the form (9.4).

Theorem 9.10. *The optimal solution set* $\mathcal{S}$ *is characterized as the solution set of parametric interval system (more precisely, its projection to* x*-variables)*

$$Ax = b,\quad x \ge 0,\quad A^T y \le c,\quad c^T x = b^T y,\quad (A, b, c) \in (\boldsymbol{A}, \boldsymbol{b}, \boldsymbol{c}).$$

Proof. Let $(A,b,c) \in (\boldsymbol{A},\boldsymbol{b},\boldsymbol{c})$. Then $x \in \mathbb{R}^n$ is optimal for the corresponding realization if and only if there is $y \in \mathbb{R}^m$ such that

$$Ax = b, \quad x \geq 0, \quad A^T y \leq c, \quad c^T x = b^T y.$$

This is a well known optimality condition, claiming that x is a feasible primal solution, y is a feasible dual solution, and the primal and dual objectives are equal. □

Parametric interval systems were addressed in Section 3.9, and they are hard to deal with. A basic way to tackle them is to relax dependencies and reduce the system to an ordinary interval linear system

$$\boldsymbol{A}x = \boldsymbol{b}, \quad x \geq 0, \quad \boldsymbol{A}^T y \leq \boldsymbol{c}, \quad \boldsymbol{c}^T x = \boldsymbol{b}^T y.$$

This is still computationally intractable problem, but the theory from Section 4.3 enables to characterize the solution set and provides tools to handle it if the dimension is small. In particular, Theorem 4.13 implies that $\mathcal{S}$ lies in the set described by the system

$$\begin{aligned}
&\underline{A}x \leq \overline{b}, \ -\overline{A}x \leq -\underline{b}, \ x \geq 0, \\
&(A^c)^T y - (A^\Delta)^T |y| \leq \overline{c}, \\
&|(c^c)^T x - (b^c)^T y| \leq (c^\Delta)^T x + (b^\Delta)^T |y|.
\end{aligned}$$

Characterization by complementarity. Another method to characterize $\mathcal{S}$ and to compute its approximation is based on optimality conditions involving complementarity (Garajová and Hladík, 2019). A vector $x \in \mathbb{R}^n$ is optimal if there exists $y \in \mathbb{R}^m$ such that

$$Ax = b, \quad x \geq 0, \quad A^T y \leq c, \quad x^T(c - A^T y) = 0,$$

This gives rise to the following charactrization of $\mathcal{S}$ by means of a parametric interval system.

Theorem 9.11. *The optimal solution set $\mathcal{S}$ is characterized as the solution set of parametric interval system (more precisely, its projection to x-variables)*

$$Ax = b, \quad x \geq 0, \quad A^T y \leq c, \quad x^T(c - A^T y) = 0, \quad (A,b,c) \in (\boldsymbol{A},\boldsymbol{b},\boldsymbol{c}). \tag{9.12}$$

The complementary condition $x^T(c - A^T y) = 0$ means that for every $i \in \{1,\dots,n\}$ we have $x_i = 0$ or $(c - A^T y)_i = 0$. Let $i \subseteq \{1,\dots,n\}$ be the set of indices for which the former is true and let $J := \{1,\dots,n\} \setminus I$ be its complement. Then (9.12) reduced to the parametric interval system

$$A_I x_I = b, \quad x_I \geq 0, \quad (A_J)^T y \leq c_J, \quad (A_I)^T y = c_I, \quad (A,b,c) \in (\boldsymbol{A},\boldsymbol{b},\boldsymbol{c}). \tag{9.13}$$

The obvious drawback of this approach is that it is explicitly exponential as there are 2^n possibilities for I. On the other hand, system (9.13) is less affected by dependencies. Indeed, only matrix A_I occurs twice there. Thus, the relaxation

$$\boldsymbol{A}_I x_I = \boldsymbol{b}, \quad x_I \geq 0, \quad (\boldsymbol{A}_J)^T y \leq \boldsymbol{c}_J, \quad (\boldsymbol{A}_I)^T y = \boldsymbol{c}_I.$$

is supposed to produce minor overestimation. Moreover, this interval system can be spit to two independent sub-systems

$$\boldsymbol{A}_I x_I = \boldsymbol{b}, \quad x_I \geq 0$$

and

$$(\boldsymbol{A}_J)^T y \leq \boldsymbol{c}_J, \quad (\boldsymbol{A}_I)^T y = \boldsymbol{c}_I.$$

If $\boldsymbol{A}$ is real (its radius is zero), then there are no dependencies in (9.13) and the relaxation produces no overestimation.

Notes and further reading. Garajová and Hladík (2019) analysed topological and geometric properties of the optimal solution set $\mathcal{S}$, including closedness, polyhedrality, convexity, connectedness and boundedness. Hladík (2012c) proposed a contractor for $\mathcal{S}$, that is, an iterative method to make an enclosure of $\mathcal{S}$ tighter. Hladík (2020b); Jansson and Rump (1991) studied an inner approximation of the optimal solution set $\mathcal{S}$.

9.3 Basis stability

Basis stability refers to the situation, where there is a basis that is optimal for each realization. If this situation happens, many problems become much easier to solve. Throughout this section, we consider the interval LP problem (9.4).

Basis feasibility and optimality. Consider a fixed LP realization $(A, b, c) \in (\boldsymbol{A}, \boldsymbol{b}, \boldsymbol{c})$. Recall that *a basis* is an index set $B \subseteq \{1, \dots, n\}$ such that A_B is nonsingular, where A_B stands for the restriction of A to the columns indexed by B. Analogously, $N \coloneqq \{1, \dots, n\} \setminus B$ denotes the nonbasic variables.

A basis B is optimal if and only if three conditions hold at once

$$\text{(regularity)} \quad A_B \text{ is nonsingular}, \tag{9.14a}$$

$$\text{(feasibility)} \quad A_B^{-1} b \geq 0, \tag{9.14b}$$

$$\text{(optimality)} \quad A_N^T (A_B^T)^{-1} c_B \leq c_N. \tag{9.14c}$$

The solution associated to a feasible basis B then has the form $x_B = A_B^{-1} b$ and $x_N = 0$.

Basis stability. We say that a basis B is *stable* if it is optimal for each realization $(A, b, c) \in (\boldsymbol{A}, \boldsymbol{b}, \boldsymbol{c})$. In order to verify stability of B, we need to check if conditions (9.14) are satisfied for each realization.

Checking for stability of a given basis is known to be a *co-NP*-hard problem (Hladík, 2014a). Below, we present a characterization of basis stability via the three condition, and we also show some simple sufficient tests.

Regularity. Condition (9.14a) is satisfied for each realization if and only if the interval matrix $\boldsymbol{A}_B$ is regular. Regularity was discussed in Section 3.6. Hoewever, the methods checking for conditions (9.14b)–(9.14c) usually verify regularity of $\boldsymbol{A}_B$ implicitly, so there is no need to do it explicitly.

Feasibility. Condition (9.14b) is satisfied for each realization if and only if the interval system

$$\boldsymbol{A}_B x = \boldsymbol{b}, \ x \geq 0$$

is strongly solvable. This can be verified by means of Theorem 4.18.

An alternative way is to compute the solution set of the interval system of linear equations $\boldsymbol{A}_B x = \boldsymbol{b}$ and check it lies in the nonnegative orthant. This approach has the benefit of a simple sufficient condition: Compute $\boldsymbol{x} \in \mathbb{IR}^n$, an enclosure to the solution set of $\boldsymbol{A}_B x = \boldsymbol{b}$, for which we know many methods from Sections 3.3–3.5, and then check if $\underline{x} \geq 0$.

Optimality. To check condition (9.14c), it is convenient to substitute $y \coloneqq (A_B^T)^{-1} c_B$. Now, (9.14c) is satisfied for each realization if and only if the interval system

$$\boldsymbol{A}_N^T y \leq \boldsymbol{c}_N, \ \boldsymbol{A}_B^T y = \boldsymbol{c}_B$$

is strongly solvable. Herein, we can again employ techniques of strong solvability developed in Section 4.4.

A computationally cheaper sufficient condition works as follows. Compute $\boldsymbol{y} \in \mathbb{IR}^n$, an enclosure to the solution set of $\boldsymbol{A}_B^T y = \boldsymbol{c}_B$ and then check if

$$\sup(\boldsymbol{A}_N^T \boldsymbol{y}) \leq \underline{c}_N,$$

where the left-hand side is evaluated by interval arithmetic.

Consequences of basis stability. Under basis stability, we can easily determine the optimal solution set and the optimal value range. Moreover, the optimal value function reads $f(A,b,c) = c_B^T A_B^{-1} b$, so it is continuous on $(A,b,c) \in (\boldsymbol{A},\boldsymbol{b},\boldsymbol{c})$. As a consequence, $[\underline{f},\overline{f}] = f(\boldsymbol{A},\boldsymbol{b},\boldsymbol{c})$, that is, each value in the interval $[\underline{f},\overline{f}]$ is an optimal value $f(A,b,c)$ of some realization, and there are no gaps in the optimal value range.

Theorem 9.12 (Beeck, 1978)**.** *If basis B is stable, then each solution of*

$$\underline{A}_B x_B \le \overline{b},\ -\overline{A}_B x_B \le -\underline{b},\ x_B \ge 0,\ x_N = 0. \tag{9.15}$$

belongs to $\mathcal{S}$. Conversely, each LP realization has at least one optimal solution satisfying (9.15).

Proof. Since B is stable, the vector with components $x_B = A_B^{-1} b$ and $x_N = 0$ is optimal for each realization $(A,b,c) \in (\boldsymbol{A},\boldsymbol{b},\boldsymbol{c})$. Thus, the set in question is defined as the solution set of the interval system

$$\boldsymbol{A}_B x_B = \boldsymbol{b},\quad x_B \ge 0,\quad x_N = 0.$$

According to Corollary 3.7, the solution set is described by (9.15). □

The reason that (9.15) does not describe the whole optimal solution set $\mathcal{S}$ is that some instances might be degenerate and there can be more optimal bases and optimal solutions even for one degenerate LP problem. To enforce that (9.15) characterizes $\mathcal{S}$ exactly, strict inequalities must be fulfilled in conditions (9.14b)–(9.14c).

Theorem 9.13 (Beeck, 1978)**.** *If basis B is stable, then*

$$\underline{f} = \min \underline{c}_B^T x_B \quad \textit{subject to}\ \ \underline{A}_B x_B \le \overline{b},\ -\overline{A}_B x_B \le -\underline{b},\ x_B \ge 0, \tag{9.16a}$$

$$\overline{f} = \max \overline{c}_B^T x_B \quad \textit{subject to}\ \ \underline{A}_B x_B \le \overline{b},\ -\overline{A}_B x_B \le -\underline{b},\ x_B \ge 0. \tag{9.16b}$$

Proof. In view of Theorem 9.12, it is sufficient to find the minimal and maximal values of the objective function on the set described by (9.15). Due to nonnegativity of variables, we take the objective vector $\underline{c}$ for minimization and $\overline{c}$ for maximization. □

Notes and further reading. The conditions for checking basis stability (and nondegenerate basis stability) were proposed in Koníčková (2001); Krawczyk (1975). Characterization of basis stability by a reduction to certain (exponentially many) instances was developed in Rohn (1993d).

Král (2020) extended basis stability to the generally constrained interval LP problems. Based on stability of complementarity in the Karush–Kuhn–Tucker conditions, Oelschlägel and Süße (1980); Lange (2020) extended basis stability to interval convex programming problems.

9.4 Applications

Interval linear programming has been utilized in many real-world applications such as in economics (e.g., the portfolio selection problem), project management, product and process scheduling, environmental management and planning, solid waste management planning, structural optimization, chemical and metallurgical process design, statistics and data analysis, ... There are simply too many to mention them all. The aim of this section how interval linear programming influences other disciplines and how it was utilized there.

We begin, however, with an illustration of interval linear programming on a diet problem.

Example 9.14 (Stigler's nutrition model)**.** The nutrition problem is the classical problem, in which we seek for a combination of n different types of food such that m nutritional demands are satisfied and the overall cost in minimized. It is formulated as an LP problem

$$\min\ c^T x\ \ \text{subject to}\ \ Ax \ge b,\ x \ge 0.$$

where x_j is the number of units of food j to be consumed, b_i is the required minimal amount of nutrient i, c_j is the price per unit of food j, and a_{ij} is the amount of nutrient i contained in one unit of food j.

	calories (1000)	protein (g)	calcium (g)	iron (mg)	vitamin-a (1000iu)	vitamin-b1 (mg)	vitamin-b2 (mg)	niacin (mg)	vitamin-c (mg)
wheat	44.7	1411	2.0	365		55.4	33.3	441	
cornmeal	36	897	1.7	99	30.9	17.4	7.9	106	
canned milk	8.4	422	15.1	9	26	3	23.5	11	60
margarine	20.6	17	.6	6	55.8	.2			
cheese	7.4	448	16.4	19	28.1	.8	10.3	4	
peanut-b	15.7	661	1	48		9.6	8.1	471	
lard	41.7				.2		.5	5	
liver	2.2	333	.2	139	169.2	6.4	50.8	316	525
pork roast	4.4	249	.3	37		18.2	3.6	79	
salmon	5.8	705	6.8	45	3.5	1	4.9	209	
green beans	2.4	138	3.7	80	69	4.3	5.8	37	862
cabbage	2.6	125	4	36	7.2	9	4.5	26	5369
onions	5.8	166	3.8	59	16.6	4.7	5.9	21	1184
potatoes	14.3	336	1.8	118	6.7	29.4	7.1	198	2522
spinach	1.1	106		138	918.4	5.7	13.8	33	2755
sweet-pot	9.6	138	2.7	54	290.7	8.4	5.4	83	1912
peaches	8.5	87	1.7	173	86.8	1.2	4.3	55	57
prunes	12.8	99	2.5	154	85.7	3.9	4.3	65	257
lima beans	17.4	1055	3.7	459	5.1	26.9	38.2	93	
navy beans	26.9	1691	11.4	792		38.4	24.6	217	
demand	3	70	.8	12	5	1.8	2.7	18	75

Table 9.1: (Example 9.14) Nutritive value of foods (per dollar spent) and the nutritional demand b.

For concreteness, consider Stigler's nutrition model (Dantzig, 1963). The problem involves $m = 9$ nutrients and $n = 20$ types of food. Nutritive values of foods (i.e., the components of matrix A) are displayed in Table 9.1. The entries in A are normalized such that they give nutritive values of foods per dollar spent. This means that the objective vector is $c = e^T = (1, \dots, 1)^T$.

Since the amount of nutrients in foods is not fixed and hardly known exactly, it is reasonable to assume that the quantities are obtained with a certain accuracy. Suppose first that the entries a_{ij} are known with 10% accuracy. That is, $\boldsymbol{A}$ has interval entries $\boldsymbol{a}_{ij} = [a_{ij} - 0.1|a_{ij}|, a_{ij} + 0.1|a_{ij}|]$. Due to the structure of the LP problem, the optimal value range is computed efficiently by means of Theorem 4.2 and we get $[\underline{f}, \overline{f}] = [0.09878, 0.12074]$.

Checking the conditions for basis stability, we observe that the problem is not basis stable. In spite of this, we can compute an interval enclosure of the optimal solution set $\mathcal{S}$ by means of the interval system based on the duality characterization. Using the contractor method from Hladík (2012c) to make it tighter, we obtain the enclosure

$$\begin{aligned}\big(&[0, 0.0734],\ [0, 0.0438],\ [0, 0.0576],\ [0, 0.0283],\ [0, 0.0535],\\ &[0, 0.0315],\ [0, 0.0339],\ [0, 0.0300],\ [0, 0.0246],\ [0, 0.0337],\\ &[0, 0.0358],\ [0, 0.0387],\ [0, 0.0396],\ [0, 0.0429],\ [0, 0.0370],\\ &[0, 0.0443],\ [0, 0.0290],\ [0, 0.0330],\ [0, 0.0472],\ [0, 0.1057]\big)^T,\end{aligned}$$

which is still somewhat overestimated, but we computed it in polynomial time.

Second, suppose that the entries a_{ij} are known with 1% accuracy. The optimal value ranges in the interval $[\underline{f}, \overline{f}] = [0.10758, 0.10976]$. Since the intervals in $\boldsymbol{A}$ are narrow enough, the problem becomes basis stable and the stable optimal basis is $B = (1, 8, 12, 15, 20)$. Therefore the solution set has a simple description and we easily determine its interval hull,

$$\begin{aligned}&\boldsymbol{x}_1 = [0.0282, 0.0309],\ \ \boldsymbol{x}_8 = [0.0007, 0.0031],\ \ \boldsymbol{x}_{12} = [0.0110, 0.0114],\\ &\boldsymbol{x}_{15} = [0.0047, 0.0053],\ \ \boldsymbol{x}_{20} = [0.0600, 0.0621];\end{aligned}$$

the other components are zero. □

In Section 9.4.1–9.4.3, we will describe several areas of applications of interval linear programming. Beside those, we also briefly mention some other important problems, where interval LP techniques were utilized:

- Based on the theory of interval linear programming, Rohn (1989b, 2012b) introduced a novel kind of a condition number of real-valued linear programs.
- Rounding errors may (and *do*) influence the results of calculations on computers. Thus, when solving real-valued LP problems on computer, the results are not completely reliable. Interval techniques are suitable to make the calculations numerically rigorous. The goal of *verification* is to produce numerically guaranteed bounds for the true results. We discuss this topic more in detail in Section 10.1.
- Mohammadi and Gentili (2021) introduced the so-called *outcome range problem*. For an interval LP problem, an outcome function is an additional objective function of interest. Even though this model might seem to be similar to mutliobjective or bilevel programming, it is not the same. The authors applied the model in healthcare access measurement, analysing the primary care service for children in the state of Mississippi.

9.4.1 Game theory

Zero-sum matrix games. Game theory studies interactions and strategies of competing players (von Neumann and 2007). The most simple model of non-cooperative game theory is *the two-person zero-sum matrix game*. It is represented by a matrix $A \in \mathbb{R}^{m\times n}$. If player I chooses (pure) strategy i and player II chooses j, then the value of a_{ij} gives the payoff of player I and $-a_{ij}$ gives the payoff of player II. We can without loss of generality, assume that $A \geq 0$ since otherwise we can shift $a_{ij} \to a_{ij} + \alpha$ by a sufficiently large constant α.

It turned out that, besides pure strategies, it is essential for the players to play *mixed strategies*, which are probability distributions on pure strategies. A mixed strategy for player I is thus defined as any vector $x \in \mathbb{R}^m$ such that $x \geq 0$ and $e^T x = 1$, and similarly for player II. If player I chooses mixed strategy x and player II mixed strategy y, then the expected payoffs of players I and II are $x^T Ay$ and $-x^T Ay$, respectively.

What is a solution of a game? A common concept is *a Nash equilibrium*, which is a pair of mixed strategies $(\hat{x}, \hat{y})$ such that

$$\begin{aligned} \hat{x}^T A\hat{y} &\geq x^T A\hat{y}, \quad \forall x \geq 0,\ e^T x = 1, \\ \hat{x}^T A\hat{y} &\leq \hat{x}^T Ay, \quad \forall y \geq 0,\ e^T y = 1. \end{aligned}$$

In other words, Nash equilibrium is such a situation, in which no player can gain more by changing his strategy.

It is an important fact that (at least one) Nash equilibrium always exists. Nash equilibria can be computed by means linear programming. Consider an LP problem

$$\max\ u \ \text{ subject to } \ A^T x \geq eu,\ e^T x = 1,\ x \geq 0,$$

and its dual

$$\min\ v \ \text{ subject to } \ Ay \leq ev,\ e^T y = 1,\ y \geq 0.$$

Let $(\hat{x}, \hat{u})$ and $(\hat{y}, \hat{v})$ be their optimal solutions, respectively. Then $(\hat{x}, \hat{y})$ with a Nash equilibrium. Conversely, every Nash equilibrium can be expressed in this way for some optimal solutions $(\hat{x}, \hat{u})$ and $(\hat{y}, \hat{v})$. *The value of the game* is the quantity $w(A) := \hat{u} = \hat{v}$, which represents the expected payoff of player I.

Interval zero-sum matrix games. Consider now an interval payoff matrix $\boldsymbol{A} \in \mathbb{IR}^{m\times n}$ and the corresponding family of matrix games. The interval payoffs usually represent certain uncertainty in the knowledge of the payoffs.

Since real matrix games are solved by linear programming, interval matrix games can be handled by the techniques of interval linear programming. In particular, consider the interval LP problem

$$\max\ u\ \text{ subject to }\ \boldsymbol{A}^T x \geq eu,\ e^T x = 1,\ x \geq 0. \tag{9.17}$$

In this context, the optimal value range represents the range of the values of the games $w(\boldsymbol{A}) \coloneqq \{w(A);\ A \in \boldsymbol{A}\}$, and it is easily computed by solving two linear programs (Liu and Kao, 2009).

The optimal solution set represents the set of all Nash equilibria over all realizations of the payoff matrix. Therefore the results from Section 9.2 apply directly to characterize and approximate the set of Nash equilibria.

Some of the components of a Nash equilibrium $(\hat{x}, \hat{y})$ are positive and the others are zero. The former are in some sense more important (we take the corresponding strategies into account for the mixed strategy), while the latter are less important. Basis stability then depicts the situation, in which the Nash equilibrium consists of the same subset of pure strategies for any realization $A \in \boldsymbol{A}$.

Notes and further reading. While zer-sum matrix game are related to linear programming, bimatrix games are closely related to integer linear programming, and therefore more difficult to solve. Hladík (2010b) investigated different issues in interval bimatrix games, including the set of all equilibria or several types of stabilities of supports of the strategies (i.e., which strategies are involved in the equilibrium).

Cooperative interval games were analysed in Bok and Hladík (2015), among others.

9.4.2 Linearization in the constraint satisfaction problem

Interval linear programming also provides tool to handle nonconvex function in the constraint satisfaction problem and global optimization (Chapter 8); in the sequel, we consider just the former since the objective function can easily be involved.

Let $\boldsymbol{x} \in \mathbb{IR}^n$ be an initial box, in which we seek for the solutions of the system of nonlinear equations and inequalities

$$f_j(x) = 0, \quad j = 1, \dots, j^*, \tag{9.18a}$$
$$g_\ell(x) \leq 0, \quad \ell = 1, \dots, \ell^*, \tag{9.18b}$$

or, in compact form,

$$f(x) = 0, \tag{9.19a}$$
$$g(x) \leq 0, \tag{9.19b}$$

where $f_j, g_\ell \colon \mathbb{R}^n \mapsto \mathbb{R}$ are real-valued functions and $f(x) = (f_1(x), \dots, f_{j^*}(x))$, $g(x) = (g_1(x), \dots, g_{\ell^*}(x))$.

Our aim is to linearize the noncovnex functions in such a way that we enclose them by interval linear functions. Let $h \colon \mathbb{R}^n \mapsto \mathbb{R}$ be a function and $x^0 \in \boldsymbol{x}$ an arbitrary point. Under general assumptions, the function is enclosed as

$$h(x) \subseteq S_h(\boldsymbol{x}, x^0)(x - x^0) + h(x^0), \quad \forall x \in \boldsymbol{x}, \tag{9.20}$$

where $S_h \colon \mathbb{IR}^n \times \mathbb{R}^n \mapsto \mathbb{IR}^n$ is a suitable interval-valued function. For this purpose, we can utilize mean value form enclosure (6.2), slopes or some other centered form.

Using this type of enclosure for every constraint function, we transform the system (9.19) into an interval system of linear equations and inequalities

$$\boldsymbol{A}(x - x^0) + f(x^0) = 0, \tag{9.21a}$$
$$\boldsymbol{B}(x - x^0) + g(x^0) \leq 0, \tag{9.21b}$$

where $\boldsymbol{A} \in \mathbb{IR}^{j^* \times n}$ and $\boldsymbol{B} \in \mathbb{IR}^{\ell^* \times n}$.

The choice of x^0. Now, it is a question what choice of the center $x^0 \in \boldsymbol{x}$ is the best one? A vertex selection was introduced in Araya et al. (2012); Jansson (2000); Trombettoni et al. (2011). If x^0 is a vertex of the box $\boldsymbol{x}$, then the solution set of (9.21) is a convex polyhedron described by linear constraints. For example, if we put $x^0 \coloneqq \underline{x}$, then $x - \underline{x}$ is nonnegative. By the Oettli–Prager Theorem 3.2, the interval system (9.21a) is described by linear inequalities

$$\begin{aligned} \underline{A}(x - \underline{x}) + f(\underline{x}) &\leq 0, \\ \overline{A}(x - \underline{x}) + f(\underline{x}) &\geq 0. \end{aligned}$$

By the Gerlach Theorem 4.1, the solution set to (9.21b) is described by

$$\underline{B}(x - \underline{x}) + g(\underline{x}) \leq 0.$$

If we take $x^0 \coloneqq \overline{x}$, then $x - \overline{x} \leq 0$ and the solution set to (9.21) is characterized by the systel of linear inequalities

$$\begin{aligned} \overline{A}x &\leq \overline{A}\overline{x} - f(\overline{x}), \\ \underline{A}x &\geq \underline{A}\overline{x} - f(\overline{x}), \\ \overline{B}x &\leq \overline{B}\overline{x} - g(\overline{x}). \end{aligned}$$

Similarly for any other vertex of $\boldsymbol{x}$. Basically, we can choose several centers and jin all the resulting inequalities together, hoping for tighter approximation. Araya et al. (2012) recommend to use two opposite corners of $\boldsymbol{x}$; the other vertices do not significantly increase efficiency.

Hladík and Horáček (2014) initiated a non-vertex selection of the center x^0. When x^0 is not a vertex of $\boldsymbol{x}$, then the solution set does not admit a linear description in general. By Theorem 4.13, the solution set to (9.21) is described by

$$\begin{aligned} |A^c(x - x^0) + f(x^0)| &\leq A^\Delta |x - x^0|, \\ B^c(x - x^0) &\leq B^\Delta |x - x^0| - g(x^0). \end{aligned}$$

Thi system is not linear due to the absolute values. To get rid of them, one can estimate them from above by a linear function, leading to a relaxation by means of a system of linear inequalities.

Figure 9.1 symbolically illustrates different choices of the center x^0. In the pictures, $\mathcal{S}$ denotes the set described by (9.18), the initial box $\boldsymbol{x}$ is colored in light gray, and the linear relaxation in dark gray.

9.4.3 Robust optimization

In robust optimization (Ben-Tal et al., 2009; Bertsimas et al., 2011), we deal with uncertainty in such a way that we want to be safe against the worst case situations. In particular, we focus only on those solutions that satisfy the given constraints for all possible realizations of the data.

Inequality constrained LP form with nonnegativity. Consider an LP problem in the form

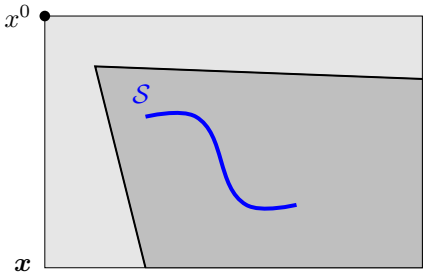
$$\min \ c^T x \ \text{ subject to } \ Ax \leq b, \ x \geq 0,$$

and suppose that (A, b) comes from a certain uncertainty set $\mathcal{U}$. We consider uncertainties in the constraints only since the objective function can be transferred into the constraints.

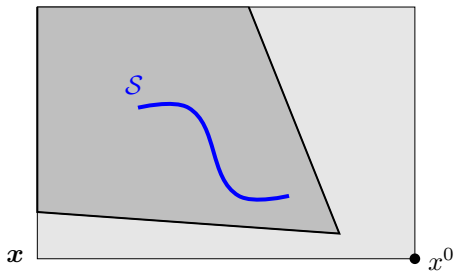
The robust optimization approaches the problem by constructing a deterministic optimization problem, called *the robust counterpart*,

$$\min \ c^T x \ \text{ subject to } \ Ax \leq b, \ x \geq 0, \ \forall (A, b) \in \mathcal{U}.$$

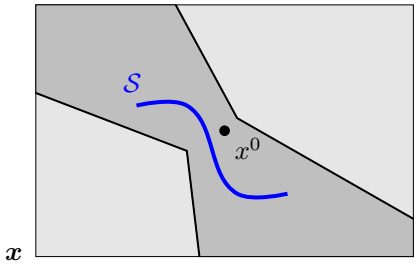
The uncertainty set $\mathcal{U}$ has often the form of a Cartesian product of intervals or ellipsoids. In which such cases, the robust counterpart problem is efficiently solvable. In this section, we consider uncertainties have the form of intervals.



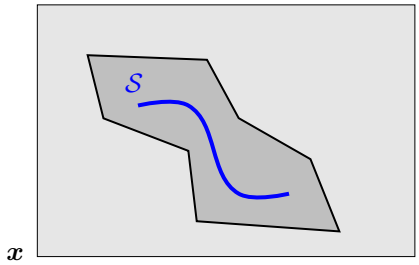
(a) Typical situation when choosing x^0 to be a vertex.



(b) Typical situation when choosing x^0 to be the opposite vertex.



(c) Typical situation when choosing $x^0 = x^c$.



(d) Typical situation when choosing all of them.

Figure 9.1: A symbolic illustration of linearization for different choices of the center x^0.

Let $\boldsymbol{A} \in \mathbb{IR}^{m\times n}$ and $\boldsymbol{b} \in \mathbb{IR}^m$ and suppose that $\mathcal{U} = \boldsymbol{A} \times \boldsymbol{b}$. Then the robust counterpart problem reads

$$\min\ c^T x \ \text{ subject to } \ Ax \le b,\ x \ge 0,\ \forall A \in \boldsymbol{A}, \forall b \in \boldsymbol{b}.$$

By Theorem 4.35 or easily by direct inspection we see that it is equivalent to the LP problem

$$\min\ c^T x \ \text{ subject to } \ \overline{A}x \le \underline{b},\ x \ge 0.$$

Inequality constrained LP form. Now, consider an LP problem in the form

$$\min\ c^T x \ \text{ subject to } \ Ax \le b$$

and the uncertainty set $\mathcal{U} = \boldsymbol{A} \times \boldsymbol{b}$. Analogously to the previous case, the robust counterpart problem is defined

$$\min\ c^T x \ \text{ subject to } \ Ax \le b,\ \forall A \in \boldsymbol{A}, \forall b \in \boldsymbol{b}.$$

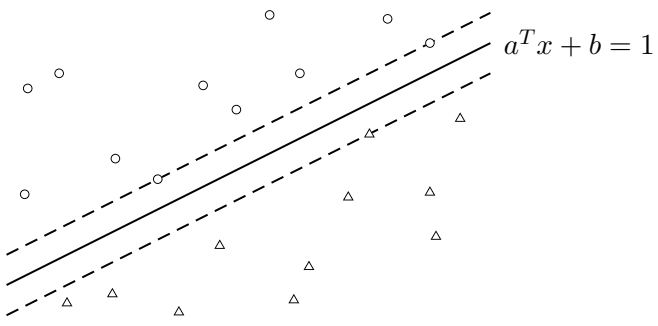
This means that the robust feasible solutions are the strong feasible solutions discussed in Section 4.4. By Corollary 4.36, the problem is equivalent to an LP problem

$$\min\ c^T x^1 - c^T x^2 \ \text{ subject to } \ \overline{A}x^1 - \underline{A}x^2 \le \underline{b},\ x^1, x^2 \ge 0.$$

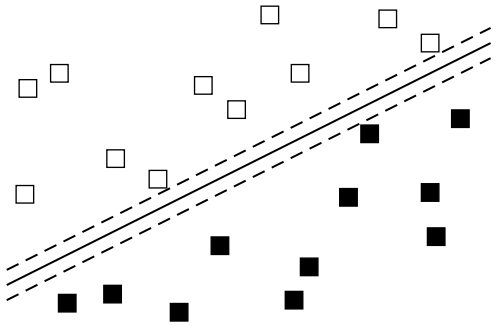
Example 9.15 (Robust classification)**.** Consider the classification problem with two classes of data – the first one consists of points $x_1, \dots, x_m \in \mathbb{R}^n$ and the second one consists of points $y_1, \dots, y_k \in \mathbb{R}^n$. The goal is to construct a classifier that predicts to which class a new point belongs to. A basic linear classifier separates the two classes of points by a widest separating band. That is, we seek for a hyperplane $a^T x + b = 1$ such that one class of points belongs to the positive half-space, the second one belongs to the negative half-space, and the separating band is maximal; see an illustration in Figure 9.2a. This leads to a convex quadratic program in variables $a \in \mathbb{R}^n$ and $b \in \mathbb{R}$

$$\min\ \|a\|_2 \ \text{ subject to } \ a^T x_i + b \ge 1\ \forall i,\ a^T y_j + b \le -1\ \forall j,$$

which is efficiently solvable.



(a) The widest separating band for real data.



(b) The widest separating band for interval data.

Figure 9.2: (Example 9.15) A linear classifier for real data and the robust linear classifier for interval data.

Data are often not measured exactly and we observe them with a certain accuracy only. This means that the input data have the form of interval vectors $\boldsymbol{x}_i \in \mathbb{IR}^n$, $i = 1, \ldots, m$, and $\boldsymbol{y}_j \in \mathbb{IR}^n$, $j = 1, \ldots, k$, enclosing the true, albeit unobservable, data. Utilizing the robust approach as described above, we formulate the robust counterpart model

$$\begin{aligned} \min \ \|a^1 - a^2\|_2 \ \text{ subject to } \ & \underline{x}_i^T a^1 - \overline{x}_i^T a^2 + b \geq 1 \ \forall i, \\ & \overline{y}_j^T a^1 - \underline{y}_j^T a^2 + b \leq -1 \ \forall j, \\ & a^1, a^2 \geq 0. \end{aligned}$$

It is again a convex quadratic program (in variables $a^1, a^2 \in \mathbb{R}^n$ and $b \in \mathbb{R}$). The program calculates the widest band that separates all possible instances of the interval data simultaneously; see Figure 9.2b. □

Chapter 10

Other applications

10.1 Verification

A computer solving a given problem performs a series of floating-point operations. Due to rounding errors, we can hardly expect to obtain the exact solution in the output. *Verification* is a technique that, given an approximate solution, a posteriori calculates rigorous bounds for the true solution. Here, the term "rigorous" means that the bounds are provably correct taking into account possible numerical aspects and rounding errors.

The introductory Example 2.1 convinced us that even a simple expression evaluation may result in a completely wrong value, so verification is necessary when we want to make the results of unreliable calculations to be numerically rigorous. The basic paradigm of verification is that *every* computation on a computer should be done in a verified way.

More-or-less efficient methods for verification are known for most of the standard numerical problems, including solving linear and nonlinear equations, computing eigenvalues or solving certain classes of optimization problems, among others. Naturally, we would like the verification procedure to be less expensive than solving the original problem. This is not the case for all problems, but we are approaching this goal.

Using a floating point based architecture, verification is in principle possible only for those problems that are stable in some sense. On the one hand, we can in essence verify nonsingularity of a matrix since the set of nonsingular matrices is open, so there are no singular matrices in a small neighbourhood. On the other hand, we cannot verify singularity since it is not preserved for arbitrarily small perturbations.

The following subsections discuss verification for basic linear algebraic problems, such as solving a system of linear and nonlinear equations or computing eigenvalues of a matrix.

10.1.1 Verification for linear systems

Let $A \in \mathbb{R}^{n\times n}$ and $b \in \mathbb{R}^n$. Given an approximate solution $x^* \approx A^{-1}b$ of the linear system $Ax = b$, we want to find as small as possible $\boldsymbol{y} \in \mathbb{IR}^n$ such that the true solution $A^{-1}b \in x^* + \boldsymbol{y}$. This gives us rigorous lower and upper bounds on the true solution. Figure 10.1 symbolically illustrates the idea verification.

To this end, we can apply any of the methods for solving interval linear equations from Section 3.2. We just call them for intervals $\boldsymbol{A} \coloneqq A$ and $\boldsymbol{b} \coloneqq b$, and they yield guaranteed enclosures of the solution $A^{-1}b$. However, as the following example shows, they can give very wide interval enclosures. The general recommendation here is to postpone interval arithmetic to the very end or when really necessary; otherwise, the outer rounding accumulates the overestimation throughout the computation.

Example 10.1 (Amplification factor for the interval Gaussian elimination, Rump (2010))**.** An enclosure of the true solution $A^{-1}b$ can be found by the interval Gaussian elimination. During the iterations and transformations of the matrix $(A \mid b)$, however, the initial tiny intervals are becoming wider and wider. The following table shows a lower bound on the amplification factor of the radii:

n	20	50	100	170
amplification	10^2	10^5	10^{10}	10^{16}

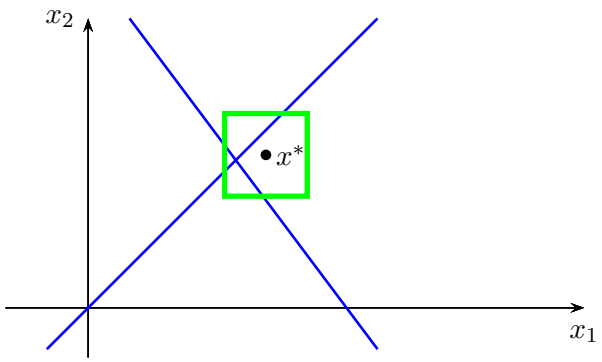


Figure 10.1: Illustration of the verification of an approximate solution x^* to $Ax = b$.

This confirms that the interval Gaussian elimination is not suitable for verification of even medium sized systems. □

On the other hand, the ε-inflation method from Section 3.4.3 can be efficiently applied to do this job. Calling it for degenerate intervals $\boldsymbol{A} := A$ and $\boldsymbol{b} := b$, and if it succeeds, then it yields $\boldsymbol{y}$ satisfying the condition

$$C(b - Ax^*) + (I - CA)\boldsymbol{y} \subseteq \operatorname{int} \boldsymbol{y}.$$

This proves nonsingularity of A and guarantees that $A^{-1}b \in x^* + \boldsymbol{y}$.

Example 10.2. Let A be the Hilbert matrix of size 10, that is, $a_{ij} = \frac{1}{i+j-1}$, and $b := Ae$. Then the linear system $Ax = b$ has the solution $x = e = (1, \dots, 1)^T$. Below, there is displayed the approximate solution computed by MATLAB, and the enclosure calculated in two iterations of the ε-inflation method:

Approximate solution by MATLAB	Enclosing interval by ε-inflation method
0.999999999235452	[0.99999973843401, 1.00000026238575]
1.000000065575364	[0.99999843048508, 1.00000149895660]
0.999998607887449	[0.99997745481481, 1.00002404324710]
1.000012638750021	[0.99978166603900, 1.00020478046370]
0.999939734980300	[0.99902374408278, 1.00104070076742]
1.000165704992114	[0.99714060702796, 1.00268292103727]
0.999727989024899	[0.99559932282378, 1.00468935360003]
1.000263042205847	[0.99546972629357, 1.00425202249136]
0.999861803020249	[0.99776781605377, 1.00237789028988]
1.000030414871015	[0.99947719419921, 1.00049082925529]

We can see that the resulting intervals overestimate the true distance of the approximate solution to the true solution by the factor of about 20. This is not a bad result, taking into account that the condition number of A is $\approx 1.6 \cdot 10^{13}$. □

Rump (2010) claims that this method of verification theoretically takes 9 to 12 times the computing time of the Gaussian elimination itself for finding x^*. When tested for random instances of dimension 100 to 2000, the ratio turned out to be slightly better, so that the verification is about 7 times slower than solving the original problem.

Notes and further reading. For other methods, see Rump (2013a,b, 2010).

10.1.2 Verification for nonlinear systems

Let $f\colon \mathbb{R}^n \to \mathbb{R}^n$ be differentiable in each component, and consider the square system of nonlinear equations $f(x) = 0$. Given an approximate solution x^*, we want to find $\boldsymbol{y} \in \mathbb{IR}^n$ such that there is a solution in $x^* + \boldsymbol{y}$.

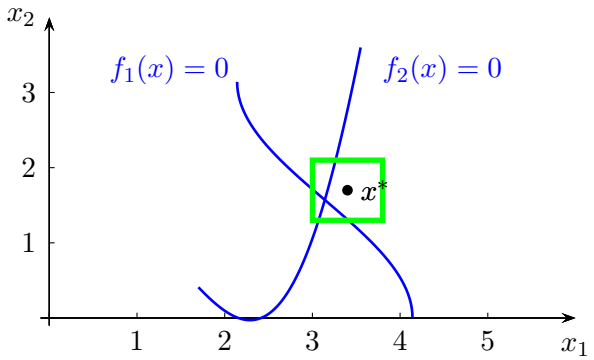


Figure 10.2: (Example 10.3) Illustration of the verification of x^* to $f(x) = 0$.

Similarly as for the linear case, we adapt Theorem 8.15 to the uniqueness of a root and come across the ε-inflation method. First, we put $\boldsymbol{y} \coloneqq -Cf(x^0)$. Then we repeat inflating $\boldsymbol{z} \coloneqq [0.9, 1.1]\boldsymbol{y} + 10^{-20}[-1, 1]$ and updating

$$\boldsymbol{y} \coloneqq -Cf(x^0) + (I_n - C\nabla f(\boldsymbol{x}))\boldsymbol{z} \tag{10.1}$$

until $\boldsymbol{y} \subseteq \operatorname{int} \boldsymbol{z}$, or we exceed the maximum number of iterations. If $\boldsymbol{y} \subseteq \operatorname{int} \boldsymbol{z}$, then there is a unique solution in $x^* + \boldsymbol{y}$.

Example 10.3. Consider the nonlinear system

$$\begin{aligned} \pi^2(x_2 - \pi/2) + 4x^2 \sin(x_1) &= 0, \\ x_1 - \pi - \cos(x_2) &= 0 \end{aligned}$$

and let us have an approximate solution $x^* = (3.1415, 1.5708)^T$ of the true solution $(\pi, \pi/2)^T$.

For $\boldsymbol{z} \coloneqq ([-10^{-5}, 10^{-5}], [-10^{-5}, 10^{-5}])^T$, the verification fails, but for $\boldsymbol{z} \coloneqq ([-10^{-4}, 10^{-4}], [-10^{-4}, 10^{-4}])^T$ the condition (10.1) succeeds. Thus, each entry of x^* is precise with accuracy 10^{-4}. Figure 10.2 symbolically illustrates the idea verification. □

Multiple roots. A double root x^* of a function $f\colon \mathbb{R} \to \mathbb{R}$ can be verified by a reduction to verification of the simple root $(x^*, 0)$ of the function

$$g(x, \varepsilon) = \big(f(x) - \varepsilon,\ f'(x)\big)^T.$$

We proceed similarly for a higher order multiple roots.

Other approaches. The tests for the existence of a solution we presented here (and in Section 8.1) were based on an interval operator mapping a box into itself and calling the Brouwer's fixed-point theorem. Besides Brouwer's theorem, there exist also another techniques that can be used. They include Miranda's theorem or the Newton–Kantorovich theorem; see Mayer (2017).

Verification of eigenvalues. Verification of eigenvalues and eigenvectors of a matrix $A \in \mathbb{R}^{n\times n}$ is a specific problem dealt with in more detail in Mayer (2017); Rump (2010). In essence, the verification can be performed by the above approach applied to the nonlinear system $f(x, \lambda) = 0$, where

$$f(x, \lambda) = (Ax - \lambda x, \|x\|^2 - 1)^T,$$

and for the normalization, we can use the Euclidean norm or the maximum norm, for instance.

Verification of singular values. Verification of the singular values of a matrix $A \in \mathbb{R}^{m\times n}$ can be reduced to the verification of the eigenvalues of a symmetric matrix via relation from Theorem 1.14(2). Another way uses the reduction to the verification of the roots of a nonlinear system $f(x, y, \sigma, \varepsilon) = 0$, where

$$f(x, y, \sigma, \varepsilon) = (Ax - \sigma y,\ A^T y - (\sigma + \varepsilon)x,\ \|x\|^2 - 1,\ \|y\|^2 - 1)^T.$$

Notes. Adaptation of ε-inflation method to nonlinear systems is due to Rump (1983). Variants of this method and are discussed, e.g., in Kreinovich et al. (1997); Rump (1998).

10.2 Tolerance analysis and inverse problems

Notes. Rohn (1989b, 2012b) proposed a condition number for linear programming problems based on a relative change of the optimal value subject to infinitesimal perturbations of the data.

10.3 Solving quantified systems

The interval computation approach to constraint satisfaction and global optimization solving (chapter 8) can be generalized to handle more complex constraints involving also quantifiers and logical operations. Pareto optimal solutions of multiobjective optimization problems or Nash equilibria in game problems are particular examples of such systems. Using the interval methods, we are able to determine a kind of ε-approximation of all solutions (Kreinovich and Kubica, 2010), even though the problem itself is potentially undecidable.

10.4 Statistics

Kreinovich and Xiang (2008)

Berleant (1996) discusses how to rigorously represent and operate on a probability distribution function when it is discretized. Discretization leads to intervalization of the domain, so a verified lower and upper bound on the distribution function are obtained by interval computation.

10.5 Real world applications

Links to diverse applications of interval computation are collected at

`http://www.cs.utep.edu/interval-comp/appl.html`

10.5.1 Robotics

Application of interval computation in robotics is one of the main research topics of the COPRIN team, established at 2001 under the French institute INRIA. Problems that are studied:

- *Forward kinematics problem* (Merlet, 1995, 2004)

  The task is to determine all possible positions of the end-effectors of a robot for given joint coordinates. The related problem of *workspace calculation* aims at computing all reachable locations of the center of a robot, possibly with constraints such as mechanical restrictions on the motion of joints.

- *Robot calibration* (Daney, 2002, 2003; Hladík et al., 2011a)

  Determination or estimation of the actual values of parameters of a robot, such as position and orientation of links and joints.

- *Robot designing* (Hao and Merlet, 2005; Merlet, 2005)

  Naturally, robot parameters influence its performance. That is why one looks for a robot with parameters guaranteeing small positioning errors or reachability of some states, for example.

- *Robot singularities* (Merlet, 1998, 2007; Merlet and Donelan, 2006; Oetomo et al., 2009)

  Here, one wants to determine all values of robot parameters that may cause a deadlock or breakdown. Usually, it reduces to checking singularity of the so-called Jacobian matrix of the robot.

10.5.2 Computer graphics

The books Ratschek and Rokne (2003); Snyder (1992) show how important is interval computation to avoid errors in computer graphics. Some of the areas, where interval computation comes into play, are:

- *Precise geometric computation* (Kettner et al., 2008; Schirra, 2000)

  Precision in geometric computation is an important issue. There are recorded examples, where floating-point arithmetic led to wrong results in computing convex hulls in the plane and Delaunay triangulations in space, among others. Interval arithmetic, if used, can guarantee the correctness of calculations.

- *Processing geometric objects* (Mudur and Koparkar, 1984; Suffern and Fackerell, 1991)

  Curves, surfaces and other geometric objects can be rigorously processed in a similar manner as in constraint programming (Chapter 8). In curve tracing, we must in addition to generate the sequence of subsequent pixels the curve is passing, which makes the problem more interesting. Interval computation helps not only in plotting curves, but also in intersection detection, silhouette detection, linearity and planarity testing and rendering shaded, tessellated and contour pictures, besides others.

Besides the aforementioned problems, interval computation was also used in ray tracing (Sanjuan-Estrada et a 2003) (Hickey et al., 2000)

For a positioning of geometrical objects, see also Section 5.4.3.

Chapter 11

Open problems and conjectures

Conjectures

- **Garloff's conjecture I.** (Garloff, 1982)

  see Garloff (2009)

- **Garloff's conjecture II.**

  For a symmetric positive definite interval matrix $\boldsymbol{A}$ the minimum value of $p(A) \coloneqq \frac{\det(A)}{\det(A')} = \frac{1}{A^{-1}_{nn}}$ over $\boldsymbol{A}^S$ is attained at a matrix A_{zz} for some $z \in \{\pm 1\}^n$, where A' is obtained from A by removing the last row and column.

  This problem relates the interval Cholesky decomposition (Garloff, 2012). Denote $A = LL^T$ the Cholesky decomposition of A. We have $l_{nn} = \sqrt{p(A)}$, and similarly for the other diagonal entries of L.

- **Rohn's conjecture**

  The sign accord algorithm (Algorithm 3.1) finishes in a finite number of steps even when the signs are switched for all violated coordinates.

  This was empirically observed by Rohn's master students, moreover the number of steps seemed to be linear, but no theoretical result is known. See also (Neumaier, 1990, pg. 220).

- **Rump's conjecture** (Rump, 1997b)

  Every regular $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$ satisfies $\rho(|(A^c)^{-1}|A^\Delta) < n$.

Open problems

- **Exactness of the natural interval extension w.r.t. accuracy (Kreinovich et al., 2003)**

  It is open whether the following problem is NP-hard: For a given function $f\colon \mathbb{R}^n \to \mathbb{R}$ and an interval vector $\boldsymbol{x} \in \mathbb{IR}^n$, gives the natural interval extension the exact range $f(\boldsymbol{x})$ with respect to a given accuracy?

- *If $\boldsymbol{A} \in \mathbb{IR}^{n\times n}$ is strongly regular, then the implication $\boldsymbol{B} \subseteq \boldsymbol{A} \;\Rightarrow\; |(B^c)^{-1}|B^\Delta \leq |(A^c)^{-1}|A^\Delta$ holds? (Neumaier, 1990, pg. 124)*

- *For an interval linear system of equations with an inverse nonnegative matrix, is there a suitable row and column permutation such that the interval Gaussian elimination does not fail? (Neumaier, 1990, pg. 160)*

Appendices

Appendix A

Properties of interval arithmetic

This appendix displays tables with various properties of interval arithmetic and other interval operations. The formulae were adopted from another literature, including Mayer (2017); Moore et al. (2009); Neumaier (1984, 1990); Ratschek and Rokne (1995); Ris (1975).

If we write that some property does not hold ($\neg$ property or inequation $\neq$), then it means that the property is not true in general, but may be true for some particular cases. In case of division, we implicitly assume that the denominator does not contain zero. If not stated otherwise, we consider real (non-complex) intervals, interval vectors and interval matrices; some properties of complex matrices are presented in Alefeld and Herzberger (1983).

Table A.1: Set operations on intervals ($\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \ldots \in \mathbb{IR}$). We use the operator $\boldsymbol{x} \sqcup \boldsymbol{y} \coloneqq \square(\boldsymbol{x} \cup \boldsymbol{y})$.

$\boldsymbol{x} \subseteq \boldsymbol{y}$	$\Leftrightarrow$	$\lvert x^c - y^c \rvert \leq y^\Delta - x^\Delta$
$\boldsymbol{x} \subseteq \boldsymbol{y}$	$\Leftrightarrow$	$\operatorname{mag}(\boldsymbol{x} - y^c) \leq y^\Delta$
$\boldsymbol{x} \subseteq \boldsymbol{y}$	$\Rightarrow$	$x^\Delta \leq y^\Delta \ \wedge \ \operatorname{mag}(\boldsymbol{x}) \leq \operatorname{mag}(\boldsymbol{y})$
$\boldsymbol{x} \cap \boldsymbol{y} \neq \emptyset$	$\Leftrightarrow$	$\lvert x^c - y^c \rvert \leq x^\Delta + y^\Delta$
$\boldsymbol{x} \cap \boldsymbol{y} \neq \emptyset$	$\Leftrightarrow$	$0 \in \boldsymbol{x} - \boldsymbol{y}$
$\boldsymbol{x} \cap \boldsymbol{y} \neq \emptyset$	$\Leftrightarrow$	$\overline{x} \geq \underline{y} \ \wedge \ \underline{x} \leq \overline{y}$
$\boldsymbol{x} \cap \boldsymbol{y} \neq \emptyset$	$\Leftrightarrow$	$\boldsymbol{x} \cap \boldsymbol{y} = [\max(\underline{x}, \underline{y}), \min(\overline{x}, \overline{y})]$
$\boldsymbol{x}_1 \subseteq \boldsymbol{y}_1 \ \wedge \ \boldsymbol{x}_2 \subseteq \boldsymbol{y}_2$	$\Rightarrow$	$(\boldsymbol{x}_1 \cap \boldsymbol{x}_2) \subseteq (\boldsymbol{y}_1 \cap \boldsymbol{y}_2)$
$\boldsymbol{x} + (\boldsymbol{y} \sqcup \boldsymbol{z}) = (\boldsymbol{x} + \boldsymbol{y}) \sqcup (\boldsymbol{x} + \boldsymbol{z})$		
$\boldsymbol{x}(\boldsymbol{y} \sqcup \boldsymbol{z}) = (\boldsymbol{x}\boldsymbol{y}) \sqcup (\boldsymbol{x}\boldsymbol{z})$		

Table A.2: Endpoint analysis of interval multiplication ($\boldsymbol{x}, \boldsymbol{y} \in \mathbb{IR}$).

	$\inf(\boldsymbol{x}\boldsymbol{y})$	$\sup(\boldsymbol{x}\boldsymbol{y})$
$\boldsymbol{x} \geq 0 \ \wedge \ \boldsymbol{y} \geq 0$	$\underline{x}\underline{y}$	$\overline{x}\overline{y}$
$\boldsymbol{x} \leq 0 \ \wedge \ \boldsymbol{y} \geq 0$	$\underline{x}\overline{y}$	$\overline{x}\underline{y}$
$0 \in \boldsymbol{x} \ \wedge \ \boldsymbol{y} \geq 0$	$\underline{x}\overline{y}$	$\overline{x}\overline{y}$
$\boldsymbol{x} \geq 0 \ \wedge \ \boldsymbol{y} \leq 0$	$\overline{x}\underline{y}$	$\underline{x}\overline{y}$
$\boldsymbol{x} \leq 0 \ \wedge \ \boldsymbol{y} \leq 0$	$\overline{x}\overline{y}$	$\underline{x}\underline{y}$
$0 \in \boldsymbol{x} \ \wedge \ \boldsymbol{y} \leq 0$	$\overline{x}\underline{y}$	$\underline{x}\underline{y}$
$\boldsymbol{x} \geq 0 \ \wedge \ 0 \in \boldsymbol{y}$	$\overline{x}\underline{y}$	$\overline{x}\overline{y}$
$\boldsymbol{x} \leq 0 \ \wedge \ 0 \in \boldsymbol{y}$	$\underline{x}\overline{y}$	$\underline{x}\underline{y}$
$0 \in \boldsymbol{x} \ \wedge \ 0 \in \boldsymbol{y}$	$\min(\underline{x}\overline{y}, \overline{x}\underline{y})$	$\max(\underline{x}\underline{y}, \overline{x}\overline{y})$
$x \geq 0$ and real	$x\underline{y}$	$x\overline{y}$

Table A.3: Interval midpoints and radii ($\boldsymbol{x}, \boldsymbol{y} \in \mathbb{IR}$, $n \in \mathbb{N}$).

$$\begin{aligned}
\operatorname{mid}(\boldsymbol{x} \pm \boldsymbol{y}) &= x^c \pm y^c \\
\operatorname{rad}(\boldsymbol{x} \pm \boldsymbol{y}) &= x^\Delta + y^\Delta
\end{aligned}$$

$$\begin{aligned}
x \text{ real} \vee y \text{ real} \quad &\Rightarrow \quad \operatorname{mid}(\boldsymbol{xy}) = x^c y^c \\
x^c = 0 \ \vee \ y^c = 0 \quad &\Rightarrow \quad \operatorname{mid}(\boldsymbol{xy}) = x^c y^c \\
y \text{ real} \quad &\Rightarrow \quad \operatorname{mid}(\boldsymbol{x}/y) = x^c / y^c
\end{aligned}$$

$$\begin{aligned}
\operatorname{sgn}(\operatorname{mid}(\boldsymbol{xy})) &= \operatorname{sgn}(x^c y^c) \\
\operatorname{mid}(1/\boldsymbol{x}) &= x^c / (\operatorname{mag}(\boldsymbol{x}) \operatorname{mig}(\boldsymbol{x})) \\
\operatorname{mid}(\boldsymbol{x}/\boldsymbol{y}) &= \operatorname{mid}(\boldsymbol{xy}) / (\operatorname{mag}(\boldsymbol{y}) \operatorname{mig}(\boldsymbol{y}))
\end{aligned}$$

$$\begin{aligned}
x^\Delta &= \operatorname{mag}(\boldsymbol{x} - x^c) \\
&= \tfrac{1}{2} \max\{x - y;\ x, y \in \boldsymbol{x}\} \\
x^\Delta &\le \operatorname{mag}(\boldsymbol{x}) - \operatorname{mig}(\boldsymbol{x}) \le 2x^\Delta \\
x \in \boldsymbol{x} \ \Rightarrow \ x^\Delta &\le \operatorname{mag}(\boldsymbol{x} - x) \le 2x^\Delta \\
0 \in \boldsymbol{x} \ \Rightarrow \ x^\Delta &\ge \tfrac{1}{2} \operatorname{mag}(\boldsymbol{x})
\end{aligned}$$

$$\begin{aligned}
\operatorname{rad}(\boldsymbol{xy}) &= \max\bigl\{x^\Delta \operatorname{mag}(\boldsymbol{y}),\ \operatorname{mag}(\boldsymbol{x}) y^\Delta,\ x^\Delta |y^c| + |x^c| y^\Delta\bigr\} \\
\operatorname{rad}(1/\boldsymbol{x}) &= x^\Delta / (\operatorname{mag}(\boldsymbol{x}) \operatorname{mig}(\boldsymbol{x})) \\
\operatorname{rad}(\boldsymbol{x}/\boldsymbol{y}) &= \operatorname{rad}(\boldsymbol{xy}) / (\operatorname{mag}(\boldsymbol{y}) \operatorname{mig}(\boldsymbol{y}))
\end{aligned}$$

$$\begin{aligned}
x^\Delta \operatorname{mag}(\boldsymbol{y}) \le \ \operatorname{rad}(\boldsymbol{xy}) \ &\le x^\Delta \operatorname{mag}(\boldsymbol{y}) + |x^c| y^\Delta \\
\operatorname{mag}(\boldsymbol{x}) y^\Delta \le \ \operatorname{rad}(\boldsymbol{xy}) \ &\le \operatorname{mag}(\boldsymbol{x}) y^\Delta + x^\Delta |y^c| \\
x^\Delta / \operatorname{mig}(\boldsymbol{y}) \le \operatorname{rad}(\boldsymbol{x}/\boldsymbol{y}) &\le \bigl(x^\Delta + y^\Delta |x^c| / \operatorname{mag}(\boldsymbol{y})\bigr) / \operatorname{mig}(\boldsymbol{y}) \\
\operatorname{rad}(\boldsymbol{x}^n) \ &\le n \cdot x^\Delta \operatorname{mag}(\boldsymbol{x})^{n-1}
\end{aligned}$$

$$\begin{aligned}
x \text{ real} \quad &\Rightarrow \quad \operatorname{rad}(x\boldsymbol{y}) = |x| y^\Delta \\
y^c = 0 \quad &\Rightarrow \quad \operatorname{rad}(\boldsymbol{xy}) = \operatorname{mag}(\boldsymbol{x}) y^\Delta \\
0 \in \boldsymbol{x} \cap \boldsymbol{y} \quad &\Rightarrow \quad \operatorname{rad}(\boldsymbol{xy}) \le 2x^\Delta y^\Delta \\
0 \in \boldsymbol{x} \quad &\Rightarrow \quad \operatorname{rad}(\boldsymbol{xy}) \le x^\Delta \operatorname{mag}(\boldsymbol{y}) \\
&\qquad\qquad + |x^c| \max(0, y^\Delta - |y^c|) \\
x^c = 0 \quad &\Rightarrow \quad \operatorname{rad}(\boldsymbol{xy}) = x^\Delta \operatorname{mag}(\boldsymbol{y}) \\
y^\Delta = 0 \quad &\Rightarrow \quad \operatorname{rad}(\boldsymbol{xy}) = x^\Delta \operatorname{mag}(\boldsymbol{y}) \\
0 \in \boldsymbol{x} \ \wedge \ 0 \notin \operatorname{int} \boldsymbol{y} \quad &\Rightarrow \quad \operatorname{rad}(\boldsymbol{xy}) = x^\Delta \operatorname{mag}(\boldsymbol{y}) \\
0 \notin \operatorname{int} \boldsymbol{x} \quad &\Rightarrow \quad \operatorname{rad}(\boldsymbol{xy}) = x^\Delta \operatorname{mig}(\boldsymbol{y}) + \operatorname{mag}(\boldsymbol{x}) y^\Delta \\
0 \in \boldsymbol{x} \quad &\Rightarrow \quad \operatorname{rad}(\boldsymbol{x}/\boldsymbol{y}) = x^\Delta / \operatorname{mig}(\boldsymbol{y}) \\
0 \in \boldsymbol{x} \quad &\Rightarrow \quad \operatorname{rad}(\boldsymbol{x}^n) \le x^\Delta \operatorname{mag}(\boldsymbol{x})^{n-1}
\end{aligned}$$

Table A.4: Arithmetic on intervals: basic algebraic properties ($\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}, \ldots \in \mathbb{IR}$).

commutativity	$\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{y} + \boldsymbol{x}$
associativity	$\boldsymbol{x} + (\boldsymbol{y} + \boldsymbol{z}) = (\boldsymbol{x} + \boldsymbol{y}) + \boldsymbol{z}$
identity element	$\boldsymbol{x} + 0 = 0 + \boldsymbol{x} = \boldsymbol{x}$
cancellation	$\boldsymbol{x} + \boldsymbol{y} = \boldsymbol{x} + \boldsymbol{z} \;\Leftrightarrow\; \boldsymbol{y} = \boldsymbol{z}$
	$\boldsymbol{x} + \boldsymbol{y} \subseteq \boldsymbol{x} + \boldsymbol{z} \;\Leftrightarrow\; \boldsymbol{y} \subseteq \boldsymbol{z}$
subcancellation	$\boldsymbol{x} - \boldsymbol{y} \subseteq (\boldsymbol{x} + \boldsymbol{z}) - (\boldsymbol{y} + \boldsymbol{z})$
commutativity	$\boldsymbol{x}\boldsymbol{y} = \boldsymbol{y}\boldsymbol{x}$
associativity	$\boldsymbol{x}(\boldsymbol{y}\boldsymbol{z}) = (\boldsymbol{x}\boldsymbol{y})\boldsymbol{z}$
identity element	$1 \cdot \boldsymbol{x} = \boldsymbol{x} \cdot 1 = \boldsymbol{x}$
cancellation	$(\boldsymbol{x}\boldsymbol{y} = \boldsymbol{x}\boldsymbol{z} \;\wedge\; 0 \notin \boldsymbol{x}) \;\Rightarrow\; \boldsymbol{y} = \boldsymbol{z}$
	$(\boldsymbol{x}\boldsymbol{y} \subseteq \boldsymbol{x}\boldsymbol{z} \;\wedge\; 0 \notin \boldsymbol{x}) \;\Rightarrow\; \boldsymbol{y} \subseteq \boldsymbol{z}$
subcancellation	$\boldsymbol{x}/\boldsymbol{y} \subseteq (\boldsymbol{x}\boldsymbol{z})/(\boldsymbol{y}\boldsymbol{z})$
(sub)distributivity	$\boldsymbol{x}(\boldsymbol{y} + \boldsymbol{z}) \subseteq \boldsymbol{x}\boldsymbol{y} + \boldsymbol{x}\boldsymbol{z}$
	x real $\;\Rightarrow\; x(\boldsymbol{y} + \boldsymbol{z}) = x\boldsymbol{y} + x\boldsymbol{z}$
	$\inf(\boldsymbol{y}\boldsymbol{z}) \geq 0 \;\Rightarrow\; \boldsymbol{x}(\boldsymbol{y} + \boldsymbol{z}) = \boldsymbol{x}\boldsymbol{y} + \boldsymbol{x}\boldsymbol{z}$
no zero divisors	$0 = \boldsymbol{x}\boldsymbol{y} \;\Leftrightarrow\; (\boldsymbol{x} = 0 \;\vee\; \boldsymbol{y} = 0)$
	$0 \in \boldsymbol{x}\boldsymbol{y} \;\Leftrightarrow\; (0 \in \boldsymbol{x} \;\vee\; 0 \in \boldsymbol{y})$
inclusion isotonicity for $\circ \in \{+, -, \cdot, /\}$:	$(\boldsymbol{x}_1 \subseteq \boldsymbol{x}_2 \;\wedge\; \boldsymbol{y}_1 \subseteq \boldsymbol{y}_2) \;\Rightarrow\; (\boldsymbol{x}_1 \circ \boldsymbol{y}_1) \subseteq (\boldsymbol{x}_2 \circ \boldsymbol{y}_2)$

Table A.5: Further properties of intervals ($\boldsymbol{x}, \boldsymbol{y} \in \mathbb{IR}$, $n \in \mathbb{N}$).

$\boldsymbol{x} = x^c + x^\Delta[-1, 1]$		
$\boldsymbol{x} - \boldsymbol{x} = 2x^\Delta[-1, 1]$		
$\boldsymbol{x} + \boldsymbol{x} = 2\boldsymbol{x}$		
$(-\boldsymbol{x})(-\boldsymbol{y}) = \boldsymbol{x}\boldsymbol{y}$		
$\boldsymbol{x}/\boldsymbol{y} = \boldsymbol{x}(1/\boldsymbol{y})$		
$\boldsymbol{x} - \boldsymbol{x} = 0$	$\Leftrightarrow$	$x^\Delta = 0$
$\boldsymbol{x} \cdot \boldsymbol{x} = \boldsymbol{x}^2$	$\Leftrightarrow$	$0 \notin \operatorname{int} \boldsymbol{x}$
$\boldsymbol{x} > 0 \;\vee\; n$ odd	$\Rightarrow$	$\boldsymbol{x}^n = [\underline{x}^n, \overline{x}^n]$
$\boldsymbol{x} < 0 \;\wedge\; n$ even	$\Rightarrow$	$\boldsymbol{x}^n = [\overline{x}^n, \underline{x}^n]$
$0 \in \boldsymbol{x} \;\wedge\; n$ even	$\Rightarrow$	$\boldsymbol{x}^n = [0, \operatorname{mag}(\boldsymbol{x})^n]$

Table A.6: Symmetric intervals ($\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z} \in \mathbb{IR}$).

$\boldsymbol{x}$ is symmetric (def.)	$\Leftrightarrow$	$x^c = 0$
	$\Leftrightarrow$	$\underline{x} = -\overline{x}$
$\boldsymbol{x}, \boldsymbol{y}$ are symmetric	$\Rightarrow$	$\boldsymbol{x} \pm \boldsymbol{y}$ are symmetric
$\boldsymbol{x}, \boldsymbol{y}$ are symmetric	$\Rightarrow$	$\boldsymbol{xy}$ is symmetric
$x^c = y^c = 0$	$\Rightarrow$	$\mathrm{mid}(\boldsymbol{x} \pm \boldsymbol{y}) = 0$
$x^c = y^c = 0$	$\Rightarrow$	$\mathrm{mid}(\boldsymbol{xy}) = 0$
$(x^c = y^c = 0 \ \wedge \ 0 \notin \boldsymbol{y})$	$\Rightarrow$	$\mathrm{mid}(\boldsymbol{x}/\boldsymbol{y}) = 0$
$x^c = 0$	$\Rightarrow$	$x^\Delta = \mathrm{mag}(\boldsymbol{x})$
$x^c = 0$	$\Rightarrow$	$\boldsymbol{x} = x^\Delta[-1, 1]$
$x^c = y^c = 0$	$\Rightarrow$	$\boldsymbol{x} \pm \boldsymbol{y} = (x^\Delta + y^\Delta)[-1, 1]$
$x^c = y^c = 0$	$\Rightarrow$	$\boldsymbol{xy} = (x^\Delta y^\Delta)[-1, 1]$
$y^c = 0$	$\Rightarrow$	$\boldsymbol{xy} = \mathrm{mag}(\boldsymbol{x})\boldsymbol{y}$ $= (\mathrm{mag}(\boldsymbol{x})y^\Delta)[-1, 1]$
$x^c = y^c = z^c = 0$	$\Rightarrow$	$\boldsymbol{x}(\boldsymbol{y} \pm \boldsymbol{z}) = \boldsymbol{xy} \pm \boldsymbol{xz}$ $= x^\Delta(y^\Delta + z^\Delta)[-1, 1]$

Table A.7: The mid-rad representation and arithmetic on intervals ($\boldsymbol{x}, \boldsymbol{y} \in \mathbb{IR}$ and $\boldsymbol{A}, \boldsymbol{B}$ interval matrices of suitable sizes). The mid-rad interval form is defined $\langle x^c, x^\Delta \rangle := [x^c - x^\Delta, x^c + x^\Delta]$

		$\langle x^c, x^\Delta \rangle \langle y^c, y^\Delta \rangle \subseteq \langle x^c y^c, \ x^\Delta \lvert y^c \rvert + (\lvert x^c \rvert + x^\Delta) y^\Delta \rangle$
$\boldsymbol{x}, \boldsymbol{y} \geq 0$	$\Rightarrow$	$\langle x^c, x^\Delta \rangle \cdot \langle y^c, y^\Delta \rangle = \langle x^c y^c + x^\Delta y^\Delta, \ x^\Delta y^c + x^c y^\Delta \rangle$
$\boldsymbol{x}, \boldsymbol{y} \geq 0$	$\Rightarrow$	$\dfrac{\langle x^c, x^\Delta \rangle}{\langle y^c, y^\Delta \rangle} = \left\langle \dfrac{x^c y^c + x^\Delta y^\Delta}{(y^c)^2 - (y^\Delta)^2}, \ \dfrac{x^\Delta y^c + x^c y^\Delta}{(y^c)^2 - (y^\Delta)^2} \right\rangle$
$\boldsymbol{x}, \boldsymbol{y} \geq 0$	$\Rightarrow$	$\dfrac{\langle x^c, x^\Delta \rangle}{\langle y^c, y^\Delta \rangle} \subseteq \left\langle \dfrac{x^c}{y^c}, \ \dfrac{x^\Delta + \frac{x^c}{y^c} y^\Delta}{y^c - y^\Delta} \right\rangle$
		$\langle A^c, A^\Delta \rangle \langle B^c, B^\Delta \rangle \subseteq \langle A^c B^c, \ A^\Delta(\lvert B^c \rvert + B^\Delta) + \lvert A^c \rvert B^\Delta \rangle$
B real	$\Rightarrow$	$\langle A^c, A^\Delta \rangle B = \langle A^c B, A^\Delta \lvert B \rvert \rangle$

Table A.8: Magnitude and mignitude ($\boldsymbol{x}, \boldsymbol{y} \in \mathbb{IR}$).

$$
\begin{aligned}
\operatorname{mag}(\boldsymbol{x}) &= |x^c| + x^\Delta \\
&= \max\{|\underline{x}|, |\overline{x}|\} \\
&= \max\{|x|;\ x \in \boldsymbol{x}\} \\
[-1,1]\boldsymbol{x} &= [-1,1]\operatorname{mag}(\boldsymbol{x}) \\
2x^\Delta &= \operatorname{mag}(\boldsymbol{x} - \boldsymbol{x})
\end{aligned}
$$

$$
\begin{aligned}
\boldsymbol{x} = 0 \quad &\Leftrightarrow \quad \operatorname{mag}(\boldsymbol{x}) = 0 \\
0 \notin \operatorname{int} \boldsymbol{x} \quad &\Leftrightarrow \quad \operatorname{mig}(\boldsymbol{x}) = |x^c| - x^\Delta \\
\boldsymbol{x} \subseteq \boldsymbol{y} \quad &\Rightarrow \quad \operatorname{mag}(\boldsymbol{x}) \leq \operatorname{mag}(\boldsymbol{y}) \\
\boldsymbol{x} \subseteq \boldsymbol{y} \quad &\Rightarrow \quad \operatorname{mig}(\boldsymbol{x}) \geq \operatorname{mig}(\boldsymbol{y})
\end{aligned}
$$

$$
\begin{aligned}
\operatorname{mag}(\boldsymbol{x} \pm \boldsymbol{y}) &\leq \operatorname{mag}(\boldsymbol{x}) + \operatorname{mag}(\boldsymbol{y}) \\
\operatorname{mag}(\boldsymbol{x} \pm \boldsymbol{y}) &\geq \operatorname{mag}(\boldsymbol{x}) - \operatorname{mig}(\boldsymbol{y}) \\
y^c = 0 \quad \Rightarrow \quad \operatorname{mag}(\boldsymbol{x} \pm \boldsymbol{y}) &= \operatorname{mag}(\boldsymbol{x}) + \operatorname{mag}(\boldsymbol{y})
\end{aligned}
$$

$$
\begin{aligned}
\operatorname{mag}(\boldsymbol{x}\boldsymbol{y}) &= \operatorname{mag}(\boldsymbol{x}) \operatorname{mag}(\boldsymbol{y}) \\
\operatorname{mag}(\boldsymbol{x}/\boldsymbol{y}) &= \operatorname{mag}(\boldsymbol{x}) / \operatorname{mig}(\boldsymbol{y}) \\
\operatorname{mag}(1/\boldsymbol{x}) &= 1 / \operatorname{mig}(\boldsymbol{x}) \\
x \text{ real} \quad \Rightarrow \quad \operatorname{mag}(x\boldsymbol{y}) &= |x| \operatorname{mag}(\boldsymbol{y}) \\
x^c = 0 \quad \Rightarrow \quad \operatorname{mag}(\boldsymbol{x}\boldsymbol{y}) &= x^\Delta \operatorname{mag}(\boldsymbol{y}) \\
x^c = 0 \quad \Rightarrow \quad \boldsymbol{x}\boldsymbol{y} &= \boldsymbol{x} \operatorname{mag}(\boldsymbol{y})
\end{aligned}
$$

$$
\begin{aligned}
\operatorname{mig}(\boldsymbol{x}) &\geq |x^c| - x^\Delta \\
\operatorname{mig}(\boldsymbol{x} \pm \boldsymbol{y}) &\leq \operatorname{mig}(\boldsymbol{x}) + \operatorname{mag}(\boldsymbol{y}) \\
\operatorname{mig}(\boldsymbol{x} \pm \boldsymbol{y}) &\geq \operatorname{mig}(\boldsymbol{x}) - \operatorname{mag}(\boldsymbol{y}) \\
\operatorname{mig}(\boldsymbol{x}\boldsymbol{y}) &= \operatorname{mig}(\boldsymbol{x}) \operatorname{mig}(\boldsymbol{y}) \\
\operatorname{mig}(\boldsymbol{x}/\boldsymbol{y}) &= \operatorname{mig}(\boldsymbol{x}) / \operatorname{mag}(\boldsymbol{y}) \\
\operatorname{mig}(1/\boldsymbol{x}) &= 1 / \operatorname{mag}(\boldsymbol{x})
\end{aligned}
$$

Table A.9: (Hausdorff) distance on intervals ($\boldsymbol{x}, \boldsymbol{x}', \boldsymbol{x}_k, \boldsymbol{y}, \boldsymbol{y}', \boldsymbol{z} \in \mathbb{IR}$ and $n \in \mathbb{N}$).

distance definition	$\begin{aligned}\operatorname{dist}(\boldsymbol{x},\boldsymbol{y}) &= \max\{|\underline{x}-\underline{y}|, |\overline{x}-\overline{y}|\}\\ &= |x^c - y^c| + |x^\Delta - y^\Delta|\\ &= \min\{q \geq 0;\, \boldsymbol{x} \subseteq \boldsymbol{y} + [-q,q],\, \boldsymbol{y} \subseteq \boldsymbol{x} + [-q,q]\}\end{aligned}$
metric axioms	$\operatorname{dist}(\boldsymbol{x},\boldsymbol{y}) \geq 0$ $\operatorname{dist}(\boldsymbol{x},\boldsymbol{y}) = 0 \;\Leftrightarrow\; \boldsymbol{x} = \boldsymbol{y}$ $\operatorname{dist}(\boldsymbol{x},\boldsymbol{y}) = \operatorname{dist}(\boldsymbol{y},\boldsymbol{x})$ $\operatorname{dist}(\boldsymbol{x},\boldsymbol{y}) \leq \operatorname{dist}(\boldsymbol{x},\boldsymbol{z}) + \operatorname{dist}(\boldsymbol{z},\boldsymbol{y})$
basic properties	$\begin{aligned}\operatorname{mag}(\boldsymbol{x}) &= \operatorname{dist}(\boldsymbol{x},0)\\ &= \operatorname{dist}(\boldsymbol{x}+\boldsymbol{y},\boldsymbol{y})\end{aligned}$ $\operatorname{dist}(\boldsymbol{x},\boldsymbol{y}) = \operatorname{dist}(\boldsymbol{x}+\boldsymbol{z},\boldsymbol{y}+\boldsymbol{z})$ $\operatorname{dist}(\boldsymbol{x}+\boldsymbol{x}',\boldsymbol{y}+\boldsymbol{y}') \leq \operatorname{dist}(\boldsymbol{x},\boldsymbol{y}) + \operatorname{dist}(\boldsymbol{x}',\boldsymbol{y}')$ $\operatorname{dist}(\boldsymbol{z}\boldsymbol{x},\boldsymbol{z}\boldsymbol{y}) \leq \operatorname{mag}(\boldsymbol{z})\operatorname{dist}(\boldsymbol{x},\boldsymbol{y})$ $\operatorname{dist}(\boldsymbol{x}/\boldsymbol{z},\boldsymbol{y}/\boldsymbol{z}) \leq \operatorname{dist}(\boldsymbol{x},\boldsymbol{y})/\operatorname{mig}(\boldsymbol{z}), \quad 0 \notin \boldsymbol{z}$ $\operatorname{dist}(\boldsymbol{z}/\boldsymbol{x},\boldsymbol{z}/\boldsymbol{y}) \leq \frac{\operatorname{mag}(\boldsymbol{z})}{\operatorname{mig}(\boldsymbol{x})\operatorname{mig}(\boldsymbol{y})}\operatorname{dist}(\boldsymbol{x},\boldsymbol{y})$ $\operatorname{dist}(\boldsymbol{x},\boldsymbol{y}) \geq |x^\Delta - y^\Delta|$ $\operatorname{dist}(\boldsymbol{x},\boldsymbol{y}) \geq |\operatorname{mag}(\boldsymbol{x}) - \operatorname{mag}(\boldsymbol{y})|$ $\operatorname{dist}(\boldsymbol{x},\boldsymbol{y}) \geq \operatorname{mag}(\boldsymbol{x}-\boldsymbol{y}) - 2x^\Delta$ $\boldsymbol{x} \subseteq \boldsymbol{y} + \operatorname{dist}(\boldsymbol{x},\boldsymbol{y})[-1,1]$
further properties	$\begin{aligned} y \text{ real} &\;\Rightarrow\; \operatorname{dist}(\boldsymbol{x},y) = \operatorname{mag}(\boldsymbol{x}-y)\\ z \text{ real} &\;\Rightarrow\; \operatorname{dist}(z\boldsymbol{x},z\boldsymbol{y}) = |z|\operatorname{dist}(\boldsymbol{x},\boldsymbol{y})\\ \boldsymbol{x} \subseteq \boldsymbol{y} &\;\Rightarrow\; \operatorname{dist}(\boldsymbol{x},\boldsymbol{y}) = \operatorname{mag}(\boldsymbol{y} - x^c) - x^\Delta\\ \boldsymbol{x} \subseteq \boldsymbol{y} &\;\Rightarrow\; \operatorname{dist}(\boldsymbol{x},\boldsymbol{y}) \leq 2(y^\Delta - x^\Delta)\\ \boldsymbol{x} \subseteq \boldsymbol{y} &\;\Rightarrow\; \operatorname{dist}(\boldsymbol{x},\boldsymbol{y}) \geq y^\Delta - x^\Delta\\ \boldsymbol{x} \subseteq \boldsymbol{y} \subseteq \boldsymbol{z} &\;\Rightarrow\; \operatorname{dist}(\boldsymbol{x},\boldsymbol{y}) \leq \operatorname{dist}(\boldsymbol{x},\boldsymbol{z})\\ \boldsymbol{x} \subseteq \boldsymbol{y} \subseteq \boldsymbol{z} &\;\Rightarrow\; \operatorname{dist}(\boldsymbol{y},\boldsymbol{z}) \leq \operatorname{dist}(\boldsymbol{x},\boldsymbol{z})\\ \boldsymbol{x} \subseteq \boldsymbol{y} \ni y &\;\Rightarrow\; \operatorname{dist}(y,\boldsymbol{x}) \leq \operatorname{dist}(y,\boldsymbol{y})\\ (\boldsymbol{x} \subseteq \boldsymbol{y} \subseteq \boldsymbol{z} \;\wedge\; \boldsymbol{x} \subseteq \boldsymbol{y}' \subseteq \boldsymbol{z}) &\;\Rightarrow\; \operatorname{dist}(\boldsymbol{y},\boldsymbol{y}') \leq \max\{\operatorname{dist}(\boldsymbol{x},\boldsymbol{y}), \operatorname{dist}(\boldsymbol{y},\boldsymbol{z})\}\\ \boldsymbol{x},\boldsymbol{y} \subseteq \boldsymbol{z} &\;\Rightarrow\; \operatorname{dist}(\boldsymbol{y}^n,\boldsymbol{z}^n) \leq n \cdot \operatorname{mag}(\boldsymbol{z})^{n-1}\operatorname{dist}(\boldsymbol{x},\boldsymbol{y})\end{aligned}$
convergence	$\begin{aligned}\lim_{k\to\infty} \boldsymbol{x}_k = \boldsymbol{y} &\;\Leftrightarrow\; \left(\lim_{k\to\infty}\underline{x}_k = \underline{y} \;\wedge\; \lim_{k\to\infty}\overline{x}_k = \overline{y}\right)\\ &\;\Leftrightarrow\; \left(\lim_{k\to\infty} x_k^c = y^c \;\wedge\; \lim_{k\to\infty} x_k^\Delta = y^\Delta\right)\\ \boldsymbol{x}_1 \supseteq \boldsymbol{x}_2 \supseteq \ldots &\;\Rightarrow\; \lim_{k\to\infty}\boldsymbol{x}_k = \cap_{k=1}^\infty \boldsymbol{x}_k\end{aligned}$ $\operatorname{mid}(\lim_{k\to\infty}\boldsymbol{x}_k) = \lim_{k\to\infty} x_k^c$ $\operatorname{rad}(\lim_{k\to\infty}\boldsymbol{x}_k) = \lim_{k\to\infty} x_k^\Delta$ the metric space $(\mathbb{IR}, \operatorname{dist})$ is complete arithmetic operations $+, -, \cdot$ and $/$ are continuous functions $\operatorname{mid}, \operatorname{rad}, \operatorname{inf}, \operatorname{sup}, \operatorname{mag}$ and mig are continuous

Table A.10: Arithmetic on complex intervals in rectangular form ($\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}}, \tilde{\boldsymbol{z}} \in \mathbb{IC}$).

commutativity	$\tilde{\boldsymbol{x}} + \tilde{\boldsymbol{y}} = \tilde{\boldsymbol{y}} + \tilde{\boldsymbol{x}}$
associativity	$\tilde{\boldsymbol{x}} + (\tilde{\boldsymbol{y}} + \tilde{\boldsymbol{z}}) = (\tilde{\boldsymbol{x}} + \tilde{\boldsymbol{y}}) + \tilde{\boldsymbol{z}}$
identity element	$\tilde{\boldsymbol{x}} + 0 = 0 + \tilde{\boldsymbol{x}} = \tilde{\boldsymbol{x}}$
commutativity	$\tilde{\boldsymbol{x}}\tilde{\boldsymbol{y}} = \tilde{\boldsymbol{y}}\tilde{\boldsymbol{x}}$
¬ associativity	$\tilde{\boldsymbol{x}}(\tilde{\boldsymbol{y}}\tilde{\boldsymbol{z}}) \neq (\tilde{\boldsymbol{x}}\tilde{\boldsymbol{y}})\tilde{\boldsymbol{z}}$
identity element	$1 \cdot \tilde{\boldsymbol{x}} = \tilde{\boldsymbol{x}} \cdot 1 = \tilde{\boldsymbol{x}}$
subdistributivity	$\tilde{\boldsymbol{x}}(\tilde{\boldsymbol{y}} + \tilde{\boldsymbol{z}}) \subseteq \tilde{\boldsymbol{x}}\tilde{\boldsymbol{y}} + \tilde{\boldsymbol{x}}\tilde{\boldsymbol{z}}$
	$x(\tilde{\boldsymbol{y}} + \tilde{\boldsymbol{z}}) = x\tilde{\boldsymbol{y}} + x\tilde{\boldsymbol{z}}$
no zero divisors	$\tilde{\boldsymbol{x}}\tilde{\boldsymbol{y}} = 0 \;\Leftrightarrow\; (\tilde{\boldsymbol{x}} = 0 \;\vee\; \tilde{\boldsymbol{y}} = 0)$
inclusion isotonicity	
for $\circ \in \{+, -, \cdot, /\}$:	$(\tilde{\boldsymbol{x}}_1 \subseteq \tilde{\boldsymbol{x}}_2 \;\wedge\; \tilde{\boldsymbol{y}}_1 \subseteq \tilde{\boldsymbol{y}}_2) \;\Rightarrow\; (\tilde{\boldsymbol{x}}_1 \circ \tilde{\boldsymbol{y}}_1) \subseteq (\tilde{\boldsymbol{x}}_2 \circ \tilde{\boldsymbol{y}}_2)$

Table A.11: Arithmetic on complex intervals in circular form ($\tilde{\boldsymbol{x}}, \tilde{\boldsymbol{y}}, \tilde{\boldsymbol{z}} \in \mathbb{IC}$).

commutativity	$\tilde{\boldsymbol{x}} + \tilde{\boldsymbol{y}} = \tilde{\boldsymbol{y}} + \tilde{\boldsymbol{x}}$
associativity	$\tilde{\boldsymbol{x}} + (\tilde{\boldsymbol{y}} + \tilde{\boldsymbol{z}}) = (\tilde{\boldsymbol{x}} + \tilde{\boldsymbol{y}}) + \tilde{\boldsymbol{z}}$
identity element	$\tilde{\boldsymbol{x}} + 0 = 0 + \tilde{\boldsymbol{x}} = \tilde{\boldsymbol{x}}$
commutativity	$\tilde{\boldsymbol{x}}\tilde{\boldsymbol{y}} = \tilde{\boldsymbol{y}}\tilde{\boldsymbol{x}}$
associativity	$\tilde{\boldsymbol{x}}(\tilde{\boldsymbol{y}}\tilde{\boldsymbol{z}}) = (\tilde{\boldsymbol{x}}\tilde{\boldsymbol{y}})\tilde{\boldsymbol{z}}$
identity element	$1 \cdot \tilde{\boldsymbol{x}} = \tilde{\boldsymbol{x}} \cdot 1 = \tilde{\boldsymbol{x}}$
subdistributivity	$\tilde{\boldsymbol{x}}(\tilde{\boldsymbol{y}} + \tilde{\boldsymbol{z}}) \subseteq \tilde{\boldsymbol{x}}\tilde{\boldsymbol{y}} + \tilde{\boldsymbol{x}}\tilde{\boldsymbol{z}}$
	$x(\tilde{\boldsymbol{y}} + \tilde{\boldsymbol{z}}) = x\tilde{\boldsymbol{y}} + x\tilde{\boldsymbol{z}}$
no zero divisors	$\tilde{\boldsymbol{x}}\tilde{\boldsymbol{y}} = 0 \;\Leftrightarrow\; (\tilde{\boldsymbol{x}} = 0 \;\vee\; \tilde{\boldsymbol{y}} = 0)$
inclusion isotonicity	
for $\circ \in \{+, -, \cdot, /\}$:	$(\tilde{\boldsymbol{x}}_1 \subseteq \tilde{\boldsymbol{x}}_2 \;\wedge\; \tilde{\boldsymbol{y}}_1 \subseteq \tilde{\boldsymbol{y}}_2) \;\Rightarrow\; (\tilde{\boldsymbol{x}}_1 \circ \tilde{\boldsymbol{y}}_1) \subseteq (\tilde{\boldsymbol{x}}_2 \circ \tilde{\boldsymbol{y}}_2)$

Appendix B

Properties of interval matrices

Herein, we extend properties of interval arithmetic (Appendix A) to arithmetic of interval matrices. We include also properties of certain types of interval matrices.

Table B.1: Set operations on interval matricess ($\boldsymbol{A}, \boldsymbol{B}, \ldots \in \mathbb{IR}^{m\times n}$).

$\boldsymbol{A} \subseteq \boldsymbol{B}$	$\Leftrightarrow$	$\lvert A^c - B^c\rvert \le B^\Delta - A^\Delta$
$\boldsymbol{A} \subseteq \boldsymbol{B}$	$\Leftrightarrow$	$\mathrm{mag}(\boldsymbol{A} - B^c) \le B^\Delta$
$\boldsymbol{A} \subseteq \boldsymbol{B}$	$\Rightarrow$	$A^\Delta \le B^\Delta \ \wedge\ \mathrm{mag}(\boldsymbol{A}) \le \mathrm{mag}(\boldsymbol{B})$
$\boldsymbol{A} \cap \boldsymbol{B} \neq \emptyset$	$\Leftrightarrow$	$\lvert A^c - B^c\rvert \le A^\Delta + B^\Delta$
$\boldsymbol{A} \cap \boldsymbol{B} \neq \emptyset$	$\Leftrightarrow$	$0 \in \boldsymbol{A} - \boldsymbol{B}$
$\boldsymbol{A} \cap \boldsymbol{B} \neq \emptyset$	$\Leftrightarrow$	$\overline{A} \ge \underline{B} \ \wedge\ \underline{A} \le \overline{B}$
$\boldsymbol{A} \cap \boldsymbol{B} \neq \emptyset$	$\Leftrightarrow$	$\boldsymbol{A} \cap \boldsymbol{B} = [\max(\underline{A}, \underline{B}), \min(\overline{A}, \overline{B})]$
$\boldsymbol{A}_1 \subseteq \boldsymbol{B}_1 \ \wedge\ \boldsymbol{A}_2 \subseteq \boldsymbol{B}_2$	$\Rightarrow$	$(\boldsymbol{A}_1 \cap \boldsymbol{A}_2) \subseteq (\boldsymbol{B}_1 \cap \boldsymbol{B}_2)$

Table B.2: Arithmetic on interval matrices: basic algebraic properties ($\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}, \ldots$ are interval matrices of suitable sizes and $\boldsymbol{a} \in \mathbb{IR}$).

commutativity	$\boldsymbol{A} + \boldsymbol{B} = \boldsymbol{B} + \boldsymbol{A}$
associativity	$\boldsymbol{A} + (\boldsymbol{B} + \boldsymbol{C}) = (\boldsymbol{A} + \boldsymbol{B}) + \boldsymbol{C}$
	a real $\Rightarrow\ \boldsymbol{A}(a\boldsymbol{B}) = a(\boldsymbol{A}\boldsymbol{B})$
subdistributivity	$\boldsymbol{a}(\boldsymbol{B} + \boldsymbol{C}) \subseteq \boldsymbol{a}\boldsymbol{B} + \boldsymbol{a}\boldsymbol{C}$
identity element	$\boldsymbol{A} + 0 = 0 + \boldsymbol{A} = \boldsymbol{A}$
cancellation	$\boldsymbol{A} + \boldsymbol{B} = \boldsymbol{A} + \boldsymbol{C} \ \Leftrightarrow\ \boldsymbol{B} = \boldsymbol{C}$
	$\boldsymbol{A} + \boldsymbol{B} \subseteq \boldsymbol{A} + \boldsymbol{C} \ \Leftrightarrow\ \boldsymbol{B} \subseteq \boldsymbol{C}$
associativity	$(\boldsymbol{A}\boldsymbol{B})\boldsymbol{C} \neq \boldsymbol{A}(\boldsymbol{B}\boldsymbol{C})$
	A real $\Rightarrow\ (A\boldsymbol{B})\boldsymbol{C} \subseteq A(\boldsymbol{B}\boldsymbol{C})$
	C real $\Rightarrow\ (\boldsymbol{A}\boldsymbol{B})C \supseteq \boldsymbol{A}(\boldsymbol{B}C)$
	A, C real $\Rightarrow\ (A\boldsymbol{B})C = A(\boldsymbol{B}C)$
	A, B real $\Rightarrow\ (AB)\boldsymbol{C} \subseteq A(B\boldsymbol{C})$
	$\boldsymbol{B}, \boldsymbol{C} \ge 0 \ \Rightarrow\ (\boldsymbol{A}\boldsymbol{B})\boldsymbol{C} \subseteq \boldsymbol{A}(\boldsymbol{B}\boldsymbol{C})$
	$\boldsymbol{A}, \boldsymbol{B} \ge 0 \ \Rightarrow\ (\boldsymbol{A}\boldsymbol{B})\boldsymbol{C} \supseteq \boldsymbol{A}(\boldsymbol{B}\boldsymbol{C})$
	$A^c = 0$, C real $\Rightarrow\ (\boldsymbol{A}\boldsymbol{B})C \supseteq \boldsymbol{A}(\boldsymbol{B}C)$
	$\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C} \ge 0 \ \Rightarrow\ (\boldsymbol{A}\boldsymbol{B})\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{B}\boldsymbol{C})$
	$\underline{A} = 0$, $\underline{B} = 0 \ \Rightarrow\ (\boldsymbol{A}\boldsymbol{B})\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{B}\boldsymbol{C})$
	$A^c = 0$, $B^c = 0 \ \Rightarrow\ (\boldsymbol{A}\boldsymbol{B})\boldsymbol{C} = \boldsymbol{A}(\boldsymbol{B}\boldsymbol{C})$
(sub)distributivity	$\boldsymbol{A}(\boldsymbol{B} + \boldsymbol{C}) \subseteq \boldsymbol{A}\boldsymbol{B} + \boldsymbol{A}\boldsymbol{C}$
	$(\boldsymbol{A} + \boldsymbol{B})\boldsymbol{C} \subseteq \boldsymbol{A}\boldsymbol{C} + \boldsymbol{B}\boldsymbol{C}$
	A real $\Rightarrow\ A(\boldsymbol{B} + \boldsymbol{C}) = A\boldsymbol{B} + A\boldsymbol{C}$
	C real $\Rightarrow\ (\boldsymbol{A} + \boldsymbol{B})C = \boldsymbol{A}C + \boldsymbol{B}C$
	$\boldsymbol{B}, \boldsymbol{C} \ge 0 \ \Rightarrow\ \boldsymbol{A}(\boldsymbol{B} + \boldsymbol{C}) = \boldsymbol{A}\boldsymbol{B} + \boldsymbol{A}\boldsymbol{C}$
	$\boldsymbol{B}, \boldsymbol{C} \le 0 \ \Rightarrow\ \boldsymbol{A}(\boldsymbol{B} + \boldsymbol{C}) = \boldsymbol{A}\boldsymbol{B} + \boldsymbol{A}\boldsymbol{C}$
inclusion isotonicity for $\circ \in \{+, -, \cdot\}$:	$(\boldsymbol{A}_1 \subseteq \boldsymbol{A}_2 \ \wedge\ \boldsymbol{B}_1 \subseteq \boldsymbol{B}_2) \ \Rightarrow\ (\boldsymbol{A}_1 \circ \boldsymbol{B}_1) \subseteq (\boldsymbol{A}_2 \circ \boldsymbol{B}_2)$

Table B.3: Interval matrix midpoints and radii ($\boldsymbol{A}, \boldsymbol{B}$ are interval matrices of suitable sizes).

$\operatorname{mid}(\boldsymbol{A} \pm \boldsymbol{B}) = A^c \pm B^c$
$\operatorname{rad}(\boldsymbol{A} \pm \boldsymbol{B}) = A^\Delta + B^\Delta$
A real $\vee$ B real $\Rightarrow$ $\operatorname{mid}(\boldsymbol{AB}) = A^c B^c$
$A^c = 0 \ \vee \ B^c = 0 \Rightarrow \operatorname{mid}(\boldsymbol{AB}) = A^c B^c$
$a \in \mathbb{R} \Rightarrow \operatorname{rad}(a\boldsymbol{A}) = \lvert a \rvert A^\Delta$
$A^\Delta \operatorname{mag}(\boldsymbol{B}) \leq \operatorname{rad}(\boldsymbol{AB}) \leq A^\Delta \operatorname{mag}(\boldsymbol{B}) + \lvert A^c \rvert B^\Delta$
$\operatorname{mag}(\boldsymbol{A}) B^\Delta \leq \operatorname{rad}(\boldsymbol{AB}) \leq \operatorname{mag}(\boldsymbol{A}) B^\Delta + A^\Delta \lvert B^c \rvert$
A real $\Rightarrow$ $\operatorname{rad}(A\boldsymbol{B}) = \lvert A \rvert B^\Delta$
B real $\Rightarrow$ $\operatorname{rad}(\boldsymbol{A}B) = A^\Delta \lvert B \rvert$
$B^c = 0 \Rightarrow \operatorname{rad}(\boldsymbol{AB}) = \operatorname{mag}(\boldsymbol{A}) B^\Delta$
$0 \in \boldsymbol{A} \cap \boldsymbol{B} \Rightarrow \operatorname{rad}(\boldsymbol{AB}) \leq 2 A^\Delta B^\Delta$
$0 \in \boldsymbol{A} \Rightarrow \frac{1}{2} \operatorname{mag}(\boldsymbol{A}) \leq A^\Delta \leq \operatorname{mag}(\boldsymbol{A})$

Table B.4: Interval matrix magnitude and mignitude ($a \in \mathbb{R}$ and $\boldsymbol{A}, \boldsymbol{B}$ are interval matrices of suitable sizes).

$\operatorname{mag}(\boldsymbol{A}) = \lvert A^c \rvert + A^\Delta$
$= \max\{\lvert \underline{A} \rvert, \lvert \overline{A} \rvert\}$
$= \max\{\lvert A \rvert;\ A \in \boldsymbol{A}\}$
$[-1, 1]\boldsymbol{A} = [-\operatorname{mag}(\boldsymbol{A}), \operatorname{mag}(\boldsymbol{A})]$
$2A^\Delta = \operatorname{mag}(\boldsymbol{A} - \boldsymbol{A})$
$\operatorname{mag}(a\boldsymbol{A}) = \lvert a \rvert \operatorname{mag}(\boldsymbol{A})$
$\boldsymbol{A} = 0 \Leftrightarrow \operatorname{mag}(\boldsymbol{A}) = 0$
$\boldsymbol{A} \subseteq \boldsymbol{B} \Rightarrow \operatorname{mag}(\boldsymbol{A}) \leq \operatorname{mag}(\boldsymbol{B})$
$\boldsymbol{A} \subseteq \boldsymbol{B} \Rightarrow \operatorname{mig}(\boldsymbol{A}) \geq \operatorname{mig}(\boldsymbol{B})$
$\operatorname{mag}(\boldsymbol{A} \pm \boldsymbol{B}) \leq \operatorname{mag}(\boldsymbol{A}) + \operatorname{mag}(\boldsymbol{B})$
$\operatorname{mag}(\boldsymbol{A} \pm \boldsymbol{B}) \geq \operatorname{mag}(\boldsymbol{A}) - \operatorname{mig}(\boldsymbol{B})$
$\operatorname{mag}(\boldsymbol{AB}) \leq \operatorname{mag}(\boldsymbol{A}) \operatorname{mag}(\boldsymbol{B})$
$\operatorname{mag}(\boldsymbol{AB}) \geq \langle \boldsymbol{A} \rangle \operatorname{mag}(\boldsymbol{B})$
$B^c = 0 \Rightarrow \operatorname{mag}(\boldsymbol{A} \pm \boldsymbol{B}) = \operatorname{mag}(\boldsymbol{A}) + \operatorname{mag}(\boldsymbol{B})$
$B^c = 0 \Rightarrow \operatorname{mag}(\boldsymbol{AB}) = \operatorname{mag}(\boldsymbol{A}) \operatorname{mag}(\boldsymbol{B})$
$B^c = 0 \Rightarrow \boldsymbol{AB} = \operatorname{mag}(\boldsymbol{A}) \boldsymbol{B}$
$\operatorname{mig}(\boldsymbol{A}) \geq \lvert A^c \rvert - A^\Delta$
$\operatorname{mig}(\boldsymbol{A} \pm \boldsymbol{B}) \leq \operatorname{mig}(\boldsymbol{A}) + \operatorname{mag}(\boldsymbol{B})$
$\operatorname{mig}(\boldsymbol{A} \pm \boldsymbol{B}) \geq \operatorname{mig}(\boldsymbol{A}) - \operatorname{mag}(\boldsymbol{B})$

Table B.5: Further properties of interval matrices ($\boldsymbol{A} \in \mathbb{IR}^{m \times n}$).

$\boldsymbol{A} = A^c + A^\Delta [-1, 1]$
$\boldsymbol{A} - \boldsymbol{A} = 2A^\Delta [-1, 1]$
$\boldsymbol{A} - \boldsymbol{A} = 0 \Leftrightarrow A^\Delta = 0$
$\boldsymbol{A} + \boldsymbol{A} = 2\boldsymbol{A}$

Table B.6: Interval matrix product ($\boldsymbol{A}, \boldsymbol{B}$ are interval matrices of suitable sizes).

$$\boldsymbol{AB} \supsetneq \{AB;\ A \in \boldsymbol{A},\ B \in \boldsymbol{B}\}$$
$$\boldsymbol{AB} = \square\{AB;\ A \in \boldsymbol{A},\ B \in \boldsymbol{B}\}$$

$\boldsymbol{A}, \boldsymbol{B} \geq 0$	$\Rightarrow$	$\boldsymbol{AB} = [\underline{A}\underline{B}, \overline{A}\overline{B}]$
$\boldsymbol{A} \geq 0,\ 0 \in \boldsymbol{B}$	$\Rightarrow$	$\boldsymbol{AB} = [\overline{A}\underline{B}, \overline{A}\overline{B}]$
$\boldsymbol{A} \geq 0,\ \boldsymbol{B} \leq 0$	$\Rightarrow$	$\boldsymbol{AB} = [\overline{A}\underline{B}, \underline{A}\overline{B}]$
$A \geq 0$ and real	$\Rightarrow$	$A\boldsymbol{B} = [A\underline{B}, A\overline{B}]$
$\underline{A} = 0$	$\Rightarrow$	$\boldsymbol{AB} = [\overline{A}\min\{\underline{B}, 0\}, \overline{A}\max\{\overline{B}, 0\}]$
$B^c = 0$	$\Rightarrow$	$\boldsymbol{AB} = \operatorname{mag}(\boldsymbol{A})\boldsymbol{B}$
		$= [-\operatorname{mag}(\boldsymbol{A})\operatorname{mag}(\boldsymbol{B}), \operatorname{mag}(\boldsymbol{A})\operatorname{mag}(\boldsymbol{B})]$

Table B.7: Interval matrix-vector and scalar multiplication ($a \in \mathbb{R}$, $x \in \mathbb{R}^n$, $A \in \mathbb{R}^{m\times n}$, $\boldsymbol{a} \in \mathbb{IR}$, $\boldsymbol{x} \in \mathbb{IR}^n$ and $\boldsymbol{A} \in \mathbb{IR}^{m\times n}$).

$$\boldsymbol{A}x = \{Ax;\ A \in \boldsymbol{A}\}$$
$$a\boldsymbol{A} = \{aA;\ A \in \boldsymbol{A}\}$$

$$\boldsymbol{a}A \supseteq \{aA;\ a \in \boldsymbol{a}\}$$
$$A\boldsymbol{x} \supseteq \{Ax;\ x \in \boldsymbol{x}\}$$

$$\boldsymbol{a}(\boldsymbol{A}\boldsymbol{x}) \subsetneq (\boldsymbol{a}\boldsymbol{A})\boldsymbol{x} = \boldsymbol{A}(\boldsymbol{a}\boldsymbol{x})$$

Table B.8: Comparison matrix ($\boldsymbol{A}, \boldsymbol{B}$ are interval matrices of suitable sizes).

$$\langle \boldsymbol{A} \rangle \geq \langle A^c \rangle - A^\Delta$$
$$\langle \boldsymbol{A} \rangle = \langle A^c \rangle - A^\Delta \ \Leftrightarrow\ 0 \notin \boldsymbol{a}_{ii}\ \forall i$$
$$\langle \boldsymbol{A} \rangle \geq \langle \boldsymbol{B} \rangle - \operatorname{dist}(\boldsymbol{A}, \boldsymbol{B})$$
$$\langle \boldsymbol{A} \rangle \leq \langle \boldsymbol{B} \rangle + \operatorname{dist}(\boldsymbol{A}, \boldsymbol{B})$$
$$\langle \boldsymbol{A} \rangle \leq \langle \boldsymbol{B} \rangle - \operatorname{mag}(\boldsymbol{A} \pm \boldsymbol{B})$$
$$\langle \boldsymbol{A} \pm \boldsymbol{B} \rangle \leq \langle \boldsymbol{A} \rangle + \operatorname{mag}(\boldsymbol{B})$$
$$\langle \boldsymbol{A} \pm \boldsymbol{B} \rangle \geq \langle \boldsymbol{A} \rangle - \operatorname{mag}(\boldsymbol{B})$$

$$\boldsymbol{A} \subseteq \boldsymbol{B} \ \Rightarrow\ \langle \boldsymbol{A} \rangle \geq \langle \boldsymbol{B} \rangle$$

Table B.9: inverse nonnegative matrix, M-matrix, H-matrix ($\boldsymbol{A}, \boldsymbol{B}$ are interval matrices of suitable sizes).

$\boldsymbol{A}$ is inverse nonnegative	$\Leftrightarrow$	$\underline{A}$ and $\overline{A}$ are inverse nonnegative
$\boldsymbol{A}$ is inverse nonnegative	$\Rightarrow$	$\boldsymbol{A}$ is regular $\wedge\ 0 \leq \overline{A}^{-1} \leq A^{-1} \leq \underline{A}^{-1}\ \forall A \in \boldsymbol{A}$
$\boldsymbol{A}$ is an M-matrix	$\Leftrightarrow$	$\underline{A}$ and $\overline{A}$ are M-matrices
$\boldsymbol{A}$ is an M-matrix	$\Rightarrow$	$\boldsymbol{A}$ is regular $\wedge\ 0 \leq \overline{A}^{-1} \leq A^{-1} \leq \underline{A}^{-1}\ \forall A \in \boldsymbol{A}$
$\boldsymbol{A}$ is an M-matrix	$\Rightarrow$	$\lvert A^{-1}\rvert \leq \lvert\underline{A}^{-1}\rvert = \langle \boldsymbol{A} \rangle^{-1} = \langle \underline{A} \rangle^{-1} = \underline{A}^{-1}\ \forall A \in \boldsymbol{A}$
$\boldsymbol{A}$ is an M-matrix $\wedge\ \boldsymbol{B} \subseteq \boldsymbol{A}$	$\Rightarrow$	$\boldsymbol{B}$ is an M-matrix
$\boldsymbol{A}$ is an H-matrix	$\Leftrightarrow$	$\langle \boldsymbol{A} \rangle$ is an M-matrix
$\boldsymbol{A}$ is an H-matrix	$\Leftrightarrow$	A^c is an H-matrix $\wedge\ \rho(\langle A^c \rangle^{-1} A^\Delta) < 1$
$\boldsymbol{A}$ is an H-matrix	$\Rightarrow$	$\boldsymbol{A}$ is regular $\wedge\ \lvert A^{-1}\rvert \leq \langle \boldsymbol{A} \rangle^{-1}\ \forall A \in \boldsymbol{A}$
$\boldsymbol{A}$ is an H-matrix $\wedge\ \boldsymbol{B} \subseteq \boldsymbol{A}$	$\Rightarrow$	$\boldsymbol{B}$ is an H-matrix

Table B.10: Interval matrix distance ($\boldsymbol{a} \in \mathbb{R}$ and $\boldsymbol{A}, \boldsymbol{A'}, \boldsymbol{B}, \boldsymbol{B'}, \boldsymbol{C}, \ldots$ are interval matrices of suitable sizes).

definition	$\begin{aligned}\operatorname{dist}(\boldsymbol{A},\boldsymbol{B}) &= \max\{|\underline{A}-\underline{B}|, |\overline{A}-\overline{B}|\} \\ &= |A^c - B^c| + |A^\Delta - B^\Delta| \\ &= \min\{Q \geq 0;\ \boldsymbol{A} \subseteq \boldsymbol{B} + [-Q,Q],\ \boldsymbol{B} \subseteq \boldsymbol{A} + [-Q,Q]\}\end{aligned}$
metric-like axioms	$\operatorname{dist}(\boldsymbol{A},\boldsymbol{B}) \geq 0$ $\operatorname{dist}(\boldsymbol{A},\boldsymbol{B}) = 0 \ \Leftrightarrow\ \boldsymbol{A} = \boldsymbol{B}$ $\operatorname{dist}(\boldsymbol{A},\boldsymbol{B}) = \operatorname{dist}(\boldsymbol{B},\boldsymbol{A})$ $\operatorname{dist}(\boldsymbol{A},\boldsymbol{B}) \leq \operatorname{dist}(\boldsymbol{A},\boldsymbol{C}) + \operatorname{dist}(\boldsymbol{C},\boldsymbol{B})$ $\|\operatorname{dist}(\boldsymbol{A},\boldsymbol{B})\|$ is a metric for any matrix norm
basic properties	$\begin{aligned}\operatorname{mag}(\boldsymbol{A}) &= \operatorname{dist}(\boldsymbol{A},0) \\ &= \operatorname{dist}(\boldsymbol{A}+\boldsymbol{B},\boldsymbol{B})\end{aligned}$ $\operatorname{dist}(\boldsymbol{A},\boldsymbol{B}) = \operatorname{dist}(\boldsymbol{A}+\boldsymbol{C},\boldsymbol{B}+\boldsymbol{C})$ $\operatorname{dist}(\boldsymbol{A}+\boldsymbol{A'},\boldsymbol{B}+\boldsymbol{B'}) \leq \operatorname{dist}(\boldsymbol{A},\boldsymbol{B}) + \operatorname{dist}(\boldsymbol{A'},\boldsymbol{B'})$ $\operatorname{dist}(\boldsymbol{CA},\boldsymbol{CB}) \leq \operatorname{mag}(\boldsymbol{C}) \operatorname{dist}(\boldsymbol{A},\boldsymbol{B})$ $\operatorname{dist}(\boldsymbol{AC},\boldsymbol{BC}) \leq \operatorname{dist}(\boldsymbol{A},\boldsymbol{B}) \operatorname{mag}(\boldsymbol{C})$ $\operatorname{dist}(\boldsymbol{A},\boldsymbol{B}) \geq |A^\Delta - B^\Delta|$ $\operatorname{dist}(\boldsymbol{A},\boldsymbol{B}) \geq |\operatorname{mag}(\boldsymbol{A}) - \operatorname{mag}(\boldsymbol{B})|$ $\operatorname{dist}(\boldsymbol{A},\boldsymbol{B}) \geq \operatorname{mag}(\boldsymbol{A}-\boldsymbol{B}) - 2A^\Delta$ $\boldsymbol{B} \subseteq \boldsymbol{A} + \operatorname{dist}(\boldsymbol{A},\boldsymbol{B})[-I, I]$
further properties	$\begin{aligned} B \text{ real} &\ \Rightarrow\ \operatorname{dist}(\boldsymbol{A},B) = \operatorname{mag}(\boldsymbol{A}-B) \\ a \text{ real} &\ \Rightarrow\ \operatorname{dist}(a\boldsymbol{A},a\boldsymbol{B}) = |a| \operatorname{dist}(\boldsymbol{A},\boldsymbol{B}) \\ \boldsymbol{A} \subseteq \boldsymbol{B} &\ \Rightarrow\ \operatorname{dist}(\boldsymbol{A},\boldsymbol{B}) = \operatorname{mag}(\boldsymbol{B} - A^c) - A^\Delta \\ \boldsymbol{A} \subseteq \boldsymbol{B} &\ \Rightarrow\ \operatorname{dist}(\boldsymbol{A},\boldsymbol{B}) \leq 2(B^\Delta - A^\Delta) \\ \boldsymbol{A} \subseteq \boldsymbol{B} &\ \Rightarrow\ \operatorname{dist}(\boldsymbol{A},\boldsymbol{B}) \geq B^\Delta - A^\Delta \\ \boldsymbol{A} \subseteq \boldsymbol{B} \subseteq \boldsymbol{C} &\ \Rightarrow\ \operatorname{dist}(\boldsymbol{A},\boldsymbol{B}) \leq \operatorname{dist}(\boldsymbol{A},\boldsymbol{C}) \\ \boldsymbol{A} \subseteq \boldsymbol{B} \subseteq \boldsymbol{C} &\ \Rightarrow\ \operatorname{dist}(\boldsymbol{B},\boldsymbol{C}) \leq \operatorname{dist}(\boldsymbol{A},\boldsymbol{C}) \end{aligned}$
convergence	$\begin{aligned}\lim_{k\to\infty} \boldsymbol{A}^{(k)} = \boldsymbol{B} &\ \Leftrightarrow\ \big(\lim_{k\to\infty} \underline{A}^{(k)} = \underline{B} \ \wedge\ \lim_{k\to\infty} \overline{A}^{(k)} = \overline{B}\big) \\ &\ \Leftrightarrow\ \lim_{k\to\infty} \boldsymbol{A}_{ij}^{(k)} = \boldsymbol{B}_{ij}\ \forall i,j \\ \boldsymbol{A}^{(1)} \supseteq \boldsymbol{A}^{(2)} \supseteq \ldots &\ \Rightarrow\ \lim_{k\to\infty} \boldsymbol{A}^{(k)} = \cap_{k=1}^{\infty} \boldsymbol{A}^{(k)}\end{aligned}$ $\operatorname{mid}\big(\lim_{k\to\infty} \boldsymbol{A}^{(k)}\big) = \lim_{k\to\infty} \operatorname{mid}(\boldsymbol{A}^{(k)})$ $\operatorname{rad}\big(\lim_{k\to\infty} \boldsymbol{A}^{(k)}\big) = \lim_{k\to\infty} \operatorname{rad}(\boldsymbol{A}^{(k)})$ arithmetic operations $+, -, \cdot$ are continuous operations $\operatorname{mid}, \operatorname{rad}, \operatorname{inf}, \operatorname{sup}$ and mag are continuous the interval scalar product is continuous

Bibliography

C. S. Adjiman, S. Dallwig, C. A. Floudas, and A. Neumaier. A global optimization method, αBB, for general twice-differentiable constrained NLPs – I. Theoretical advances. *Comput. Chem. Eng.*, 22(9): 1137–1158, 1998.

H.-S. Ahn and Y. Chen. Necessary and sufficient stability condition of fractional-order interval linear systems. *Automatica*, 44(11):2985–2988, 2008. 109

H.-S. Ahn and Y. Q. Chen. Exact maximum singular value calculation of an interval matrix. *IEEE Trans. Autom. Control*, 52(3):510–514, 2007. 109

H.-S. Ahn, K. L. Moore, and Y. Chen. Stability analysis of discrete-time iterative learning control systems with interval uncertainty. *Automatica*, 43(5):892–902, 2007. 109

I. G. Akrotirianakis and C. A. Floudas. A new class of improved convex underestimators for twice continuously differentiable constrained NLPs. *J. Glob. Optim.*, 30(4):367–390, 2004. 167

F. A. Al-Khayyal and J. E. Falk. Jointly constrained biconvex programming. *Math. Oper. Res.*, 8(2): 273–286, 1983. 166

G. Alefeld. *Intervallrechnung über den komplexen Zahlen und einige Anwendungen.* PhD thesis, Universität Karlsruhe, Karlsruhe, 1968. in German. 135

G. Alefeld. Über die Durchführbarkeit des Gaußschen Algorithmus bei Gleichungen mit Intervallen als Koeffizienten. *Comput.*, Suppl. 1:15–19, 1977. in German. 47, 48

G. Alefeld. Bounding the slope of polynomial operators and some applications. *Comput.*, 26(3):227–237, 1981. 125

G. Alefeld and N. Apostolatos. Praktische Anwendung von Abschätzungsformeln bei Iterationsverfahren. *ZAMM, Z. Angew. Math. Mech.*, 48(8):T46–T49, 1968. in German. 55

G. Alefeld and J. Herzberger. *Introduction to Interval Computations.* Computer Science and Applied Mathematics. Academic Press, New York, 1983. 46, 57, 118, 121, 125, 136, 138, 150, 152, 193

G. Alefeld and G. Mayer. The Cholesky method for interval data. *Linear Algebra Appl.*, 194:161–182, 1993. 81

G. Alefeld and G. Mayer. New criteria for the feasibility of the Cholesky method with interval data. *SIAM J. Matrix Anal. Appl.*, 30(4):1392–1405, 2008. 81

G. Alefeld, V. Kreinovich, and G. Mayer. On the shape of the symmetric, persymmetric, and skew-symmetric solution set. *SIAM J. Matrix Anal. Appl.*, 18(3):693–705, 1997. 74

G. Alefeld, V. Kreinovich, and G. Mayer. On the solution sets of particular classes of linear interval systems. *J. Comput. Appl. Math.*, 152(1-2):1–15, 2003. 74

G. Alefeld, A. Frommer, G. Heindl, and J. Mayer. On the existence theorems of Kantorovich, Miranda and Borsuk. *Electron. Trans. Numer. Anal.*, 17:102–111, 2004. 153

G. E. Alefeld, F. A. Potra, and W. Völker. Modifications of the Interval-Newton-Method with improved asymptotic efficiency. *BIT Numer. Math.*, 38(4):619–635, 1998. 150

I. P. Androulakis, C. D. Maranas, and C. A. Floudas. αBB: A global optimization method for general constrained nonconvex problems. *J. Glob. Optim.*, 7(4):337–363, 1995. 167

N. Apostolatos and U. Kulisch. Grundzüge einer Intervallrechnung für Matrizen und einige Anwendungen.

Elektron. Rechenanlagen, 10:73–83, 1968. in German. 34

I. Araya, G. Trombettoni, and B. Neveu. A contractor based on convex interval Taylor. In N. Beldiceanu, N. Jussien, and É. Pinson, editors, *Integration of AI and OR Techniques in Contraint Programming for Combinatorial Optimzation Problems*, volume 7298 of *LNCS*, pages 1–16. Springer, 2012. 155, 180

D. Avis and H. R. Tiwary. On the extension complexity of combinatorial polytopes. *Math. Program.*, 153(1):95–115, 2015. 100

A. B. Babichev, O. B. Kadyrova, T. P. Kashevarova, A. S. Leshchenko, and A. L. Semenov. UniCalc, a novel approach to solving systems of algebraic equations. *Interval Comput.*, 2:29–47, 1993. 139

A. Baharev, T. Achterberg, and E. Rév. Computation of an extractive distillation column with affine arithmetic. *AIChE J.*, 55(7):1695–1704, 2009. 33

I. Bar-On. Checking nonsingularity of tridiagonal matrices. *Electron. J. Linear Algebra*, 6:11–19, 2000. 70

I. Bar-On, B. Codenotti, and M. Leoncini. Checking robust nonsingularity of tridiagonal matrices in linear time. *BIT Numer. Math.*, 36(2):206–220, 1996. 70

W. Barth and E. Nuding. Optimale Lösung von Intervallgleichungssystemen. *Comput.*, 12:117–125, 1974. in German. 41, 44, 45, 49, 52

E. Baumann. Optimal centered forms. *BIT Numer. Math.*, 28(1):80–87, 1988. 120, 121

M. S. Bazaraa, H. D. Sherali, and C. M. Shetty. *Nonlinear Programming. Theory and Algorithms*. Wiley, NJ, 3rd edition, 2006. 15, 110, 160

H. Beeck. Über Struktur und Abschätzungen der Lösungsmenge von linearen Gleichungssystemen mit Intervallkoeffizienten. *Comput.*, 10(3):231–244, 1972. in German. 39, 63

H. Beeck. Charakterisierung der Lösungsmenge von Intervallgleichungssystemen. *Z. Angew. Math. Mech.*, 53:T181–T182, 1973. Sonderheft, in German. 41

H. Beeck. Zur scharfen Aussenabschätzung der Lösungsmenge bei linearen Intervallgleichungssystemen. *ZAMM, Z. Angew. Math. Mech.*, 54:T208–T209, 1974. in German. 39, 49

H. Beeck. Zur Problematik der Hüllenbestimmung von Intervallgleichungssystemen. In K. Nickel, editor, *Interval Mathemantics: Proceedings of the International Symposium on Interval Mathemantics*, volume 29 of *LNCS*, pages 150–159, Berlin, 1975. Springer. in German. 67

H. Beeck. Linear programming with inexact data. technical report TUM-ISU-7830, Technical University of Munich, Munich, 1978. 169, 170, 171, 176

G. Belforte, B. Bona, and V. Cerone. Parameter estimation algorithms for a set-membership description of uncertainty. *Automatica*, 26(5):887–898, 1990. 139

A. Ben-Tal, A. Goryashko, E. Guslitzer, and A. Nemirovski. Adjustable robust solutions of uncertain linear programs. *Math. Program.*, 99(2 (A)):351–376, 2004.

A. Ben-Tal, L. El Ghaoui, and A. Nemirovski. *Robust Optimization*. Princeton University Press, 2009. 180

F. Benhamou and L. Granvilliers. Continuous and interval constraints. In F. Rossi, P. van Beek, and T. Walsh, editors, *Handbook of Constraint Programming*, chapter 16, pages 571–603. Elsevier, 2006. 153

F. Benhamou, D. McAllester, and P. van Hentenryck. CLP(intervals) revisited. In *Proc. Int. Symp. Logic Prog.*, ILPS '94, pages 124–138, Cambridge, MA, USA, 1994. MIT Press. 158

A. H. Bentbib. Solving the full rank interval least squares problem. *Appl. Numer. Math.*, 41(2):283–294, 2002. 72

D. Berleant. Automatically verified arithmetic on probability distributions and intervals. In R. B. Kearfott and V. Kreinovich, editors, *Applications of Interval Computations*, chapter 6, pages 133–143. Kluwer, 1996. 186

D. Bertsimas, D. B. Brown, and C. Caramanis. Theory and applications of robust optimization. *SIAM*

Rev., 53(3):464–501, 2011. 180

S. Białas and J. Garloff. Intervals of P-matrices and related matrices. *Linear Algebra Appl.*, 58:33–41, 1984. 115

G. R. Bitran. Linear multiple objective problems with interval coefficients. *Manage. Sci.*, 26:694–706, 1980.

C. Bliek. *Computer Methods for Design Automation*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, July 1992. 57

R. Boche. Complex interval arithmetic with some applications. Technical Report LMSC4-22-66-1, Lockheed Missiles and Space Co., 1966. 135

J. Bok and M. Hladík. Selection-based approach to cooperative interval games. In B. Vitoriano and G. H. Parlier, editors, *Proceedings of the International Conference on Operations Research and Enterprise Systems*, pages 34–41, Lisbon, Portugal, 2015. SciTePress. ISBN 978-989-758-075-8. doi: 10.5220/0005221600340041. 179

S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004. 110

M. Bridger. *Real Analysis: A Constructive Approach Through Interval Arithmetic*, volume 38 of *Pure and Applied Undergraduate Texts*. AMS, Providence, RI, 2019. 24

Y. Candau, T. Raïssi, N. Ramdani, and L. Ibos. Complex interval arithmetic using polar form. *Reliab. Comput.*, 12(1):1–20, 2006. 138

O. Caprani and K. Madsen. Iterative methods for interval inclusion of fixed points. *BIT Numer. Math.*, 18(1):42–51, 1978. 55

O. Caprani and K. Madsen. Mean value forms in interval analysis. *Comput.*, 25:147–154, 1980. 123, 128

G. T. Cargo and O. Shisha. The Bernstein form of a polynomial. *J. Res. Nat. Bur. Stand., Sect. B*, 70B(1):79–81, 1966. URL `https://pdfs.semanticscholar.org/0164/50139348de499f5b3cc1425f40019d1ae3a5.pdf`. 129

M. Černý. Goffin's algorithm for zonotopes. *Kybernetika*, 48(5):890–906, 2012. 139

M. Černý and M. Rada. Polynomial time construction of ellipsoidal approximations of zonotopes given by generator descriptions. In M. Agrawal, S. B. Cooper, and A. Li, editors, *Theory and Applications of Models of Computation*, volume 7287 of *LNCS*, pages 156–163. Springer, Berlin, 2012. 139

M. Černý, J. Antoch, and M. Hladík. On the possibilistic approach to linear regression models involving uncertain, indeterminate or interval data. *Inf. Sci.*, 244:26–47, 2013. ISSN 0020-0255. doi: 10.1016/j.ins.2013.04.035. URL `https://doi.org/10.1016/j.ins.2013.04.035`. 72, 139

G. Chabert and L. Jaulin. Contractor programming. *Artif. Intell.*, 173(11):1079–1100, 2009. 155

D. Chablat, P. Wenger, F. Majou, and J.-P. Merlet. An interval analysis based study for the design and the comparison of 3-dof parallel kinematic machines. *Int. J. Robot. Res.*, 23(6):615–624, 2004. 109

L. Chen, A. Miné, J. Wang, and P. Cousot. Interval polyhedra: An abstract domain to infer interval linear relationships. In J. Palsberg and Z. Su, editors, *Proceedings of the 16th International Static Analysis Symposium (SAS'09), Los Angeles, CA*, volume 5673 of *LNCS*, pages 309–325, Springer, 2009. Berlin. 86

L. Chen, A. Miné, J. Wang, and P. Cousot. An abstract domain to discover interval linear equalities. In G. Barthe and M. Hermenegildo, editors, *Verification, Model Checking, and Abstract Interpretation*, pages 112–128, Berlin, Heidelberg, 2010. Springer. 43

F. L. Chernous'ko. An optimal algorithm for finding the roots of an approximately computed function. *USSR Comput. Math. Math. Phys.*, 8(4):705–724, 1968. 150, 151

J. W. Chinneck and K. Ramadan. Linear programming with interval coefficients. *J. Oper. Res. Soc.*, 51 (2):209–220, 2000. 173

W. Chuba and W. Miller. Quadratic convergence in interval arithmetic, part I. *BIT Numer. Math.*, 12 (3):284–290, 1972. 121

H. Collavizza, F. Delobel, and M. Rueher. Comparing partial consistencies. *Reliab. Comput.*, 5(3): 213–228, 1999. 155

W. D. Collins and C. Hu. Studying interval valued matrix games with fuzzy logic. *Soft Comput.*, 12(2): 147–155, 2008.

J. L. D. Comba and J. Stolfi. Affine arithmetic and its applications to computer graphics. In *Proceedings of SIBGRAPI'93, Recife, Brazil*, pages 9–18, 1993.

R. W. Cottle, J.-S. Pang, and R. E. Stone. *The Linear Complementarity Problem*. SIAM, Philadelphia, PA, revised ed. of the 1992 original edition, 2009. 14

G. E. Coxson. The P-matrix problem is co-NP-complete. *Math. Program.*, 64(1):173–178, 1994. 14

G. E. Coxson. Computing exact bounds on elements of an inverse interval matrix is NP-hard. *Reliab. Comput.*, 5(2):137–142, 1999. 115

T. Csendes. New subinterval selection criteria for interval global optimization. *J. Glob. Optim.*, 19(3): 307–327, 2001. 164

T. Csendes and D. Ratz. Subdivision direction selection in interval methods for global optimization. *SIAM J. Numer Anal.*, 34(3):922–938, 1997. 163, 164, 165

T. Csendes, R. Klatte, and D. Ratz. A posteriori direction selection rules for interval optimization methods. *Cent. Eur. J. Oper. Res.*, 8(3):225–236, 2000. 164

D. Daney. Optimal measurement configurations for Gough platform calibration. In *Proceedings of 2002 IEEE International Conference on Robotics and Automation*, volume 1, pages 147–152, Washington, USA, 2002. 186

D. Daney. Kinematic calibration of the Gough platform. *Robot.*, 21(6):677–690, 2003. 186

G. B. Dantzig. *Linear Programming and Extensions*. Princeton University Press, Princeton, NJ, 1963. 18, 177

D. P. Davey and N. F. Stewart. Guaranteed error bounds for the initial value problem using polytope arithmetic. *BIT*, 16:257–268, 1976.

D. Degrauwe, G. Lombaert, and G. De Roeck. Improving interval analysis in finite element calculations by means of affine arithmetic. *Comput. Struct.*, 88(3-4):247–254, 2010.

A. Deif. *Sensitivity Analysis in Linear Systems*. Springer, Berlin, Heidelberg, 1986. 72

A. S. Deif. The interval eigenvalue problem. *Z. Angew. Math. Mech.*, 71(1):61–64, 1991a. 109

A. S. Deif. Singular values of an interval matrix. *Linear Algebra Appl.*, 151:125–133, 1991b. 109

O. Dessombz, F. Thouverez, J. P. Laîné, and L. Jézéquel. Analysis of mechanical systems using interval computations applied to finite element methods. *J. Sound Vib.*, 239(5):949–968, 2001. 79

M. G. Dias da Silva. *Symmetries and Distances: Two Intriguing Challenges in Mathematical Programming*. PhD thesis, Université Paris-Saclay, 2017. URL `https://pastel.archives-ouvertes.fr/tel-01511408`. 109

A. D. Dimarogonas. Interval analysis of vibrating systems. *J. Sound Vib.*, 183(4):739–749, 1995. 109

B. Djanybekov. Interval Householder method for complex linear systems. *Reliab. Comput.*, 12(1):35–43, 2006. 138

F. Domes and A. Neumaier. Constraint aggregation for rigorous global optimization. *Math. Program.*, 155(1):375–401, 2016. 100

F. Domes, T. Montanher, H. Schichl, and A. Neumaier. Rigorous global filtering methods with interval unions. In O. Kosheleva et al., editor, *Beyond Traditional Probabilistic Data Processing Techniques: Interval, Fuzzy etc. Methods and Their Applications*, pages 249–267. Springer, Cham, 2020. 139

A. Dreyer. Interval methods for analog circuits. technical report 97, Fraunhofer-Institut für Techno- und Wirtschaftsmathematik ITWM, Kaiserslautern, 2006. 79

A. Dreyer. Interval analysis of linear analog circuits. In *Proceedings of the 12th GAMM–IMACS Sympo-*

sium on Scientific Computing, Computer Artithmetic and Validated Numerics, SCAN 2006., page 14, Los Alamitos, CA, USA, 2007. IEEE Computer Society. 79

L. Dymova, P. Sevastjanov, and M. Pilarek. A method for solving systems of linear interval equations applied to the Leontief input-output model of economics. *Expert Syst. Appl.*, 40(1):222–230, 2013. 84

B. Einarsson, editor. *Accuracy and Reliability in Scientific Computing*. SIAM, 2005. 23

H. A. Fazzolari and P. A. V. Ferreira. Robust stability and pole placement. *J. Control. Autom. Electr. Syst.*, 32:1498–1509, 2021. 79, 111

M. Fiedler, J. Nedoma, J. Ramík, J. Rohn, and K. Zimmermann. *Linear Optimization Problems with Inexact Data*. Springer, New York, 2006. 9, 39, 40, 42, 44, 69, 83, 84, 86

L. H. de Figueiredo and J. Stolfi. Adaptive enumeration of implicit surfaces with affine arithmetic. *Comput. Graph. Forum*, 15(5):287–296, 1996.

L. H. de Figueiredo and J. Stolfi. Affine arithmetic: concepts and applications. *Numer. Algorithms*, 37 (1-4):147–158, 2004.

R. Firment. Vyhodnocení polynomu s intervalovými koeficienty (Evaluation of interval polynomials). Bachelor's thesis, Faculty of Mathematics and Physics, Charles University in Prague, 2017. URL `https://is.cuni.cz/webapps/zzp/detail/184704`. in Czech. 130

C. A. Floudas. *Deterministic Global Optimization. Theory, Methods and Applications*, volume 37 of *Nonconvex Optimization and its Applications*. Kluwer, Dordrecht, 2000. 161, 167

C. A. Floudas and P. M. Pardalos. *A Collection of Test Problems for Constrained Global Optimization Algorithms*, volume 455 of *LNCS*. Springer, New York, 1990. 167

C. A. Floudas and P. M. Pardalos, editors. *Encyclopedia of Optimization*. Springer, New York, 2nd edition, 2009. 147, 152, 163, 164, 167

C. A. Floudas, P. M. Pardalos, C. Adjiman, W. R. Esposito, Z. H. Gümüs, S. T. Harding, J. L. Klepeis, C. A. Meyer, and C. A. Schweiger. *Handbook of Test Problems in Local and Global Optimization*. Springer, New York, 1999. 167

E. Fogel and Y. F. Huang. On the value of information in system identification-bounded noise case. *Automatica*, 18(2):229–238, 1982. 139

P. Franek, M. Krčál, and H. Wagner. Solving equations and optimization problems with uncertainty. *J Appl. Comput. Topology*, 1:297–330, 2018. 153

G. Franzè, L. Carotenuto, and A. Balestrino. New inclusion criterion for the stability of interval matrices. *IEE Proceedings - Control Theory and Applications*, 153(4):478–482, 2006. 109

A. Frommer. Proving conjectures by use of interval arithmetic. In U. Kulisch et al., editor, *Perspectives on Enclosure Methods*, pages 1–13. Springer, Wien, 2001. 22

V. Gabrel, C. Murat, and N. Remli. Best and worst optimum for linear programs with interval right hand sides. In H. A. Le Thi et al., editor, *Modelling, Computation and Optimization in Information Systems and Management Sciences. Second International Conference MCO 2008, Metz, France. Proceedings*, Berlin, 2008. Springer. 171, 172

V. Gabrel, C. Murat, and N. Remli. Linear programming with interval right hand sides. *Int. Trans. Oper. Res.*, 17(3):397–408, 2010. 171, 172

A. A. Gaganov. Computational complexity of the range of the polynomial in several variables. Master's thesis, Leningrad University, Math. Department, 1981. in Russian. 117, 127

A. A. Gaganov. Computation complexity of the range of a polynomial in several variables. *Cybern.*, 21: 418–421, 1985. 117, 127

E. Garajová and M. Hladík. On the optimal solution set in interval linear programming. *Comput. Optim. Appl.*, 72(1):269–292, 2019. ISSN 1573-2894. doi: 10.1007/s10589-018-0029-8. URL `http://link.springer.com/article/10.1007/s10589-018-0029-8`. 87, 173, 174, 175

E. Garajová and M. Rada. Exact method for the worst optimal value of an interval transportation

problem. In S. Drobne et al., editor, *Proceedings of the 16th International Symposium on Operational Research SOR'21, Bled, Slovenia, September 22-24, 2021*, pages 545–550, Ljubljana, Slovenia, 2021. BISTISK d.o.o. ISBN 978-961-6165-57-0. 173

E. Garajová, M. Hladík, and M. Rada. On the properties of interval linear programs with a fixed coefficient matrix. In A. Sforza and C. Sterle, editors, *Optimization and Decision Science: Methodologies and Applications*, volume 217 of *Springer Proceedings in Mathematics & Statistics*, pages 393–401. Springer, Cham, 2017. ISBN 978-3-319-67308-0. doi: 10.1007/978-3-319-67308-0_40. URL `https://link.springer.com/chapter/10.1007/978-3-319-67308-0_40`. 87, 95

E. Garajová, M. Hladík, and M. Rada. Interval transportation problem: The best and the worst (feasible) scenario. In S. Kapounek and H. Vránová, editors, *38th International Conference on Mathematical Methods in Economics 2020 (MME 2020). Conference Proceedings*, pages 122–127. Mendel University in Brno, 2020. ISBN 978-80-7509-734-7. URL `https://mme2020.mendelu.cz/wcd/w-rek-mme/mme2020_conference_proceedings_final.pdf`. 173

I. Gargantini. Parallel Laguerre iterations: The complex case. *Numer. Math.*, 26:317–323, 1976. 138

I. Gargantini. Further applications of circular arithmetic: Schroeder-like algorithms with error bounds for finding zeros of polynomials. *SIAM J. Numer. Anal.*, 15:497–510, 1978. 138

I. Gargantini and P. Henrici. Circular arithmetic and the determination of polynomial zeros. *Numer. Math.*, 18(4):305–320, 1972. 136, 138

J. Garloff. Convergent bounds for the range of multivariate polynomials. In K. Nickel, editor, *Interval Mathematics 1985: Proceedings of the International Symposium*, volume 212 of *LNCS*, pages 37–56, Berlin, Heidelberg, 1986. Springer. URL `http://www.nsc.ru/interval/Library/ProcBooks/IntervalMath85/037-056.pdf`. 131

J. Garloff. Block methods for the solution of linear interval equations. *SIAM J. Matrix Anal. Appl.*, 11 (1):89–106, 1990. 65

J. Garloff. Interval Gaussian elimination with pivot tightening. *SIAM J. Matrix Anal. Appl.*, 30(4): 1761–1772, 2009. 49, 189

J. Garloff. Pivot tightening for direct methods for solving symmetric positive definite systems of linear interval equations. *Comput.*, 94(2-4):97–107, 2012. 81, 189

J. Garloff and T. Hamadneh. Convergence and inclusion isotonicity of the tensorial rational Bernstein form. In M. Nehmeier et al., editor, *Scientific Computing, Computer Arithmetic, and Validated Numerics: 16th International Symposium, SCAN 2014*, volume 9553 of *LNCS*, pages 171–179. Springer, 2016. 131

J. Garloff, E. Popova, and A. P. Smith. Solving linear systems with polynomial parameter dependency. Preprint 1/2009, Institute of Mathematics and Informatics, Bulgarian Academy of Sciences, Sofia, January 2009. 73, 79

B. Gärtner and J. Matoušek. *Approximation Algorithms and Semidefinite Programming*. Springer, Berlin, Heidelberg, 2012. 105

D. M. Gay. Interval least squares - a diagnostic tool. In R. E. Moore, editor, *Reliability in computing. The role of interval methods in scientific computing*, volume 19 of *Perspectives in Computing*, pages 183–205. Academic Press, Boston, 1988. 72

W. Gerlach. Zur Lösung linearer Ungleichungssysteme bei Störung der rechten Seite und der Koeffizientenmatrix. *Math. Operationsforsch. Stat., Ser. Optimization*, 12:41–43, 1981. in German. 85

R. Ghosh, S. Sen, and K. B. Datta. An improved method for determining the stability of interval matrices. *Int. J. Syst. Sci.*, 31(2):171–176, 2000. 109

F. Gioia and C. N. Lauro. Principal component analysis on interval data. *Comput. Stat.*, 21(2):343–363, 2006. 109

A. Goldsztejn. Comparison of the Hansen-Sengupta and the Frommer-Lang-Schnurr existence tests. *Comput.*, 79(1):53–60, 2007. 152

A. Goldsztejn and A. Neumaier. On the exponentiation of interval matrices. *Reliab. Comput.*, 20:53–72, 2014. 115

F. Goualard and C. Jermann. A reinforcement learning approach to interval constraint propagation. *Constraints*, 13(1):206–226, 2008. 152

T. Gowers, editor. *The Princeton Companion to Mathematics*. Princeton University Press, Princeton, NJ, 2008. 36

L. Granvilliers. On the combination of interval constraint solvers. *Reliab. Comput.*, 7(6):467–483, 2001. 156

A. Griewank and G. F. Corliss, editors. *Automatic Differentiation of Algorithms. Theory, Implementation, and Application. Proceedings of the first SIAM workshop, held in Beckenridge, CO, USA, January 6–8, 1991*. SIAM, Philadelphia, 1991. 132

D. Y. Grigoryev and N. N. Vorobjov. Solving systems of polynomial inequalities in subexponential time. *J. Symp. Comput.*, 5(1/2):37–64, 1988. 127

K. G. Guderley and C. L. Keller. A basic theorem in the computation of ellipsoidal error bounds. *Numer. Math.*, 19:218–229, 1972.

T. C. Hales. A proof of the kepler conjecture. *Ann. Math., Second Ser.*, 162(3):1065–1185, 2005. 23

E. Hansen. On the solution of linear algebraic equations with interval coefficients. *Linear Algebra Appl.*, 2(2):153–165, 1969a. 61

E. R. Hansen. Interval arithmetic in matrix computations, Part I. *J. Soc. Ind. Appl. Math., Ser. B, Numer. Anal.*, 2(2):308–320, 1965. 44

E. R. Hansen. On solving systems of equations using interval arithmetic. *Math. Comput.*, 22(102):374–384, 1968. 123

E. R. Hansen. *Topics in Interval Analysis*. Oxford University Press, 1969b. 128, 138

E. R. Hansen. Global optimization using interval analysis - the multi-dimensional case. *Numer. Math.*, 34:247–270, 1980. 167

E. R. Hansen. Bounding the solution of interval linear equations. *SIAM J. Numer. Anal.*, 29(5):1493–1503, 1992. 57

E. R. Hansen. Sharpness in interval computations. *Reliab. Comput.*, 3(1):17–29, 1997. 117

E. R. Hansen and S. Sengupta. Bounding solutions of systems of equations using interval analysis. *BIT Numer. Math.*, 21:203–211, 1981. 52, 152

E. R. Hansen and R. Smith. Interval arithmetic in matrix computations. II. *SIAM J. Numer Anal.*, 4 (1):1–9, 1967. 44

E. R. Hansen and G. W. Walster. *Global Optimization Using Interval Analysis*. Marcel Dekker, New York, 2nd edition, 2004. 52, 117, 123, 150, 158, 162, 167

R. J. Hanson. Interval arithmetic as a closed arithmetic system on a computer. Technical report, Jet Propulsion Laboratory, 1968. 31

F. Hao and J.-P. Merlet. Multi-criteria optimal design of parallel manipulators based on interval analysis. *Mech. Mach. Theory*, 40(2):157–171, 2005. 186

D. J. Harfiel. Concerning the solution set of $Ax = b$ where $P \leq A \leq Q$ and $p \leq b \leq q$. *Numer. Math.*, 35 (3):355–359, 1980. 63

J. Hass and R. Schlafly. Double bubbles minimize. *Ann. Math.*, 151(2):459–515, 2000. 23

M. Hauenschild. Arithmetiken für komplexe Kreise. *Comput.*, 13(3-4):299–312, 1974. in German. 136

L. T. He and C. Hu. Impacts of interval computing on stock market variability forecasting. *Comput. Econ.*, 33(3):263–276, 2009. 72

H. A. Helfgott. The ternary goldbach conjecture is true. preprint, `https://arxiv.org/abs/1312.7748`, 2014. URL `https://arxiv.org/abs/1312.7748`. 23

E. M. T. Hendrix and B. Gazdag-Tóth. *Introduction to Nonlinear and Global Optimization*, volume 37 of *Optimization and Its Applications*. Springer, New York, 2010. 167

P. Henrici. Circular arithmetic and the determination of polynomial zeros. In J. L. Morris, editor, *Conference on Applications of Numerical Analysis*, volume 228 of *Lect. Notes Math.*, pages 86–92. Springer, Dundee, 1971. 136

T. Henriksen and K. Madsen. Parallel algorithms for global optimization. *Interval Comput.*, 3(5):88–95, 1992. 167

D. Hertz. The extreme eigenvalues and stability of real symmetric interval matrices. *IEEE Trans. Autom. Control*, 37(4):532–535, 1992. 104

T. J. Hickey, M. H. van Emden, and H. Wu. A unified framework for interval constraints and interval arithmetic. In M. Maher and J.-F. Puget, editors, *Principles and Practice of Constraint Programming – CP'98*, volume 1520 of *LNCS*, pages 250–264. Springer, Berlin, Heidelberg, 1998. 156

T. J. Hickey, Z. Qiu, and M. H. van Emden. Interval constraint plotting for interactive visual exploration of implicitly defined relations. *Reliab. Comput.*, 6:81–92, 2000. 187

N. J. Higham. *Accuracy and Stability of Numerical Algorithms*. SIAM, Philadelphia, 1996. 114

J.-B. Hiriart-Urruty. When only global optimization matters. *J. Glob. Optim.*, 56(3):761–763, 2013. 161

M. Hladík. Tolerances in portfolio selection via interval linear programming. In P. Rehorova, K. Marsikova, and Z. Hubinka, editors, *CD-ROM Proceedings 26-th International Conference on Mathematical Methods in Economics MME08, Liberec, Czech Republic*, pages 185–191. Technical University Liberec, September 2008a.

M. Hladík. Description of symmetric and skew-symmetric solution set. *SIAM J. Matrix Anal. Appl.*, 30 (2):509–521, 2008b. 74, 79, 80

M. Hladík. On necessary efficient solutions in interval multiobjective linear programming. In C. H. Antunes, D. R. Insua, and L. C. Dias, editors, *CD-ROM Proceedings of the 25th Mini-EURO Conference Uncertainty and Robustness in Planning and Decision Making URPDM 2010, April 15-17, Coimbra, Portugal*, pages 1–10, 2010a.

M. Hladík. Interval valued bimatrix games. *Kybernetika*, 46(3):435–446, 2010b. 179

M. Hladík. Optimal value bounds in nonlinear programming with interval data. *TOP*, 19(1):93–106, 2011a. doi: 10.1007/s11750-009-0099-y. URL `https://doi.org/10.1007/s11750-009-0099-y`.

M. Hladík. Error bounds on the spectral radius of uncertain matrices. In T. Simos, editor, *Proceedings of the International Conference on Numerical Analysis and Applied Mathematics 2011 (ICNAAM-2011), G-Hotels, Halkidiki, Greece, 19-25 September*, volume 1389 of *AIP Conference Proceedings*, pages 882–885, Melville, New York, 2011b. American Institute of Physics (AIP). ISBN 978-0-7354-0954-5. doi: 10.1063/1.3636875. 107

M. Hladík. Interval linear programming: A survey. In Z. A. Mann, editor, *Linear Programming – New Frontiers in Theory and Applications*, chapter 2, pages 85–120. Nova Science Publishers, New York, 2012a. ISBN 978-1-61209-579-0. 169

M. Hladík. Complexity of necessary efficiency in interval linear programming and multiobjective linear programming. *Optim. Lett.*, 6(5):893–899, 2012b. doi: 10.1007/s11590-011-0315-1. URL `https://doi.org/10.1007/s11590-011-0315-1`. 19

M. Hladík. An interval linear programming contractor. In J. Ramík and D. Stavárek, editors, *Proceedings 30th International Conference on Mathematical Methods in Economics 2012, Karviná, Czech Republic*, pages 284–289 (Part I.). Silesian University in Opava, School of Business Administration in Karviná, September 2012c. ISBN 978-80-7248-779-0. 175, 177

M. Hladík. Enclosures for the solution set of parametric interval linear systems. *Int. J. Appl. Math. Comput. Sci.*, 22(3):561–574, 2012d. doi: 10.2478/v10006-012-0043-4. 74, 75, 77, 79

M. Hladík. Bounds on eigenvalues of real and complex interval matrices. *Appl. Math. Comput.*, 219(10): 5584–5591, 2013a. ISSN 0096-3003. doi: 10.1016/j.amc.2012.11.075. 105

M. Hladík. Weak and strong solvability of interval linear systems of equations and inequalities. *Linear Algebra Appl.*, 438(11):4156–4165, 2013b. ISSN 0024-3795. doi: 10.1016/j.laa.2013.02.012. 90, 92, 97

M. Hladík. How to determine basis stability in interval linear programming. *Optim. Lett.*, 8(1):375–389, 2014a. ISSN 1862-4472. doi: 10.1007/s11590-012-0589-y. URL `https://doi.org/10.1007/s11590-012-0589-y`. 175

M. Hladík. New operator and method for solving real preconditioned interval linear equations. *SIAM J. Numer. Anal.*, 52(1):194–206, 2014b. ISSN 0036-1429. doi: 10.1137/130914358. 60

M. Hladík. On approximation of the best case optimal value in interval linear programming. *Optim. Lett.*, 8(7):1985–1997, 2014c. ISSN 1862-4472. doi: 10.1007/s11590-013-0715-5. URL `https://doi.org/10.1007/s11590-013-0715-5`. 173

M. Hladík. AE solutions and AE solvability to general interval linear systems. *Linear Algebra Appl.*, 465 (0):221–238, 2015. ISSN 0024-3795. doi: 10.1016/j.laa.2014.09.030.

M. Hladík. An extension of the αBB-type underestimation to linear parametric Hessian matrices. *J. Glob. Optim.*, 64(2):217–231, 2016a. ISSN 0925-5001. doi: 10.1007/s10898-015-0304-5. URL `https://doi.org/10.1007/s10898-015-0304-5`. 166

M. Hladík. Optimal preconditioning for the interval parametric Gauss–Seidel method. In M. Nehmeier et al., editor, *Scientific Computing, Computer Arithmetic, and Validated Numerics: 16th International Symposium, SCAN 2014, Würzburg, Germany, September 21-26*, volume 9553 of *LNCS*, pages 116–125. Springer, 2016b. ISBN 978-3-319-31769-4. doi: 10.1007/978-3-319-31769-4_10. 76

M. Hladík. Transformations of interval linear systems of equations and inequalities. *Linear Multilinear Algebra*, 65(2):211–223, 2017. ISSN 0308-1087. doi: 10.1080/03081087.2016.1180339. 90, 91, 93

M. Hladík. Testing pseudoconvexity via interval computation. *J. Glob. Optim.*, 71(3):443–455, 2018a. ISSN 0925-5001. doi: 10.1007/s10898-017-0537-6. 111, 113, 114

M. Hladík. The worst case finite optimal value in interval linear programming. *Croat. Oper. Res. Rev.*, 9(2):245–254, 2018b. ISSN 1848-9931. doi: 10.17535/crorr.2018.0019. 173

M. Hladík. P-completeness of testing solutions of parametric interval linear systems. In M. Ceberio and V. Kreinovich, editors, *Decision Making under Constraints*, volume 276 of *Studies in Systems, Decision and Control*, pages 117–123. Springer, Cham, 2020a. ISBN 978-3-030-40813-8. doi: 10.1007/978-3-030-40814-5_14. URL `https://doi.org/10.1007/978-3-030-40814-5_14`. 76, 79

M. Hladík. Two approaches to inner estimations of the optimal solution set in interval linear programming. In S. Deb, editor, *Proceedings of the 2020 4th International Conference on Intelligent Systems, Metaheuristics & Swarm Intelligence*, ISMSI 2020, pages 99–104, New York, USA, 2020b. Association for Computing Machinery. ISBN 978-1-4503-7761-4. doi: 10.1145/3396474.3396479. URL `https://doi.org/10.1145/3396474.3396479`. 175

M. Hladík. Robust Slater's condition in an uncertain environment. In R. Hlavatý, editor, *39th International Conference on Mathematical Methods in Economics 2021. Conference Proceedings*, pages 161–166. Czech University of Life Sciences Prague, 2021. ISBN 978-80-213-3126-6. URL `https://mme2021.v2.czu.cz/dl/99363?lang=en`. 87

M. Hladík and J. Horáček. Interval linear programming techniques in constraint programming and global optimization. In M. Ceberio and V. Kreinovich, editors, *Constraint Programming and Decision Making*, volume 539 of *Studies in Computational Intelligence*, pages 47–59. Springer, 2014. ISBN 978-3-319-04279-4. doi: 10.1007/978-3-319-04280-0_6. URL `http://dx.doi.org/10.1007/978-3-319-04280-0_6`. 180

M. Hladík and I. Skalna. Relations between various methods for solving linear interval and parametric equations. *Linear Algebra Appl.*, 574:1–21, August 2019. ISSN 0024-3795. doi: 10.1016/j.laa.2019.03.019. URL `http://www.sciencedirect.com/science/article/pii/S002437951930120X`. 77, 79

M. Hladík, D. Daney, and E. Tsigaridas. Bounds on real eigenvalues and singular values of interval matrices. *SIAM J. Matrix Anal. Appl.*, 31(4):2116–2129, 2010. doi: 10.1137/090753991. 103, 109

M. Hladík, D. Daney, and E. P. Tsigaridas. An algorithm for addressing the real interval eigenvalue problem. *J. Comput. Appl. Math.*, 235(8):2715–2730, 2011a. 66, 186

M. Hladík, D. Daney, and E. P. Tsigaridas. A filtering method for the interval eigenvalue problem. *Appl. Math. Comput.*, 217(12):5236–5242, 2011b. 109

M. Hladík, D. Daney, and E. P. Tsigaridas. Characterizing and approximating eigenvalue sets of symmetric interval matrices. *Comput. Math. Appl.*, 62(8):3152–3163, 2011c. 104

M. Hladík, L. V. Kolev, and I. Skalna. Linear interval parametric approach to testing pseudoconvexity. *J. Glob. Optim.*, 79(2):351–368, 2021. ISSN 1573-2916. doi: 10.1007/s10898-020-00924-w. 111

W. Hofschuster and W. Krämer. C-XSC 2.0 – A C++ Library for Extended Scientific Computing. In R. Alt, A. Frommer, R. B. Kearfott, and W. Luther, editors, *Numerical Software with Result Verification*, volume 2991 of *LNCS*, pages 15–35. Springer, Berlin, Heidelberg, 2004. 36

L. Hogben, editor. *Handbook of Linear Algebra*. Chapman & Hall/CRC, 2007. 9

O. Holzmann, B. Lang, and H. Schütt. Newton's constant of gravitation and verified numerical quadrature. *Reliab. Comput.*, 2(3):229–239, 1996. 23

H. Hong and V. Stahl. Bernstein form is inclusion monotone. *Comput.*, 55(1):43–53, 1995. 129

R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1985. 9, 68, 103

C. Hu and L. T. He. An application of interval methods to stock market forecasting. *Reliab. Comput.*, 13(5):423–434, 2007. 72

C.-Y. Hu, T. Maekawa, N. M. Patrikalakis, and X. Ye. Robust interval algorithm for surface intersections. *Comput. Aided Des.*, 29(9):617–627, 1997. 129

J. Huang, C. Wang, and H. Li. EA solutions and EA solvability to general interval linear systems. *Linear Multilinear Algebra*, 69(15):2865–2881, 2021. 100

M. Hutchings, F. Morgan, M. Ritoré, and A. Ros. Proof of the double bubble conjecture. *Ann. Math.*, 155(2):459–489, 2002. 23

E. Hyvönen. Constraint reasoning based on interval arithmetic: The tolerance propagation approach. *Artif. Intell.*, 58:71–112, 1992. 139, 153

E. Hyvönen and S. De Pascale. Interval computations on the spreadsheet. In R. B. Kearfott and V. Kreinovich, editors, *Applications of Interval Computations*, chapter 8, pages 169–209. Kluwer, 1996. 21, 37

M. Ida. Portfolio selection problem with interval coefficients. *Appl. Math. Lett.*, 16(5):709–713, 2003.

M. Ida. Efficient solution generation for multiple objective linear programming based on extreme ray generation method. *Eur. J. Oper. Res.*, 160(1):242–251, 2005.

C. Jansson. Interval linear systems with symmetric matrices, skew-symmetric matrices and dependencies in the right hand side. *Comput.*, 46(3):265–274, 1991. 77, 78

C. Jansson. Calculation of exact bounds for the solution set of linear interval systems. *Linear Algebra Appl.*, 251:321–340, 1997. 61

C. Jansson. Convex-concave extensions. *BIT*, 40(2):291–313, 2000. 180

C. Jansson and J. Rohn. An algorithm for checking regularity of interval matrices. *SIAM J. Matrix Anal. Appl.*, 20(3):756–776, 1999. 66, 67

C. Jansson and S. M. Rump. Rigorous solution of linear programming problems with uncertain data. *Z. Oper. Res.*, 35(2):87–111, 1991. 175

C. Jansson, D. Chaykin, and C. Keil. Rigorous error bounds for the optimal value in semidefinite programming. *SIAM J. Numer. Anal.*, 46(1):180–200, 2007. 72

L. Jaulin and D. Henrion. Contracting optimally an interval matrix without loosing any positive semi-definite matrix is a tractable problem. *Reliab. Comput.*, 11(1):1–17, 2005. 110

L. Jaulin, M. Kieffer, O. Didrit, and É. Walter. *Applied Interval Analysis*. Springer, London, 2001. 153, 155

M. E. Jerrell. Applications of interval computations to regional economic input-output models. In R. B. Kearfott and V. Kreinovich, editors, *Applications of Interval Computations*, chapter 6, pages 133–143. Kluwer, 1996. 63, 115

M. E. Jerrell. Interval arithmetic for input-output models with inexact data. *Comput. Econ.*, 10(1): 89–100, 1997. 63

W. Kahan. Circumscribing an ellipsoid about the intersection of two ellipsoids. *Canad. Math. Bull.*, 11: 437–441, 1968a.

W. M. Kahan. A more complete interval arithmetic. Lecture notes prepared for a summer course at the University of Michigan, 1968b. 31, 133, 151

L. V. Kantorovich. On new approaches to computational methods and processing of observations. *Siberian Math. J.*, 3:701–709, 1962. in Russian. 33

K. Karlík. Metoda nejmenších čtverců při nepřesných datech (The least squares method with interval data). Master's thesis, Faculty of Mathematics and Physics, Charles University in Prague, 2008. in Czech. 72

E. Kaucher. Interval analysis in the extended interval space IR. *Comput. Suppl.*, 2:33–49, 1980. 133

R. B. Kearfott. Some tests of generalized bisection. *ACM Trans. Math. Softw.*, 13(3):197–220, 1987. 152

R. B. Kearfott. Preconditioners for the interval Gauss–Seidel method. *SIAM J. Numer. Anal.*, 27(3): 804–822, 1990. 45

R. B. Kearfott. *Rigorous Global Search: Continuous Problems*. Kluwer, Dordrecht, 1996a. 32, 45, 150, 167

R. B. Kearfott. Algorithm 763: INTERVAL_ARITHMETIC: A Fortran 90 module for an interval data type. *ACM Trans. Math. Softw.*, 22(4):385–392, 1996b. 37

R. B. Kearfott. Interval computations: Introduction, uses, and resources. *Euromath Bulletin*, 2:95–112, 1996c.

R. B. Kearfott. On proving existence of feasible points in equality constrained optimization problems. *Math. Program.*, 83:89–100, 1998. 162

R. B. Kearfott. On existence and uniqueness verification for non-smooth functions. *Reliab. Comput.*, 8 (4):267–282, 2002. 151

R. B. Kearfott. Interval computations, rigour and non-rigour in deterministic continuous global optimization. *Optim. Methods Softw.*, 26(2):259–279, 2011. 164, 167

R. B. Kearfott, C. Hu, and M. Novoa III. A review of preconditioners for the interval Gauss–Seidel method. *Interval Comput.*, 1991(1):59–85, 1991. 45

R. B. Kearfott, M. Dawande, K. Du, and C. Hu. Algorithm 737: INTLIB: A portable Fortran 77 interval standard-function library. *ACM Trans. Math. Softw.*, 20(4):447–459, 1994. 37

R. B. Kearfott, M. T. Nakao, A. Neumaier, S. M. Rump, S. P. Shary, and P. Van Hentenryck. Standardized notation in interval analysis. In *Proc. XIII Baikal International School-seminar Optimization methods and their applications. Vol. 4 Interval analysis*, pages 106–113, Irkutsk: Institute of Energy Systems SB RAS, 2002. 7

L. Kettner, K. Mehlhorn, S. Pion, S. Schirra, and C. Yap. Classroom examples of robustness problems in geometric computations. *Comput. Geom.*, 40(1):61–78, 2008. 187

L. G. Khachiyan. A polynomial algorithm for linear programming. *Dokl. Akad. Nauk SSSR*, 244(5): 1093–1096, 1979. 18

R. Klatte and C. Ullrich. Complex sector arithmetic. *Comput.*, 24(2-3):139–148, 1980. 138

R. Klatte, U. Kulisch, M. Neaga, D. Ratz, and C. Ullrich. *Pascal-XSC. Language Reference with Examples*. Springer, Heidelberg, 1991. 36

L. Kolev. A direct method for determining a p-solution of linear parametric systems. *J. Appl. Comput. Math.*, 5(1:294):1–5, 2016. 115

L. Kolev and S. Petrakieva. Assessing the stability of linear time-invariant continuous interval dynamic systems. *IEEE Trans. Autom. Control*, 50(3):393–397, 2005. 109

L. V. Kolev. *Interval Methods for Circuit Analysis*. World Scientific, Singapore, 1993. 79

L. V. Kolev. Solving linear systems whose elements are nonlinear functions of intervals. *Numer. Algorithms*, 37(1-4):199–212, 2004a. 73

L. V. Kolev. A method for outer interval solution of linear parametric systems. *Reliab. Comput.*, 10(3): 227–239, 2004b. 79

L. V. Kolev. Outer interval solution of the eigenvalue problem under general form parametric dependencies. *Reliab. Comput.*, 12(2):121–140, 2006a. 109

L. V. Kolev. Improvement of a direct method for outer solution of linear parametric systems. *Reliab. Comput.*, 12(3):193–202, 2006b. 79

J. Koníčková. Sufficient condition of basis stability of an interval linear programming problem. *ZAMM, Z. Angew. Math. Mech.*, 81(Suppl. 3):677–678, 2001. 176

O. Král. Methods for a computation of the optimal value range in interval linear programming. Master's thesis, Department of Applied Mathematics, Charles University, 2020. URL `https://is.cuni.cz/webapps/zzp/detail/`. in Czech. 173, 176

R. Krawczyk. Newton-Algorithmen zur Bestimmung von Nullstellen mit Fehlerschranken. *Comput.*, 4: 187–201, 1969. in German. 52, 151

R. Krawczyk. Fehlerabschätzung bei linearer Optimierung. In K. Nickel, editor, *Interval Mathemantics: Proceedings of the International Symposium, Karlsruhe, West Germany, May 20-24, 1975*, volume 29 of *LNCS*, pages 215–222. Springer, 1975. in German. 169, 176

R. Krawczyk. Zur äußeren und inneren Einschließung des Wertebereichs einer Funktion. Freiburger Intervall-Berichte 80/7, Albert-Ludwigs-Universität, 1980. in German. 121

R. Krawczyk and A. Neumaier. Interval slopes for rational functions and associated centered forms. *SIAM J. Numer. Anal.*, 22(3):604–616, 1985. 125, 126

R. Krawczyk and A. Neumaier. An improved interval Newton operator. *J. Math. Anal. Appl.*, 118: 194–207, 1986. 121

R. Krawczyk and K. Nickel. Die zentrische Form in der Intervallarithmetik, ihre quadratische Konvergenz und ihre Inklusionsisotonie. *Comput.*, 28:117–137, 1982. in German. 121

V. Kreinovich. Why intervals? A simple limit theorem that is similar to limit theorems from statistics. *Reliab. Comput.*, 1(1):33–40, 1995. 24

V. Kreinovich. Optimal finite characterization of linear problems with inexact data. *Reliab. Comput.*, 11 (6):479–489, 2005. 66, 109

V. Kreinovich. Why intervals? Why fuzzy numbers? Towards a new justification. In J. M. Mendel, T. Omori, and X. Ya, editors, *2007 IEEE Symposium on Foundations of Computational Intelligence*, pages 113–119, 2007. 24

V. Kreinovich. Use of modal interval analysis in early engineering design. technical report, 2008. URL `https://api.semanticscholar.org/CorpusID:14503700`. 84

V. Kreinovich and B. J. Kubica. From computing sets of optima, Pareto sets, and sets of Nash equilibria to general decision-related set computations. *J. Univers. Comput. Sci.*, 16(18):2657–2685, 2010. 186

V. Kreinovich and A. V. Lakeyev. Linear interval equations: Computing enclosures with bounded relative or absolute overestimation is NP-hard. *Reliab. Comput.*, 2(4):341–350, 1996. 44

V. Kreinovich and J. W. von Gudenberg. An optimality criterion for arithmetic of complex sets. *Geombinatorics*, 1(10):31–37, 2000. 138

V. Kreinovich and G. Xiang. Fast algorithms for computing statistics under interval uncertainty: An

overview. In V.-N. Huynh et al., editor, *Interval / Probabilistic Uncertainty and Non-classical Logics*, volume 46 of *Advances in Soft Computing*, pages 19–31. Springer, Berlin, 2008. 186

V. Kreinovich, S. Starks, and G. Mayer. On a theoretical justification of the choice of epsilon-inflation in PASCAL-XSC. *Reliab. Comput.*, 3(4):437–445, 1997. 186

V. Kreinovich, A. Lakeyev, J. Rohn, and P. Kahl. *Computational Complexity and Feasibility of Data Processing and Interval Computations*. Kluwer, Dordrecht, 1998. 44, 109, 113, 115

V. Kreinovich, L. Longpré, and J. J. Buckley. Are there easy-to-check necessary and sufficient conditions for straightforward interval computations to be exact? *Reliab. Comput.*, 9(5):349–358, 2003. 117, 189

N. Krier. *Komplexe Kreisarithmetik*. PhD thesis, Universität Karlsruhe, Karlsruhe, 1973. in German. 138

U. Kulisch. Advanced arithmetic for the digital computer – interval arithmetic revisited. In U. Kulisch et al., editor, *Perspectives on Enclosure Methods*, pages 15–70. Springer, Wien, 2001. 132

U. Kulisch. *Computer Arithmetic and Validity. Theory, Implementation, and Applications*. de Gruyter, Berlin, 2nd revised and extended edition, 2013. 37, 132

J. R. Kuttler. A fourth-order finite-difference approximation for the fixed membrane eigenproblem. *Math. Comput.*, 25(114):237–256, 1971. 63

K. K. Lai, S. Y. Wang, J. P. Xu, S. S. Zhu, and Y. Fang. A class of linear interval programming problems and its application to portfolio selection. *IEEE Trans. Fuzzy Syst.*, 10(6):698–704, 2002.

A. V. Lakeev and S. I. Noskov. On the solution set of a linear equation with the right-hand side and operator given by intervals. *Sib. Math. J.*, 35(5):957–966, 1994. 42, 83, 84

M. Lange. Verification methods for conic linear programming problems. *Nonlinear Theory Appl., IEICE*, 11(3):327–358, 2020. 176

S. E. Laveuve. Definition einer Kahan-Arithmetik und ihre Implementierung. In K. Nickel, editor, *Interval Mathematics*, volume 29 of *LNCS*, pages 236–245, 1975. 133

Y. Lebbah and O. Lhomme. Accelerating filtering techniques for numeric CSPs. *Artif. Intell.*, 139(1): 109–132, 2002. 155

Y. Lebbah, C. Michel, M. Rueher, D. Daney, and J.-P. Merlet. Efficient and safe global constraints for handling numerical constraint systems. *SIAM J. Numer. Anal.*, 42(5):2076–2097, 2005. 155

H. Leng and Z. He. Computation of bounds for eigenvalues of structures with interval parameters. *Appl. Math. Comput.*, 216(9):2734–2739, 2010. 109

M. Lerch, G. Tischler, J. W. von Gudenberg, W. Hofschuster, and W. Krämer. The interval library filib++ 2.0 – design, features and sample programs. Preprint 2001/4, Bergische Universität Wuppertal, 2001. URL `http://www2.math.uni-wuppertal.de/wrswt/preprints/prep_01_4.pdf`. 36

Q.-X. Li and S.-F. Liu. The foundation of the grey matrix and the grey input-output analysis. *Appl. Math. Model.*, 32(3):267–291, 2008. 84

W. Li. A note on dependency between interval linear systems. *Optim. Lett.*, 9(4):795–797, 2015. 87

Y. P. Li, G. H. Huang, P. Guo, Z. F. Yang, and S. L. Nie. A dual-interval vertex analysis method and its application to environmental decision making under uncertainty. *Eur. J. Oper. Res.*, 200(2):536–550, 2010.

S.-T. Liu and C. Kao. Matrix games with interval data. *Comput. Ind. Eng.*, 56(4):1697–1700, 2009. 179

R. J. Lohner. On the ubiquity of the wrapping effect in the computation of error bounds. In U. Kulisch et al., editor, *Perspectives on Enclosure Methods*, pages 201–216. Springer, Wien, 2001. 29

H. W. Lu, G. H. Huang, and L. He. Development of an interval-valued fuzzy linear-programming method based on infinite α-cuts for water resources management. *Environ. Model. Softw.*, 25(3):354–361, 2010.

D. Luenberger and Ye. *Linear and Nonlinear Programming*. Springer, New York, 3rd edition, 2008. 110

B. Machost. Numerische Behandlung des Simplexverfahrens mit intervallanalytischen Methoden. Techni-

cal Report 30, Berichte der Gesellschaft für Mathematik und Datenverarbeitung, 54 pages, Bonn, 1970. in German. 169, 170

K. Madsen. Parallel algorithms for global optimization. DTU report NI-91-07, Institute for Numerical Analysis, Lyngby, 1991. 167

K. Makino and M. Berz. Suppression of the wrapping effect by Taylor model-based verified integrators: Long-term stabilization by preconditioning. *Int. J. Differ. Equ. Appl.*, 10(4):105–136, 2005. 141

M. Mansour. Robust stability of interval matrices. In *Proceedings of the 28th IEEE Conference on Decision and Control*, volume 1, pages 46 –51, Tampa, Florida, 1989. 109

T. A. Manteuffel. An interval analysis approach to rank determination in linear least squares problems. *SIAM J. Sci. Stat. Comput.*, 2:335–348, 1981. 72

M. Marcus and H. Minc. *A Survey of Matrix Theory and Matrix Inequalities*. Allyn and Bacon, Boston, 1964. 13

S. M. Markov. Least squares approximations under interval input data. In C. Ullrich, editor, *Contributions to Computer Arithmetic and Self-Validating Numerical Methods*, volume 7 of *IMACS Annals on computing and applied mathematics*, pages 133–147. J.C. Baltzer Sci. Publ., Basel, 1990. 72, 121

J. A. Martínez, L. G. Casado, I. García, Y. D. Sergeyev, and B. TĂóth. On an efficient use of gradient information for accelerating interval global optimization algorithms. *Numer. Algorithms*, 37(1-4):61–69, 2004. 162

Y. V. Matiyasevich. Enumerable sets are diophantine. *Sov. Math., Dokl.*, 11(2):354–357, 1970. 143

J. Matějka. Determinanty intervalových matic) (Determinants of interval matrices). Bachelor's thesis, Faculty of Mathematics and Physics, Charles University in Prague, 2017. in Czech. 113

G. Mayer. A survey on properties and algorithms for the symmetric solution set. technical report 12/2, Institut für Mathematik, Universität Rostock, Rostock, Germany, 2012. 79

G. Mayer. *Interval Analysis and Automatic Result Verification*, volume 65 of *Studies in Mathematics*. De Gruyter, Berlin, 2017. 28, 40, 46, 47, 48, 49, 50, 52, 57, 63, 71, 74, 79, 80, 81, 118, 132, 136, 137, 138, 150, 151, 152, 185, 193

G. Mayer and J. Rohn. On the applicability of the interval Gaussian algorithm. *Reliab. Comput.*, 4(3): 205–222, 1998. 49

O. Mayer. Algebraische und metrische Strukturen in der Intervallrechnung und einige Anwendungen. *Comput.*, 5:144–162, 1970. in German. 111

G. P. McCormick. Computability of global solutions to factorable nonconvex programs: Part I – Convex underestimating problems. *Math. Program.*, 10(1):147–175, 1976. 166

J.-P. Merlet. Determination of the orientation workspace of parallel manipulators. *Int. J. Robot. Res.*, 13(2):143–160, 1995. 186

J.-P. Merlet. Determination of the presence of singularities in 6D workspace of a Gough parallel manipulator. In J. Lenarčič and M. L. Husty, editors, *Advances in Robot Kinematics: Analysis and Control*, pages 39–48. Springer, Dordrecht, 1998. 114, 186

J.-P. Merlet. Solving the forward kinematics of a Gough-type parallel manipulator with interval analysis. *Int. J. Robot. Res.*, 23(3):221–235, 2004. 186

J.-P. Merlet. Optimal design of robots. In *Proceedings of Robotics: Science and Systems*, pages 8–11, Cambridge, USA, 2005. 186

J.-P. Merlet. A formal-numerical approach for robust in-workspace singularity detection. *IEEE Trans. Robot.*, 23(3):393–402, 2007. 114, 186

J.-P. Merlet and P. Donelan. On the regularity of the inverse Jacobian of parallel robots. In J. Lennarčič and B. Roth, editors, *Advances in Robot Kinematics*, pages 41–48. Springer, 2006. 79, 114, 186

F. Messine and A. Touhami. A general reliable quadratic form: An extension of affine arithmetic. *Reliab. Comput.*, 12(3):171–192, 2006.

C. D. Meyer. *Matrix Analysis and Applied Linear Algebra*. SIAM, Philadelphia, 2000. 9, 68, 103

K. Meyer and D. Schmidt, editors. *Computer Aided Proofs in Analysis*. Springer, New York, 1991. 23

W. Miller. Quadratic convergence in interval arithmetic, part II. *BIT Numer. Math.*, 12(3):291–298, 1972. 121

M. Mohammadi and M. Gentili. Bounds on the worst optimal value in interval linear programming. *Soft Comput.*, 23(21):11055–11061, 2019. 173

M. Mohammadi and M. Gentili. The outcome range problem in interval linear programming. *Comput. Oper. Res.*, 129:105160:1–17, 2021. 178

T. Montanher, F. Domes, H. Schichl, and A. Neumaier. Using interval unions to solve linear systems of equations with uncertainties. *BIT Numer. Math.*, 57(3):901–926, 2017. 45, 139

R. E. Moore. *Interval Arithmetic and Automatic Error Analysis in Digital Computing*. PhD thesis, Stanford University, Stanford, California, USA, November 1962. Published as Applied Mathematics and Statistics Laboratories Technical Report No. 25. 33

R. E. Moore. *Interval Analysis*. Prentice-Hall, Englewood Cliffs, NJ, 1966. 31, 32, 33, 34, 119, 139, 151, 167

R. E. Moore. *Methods and Applications of Interval Analysis*. SIAM, Philadelphia, 1979. 131, 140, 141, 153, 163

R. E. Moore. The resolution of close minima. *Comput. Math. Appl.*, 25(10-11):57–58, 1993. 147

R. E. Moore and H. Ratschek. Inclusion functions and global optimization II. *Math. Program.*, 41(1): 341–356, 1988. 164, 165

R. E. Moore, R. B. Kearfott, and M. J. Cloud. *Introduction to Interval Analysis*. SIAM, Philadelphia, PA, 2009. 49, 152, 193

A. Mostafaee, M. Hladík, and M. Černý. Inverse linear programming with interval coefficients. *J. Comput. Appl. Math.*, 292:591–608, 2016. ISSN 0377-0427. doi: 10.1016/j.cam.2015.07.034. 171

S. P. Mudur and P. A. Koparkar. Interval methods for processing geometric objects. *IEEE Comput. Graph. Appl.*, 4(2):7–17, 1984. 187

K. G. Murty and F.-T. Yu. *Linear Complementarity, Linear and Nonlinear Programming*. Internet edition, 1997. URL `http://www-personal.umich.edu/~murty/books/linear_complementarity_webbook/lcp-complete.pdf`. 14

A. Narkawicz, J. Garloff, A. P. Smith, and C. A. Muñoz. Bounding the range of a rational function over a box. *Reliab. Comput.*, 17(1):34–39, 2012. URL `http://interval.louisiana.edu/reliable-computing-journal/volume-17/reliable-computing-17-pp-34-39.pdf`. 131

P. S. V. Nataraj and S. M. Sheela. A new subdivision strategy for range computations. *Reliab. Comput.*, 8(1):83–92, 2002. 164

P. K. Nayak and M. Pal. Linear programming technique to solve two person matrix games with interval pay-offs. *Asia Pac. J. Oper. Res.*, 26(2):285–305, 2009.

N. S. Nedialkov and K. R. Jackson. A new perspective on the wrapping effect in interval methods for initial value problems for ordinary differential equations. In U. Kulisch et al., editor, *Perspectives on Enclosure Methods*, pages 219–263. Springer, Wien, 2001. 141

A. Nemirovskii. Several NP-hard problems arising in robust stability analysis. *Math. Control Signals Syst.*, 6(2):99–105, 1993. 26, 109

Y. Nesterov and A. Nemirovskii. *Interior-Point Polynomial Algorithms in Convex Programming*. SIAM, Philadelphia, 1994. 105

A. Neumaier. New techniques for the analysis of linear interval equations. *Linear Algebra Appl.*, 58: 273–325, 1984. 25, 44, 45, 193

A. Neumaier. Linear interval equations. In K. Nickel, editor, *Interval Mathematics 1985*, volume 212 of *LNCS*, pages 109–120, New York, 1986. Springer. 72

A. Neumaier. Overestimation in linear interval equations. *SIAM J. Numer. Anal.*, 24:207–214, 1987. 71

A. Neumaier. Rigorous sensitivity analysis for parameter-dependent systems of equations. *J. Math. Anal. Appl.*, 144(1):16–25, 1989.

A. Neumaier. *Interval Methods for Systems of Equations*. Cambridge University Press, Cambridge, 1990. 9, 45, 46, 49, 52, 53, 54, 56, 57, 60, 62, 63, 64, 65, 67, 70, 71, 115, 121, 123, 125, 150, 151, 152, 189, 193

A. Neumaier. A simple derivation of the Hansen-Bliek-Rohn-Ning-Kearfott enclosure for linear interval equations. *Reliab. Comput.*, 5(2):131–136, 1999. 57

A. Neumaier. Complete search in continuous global optimization and constraint satisfaction. *Acta Numer.*, 13:271–369, 2004. 167

A. Neumaier. Computer-assisted proofs. In W. Luther and W. Otten, editors, *Proc. 12th GAMM-IMACS Symp. Sci. Comp., (SCAN 2006)*, pages 1–6. IEEE Computer Society, 2007. 22

A. Neumaier and A. Pownuk. Linear systems with large uncertainties, with applications to truss structures. *Reliab. Comput.*, 13(2):149–172, 2007. 79

A. Neumaier, O. Shcherbina, W. Huyer, and T. Vinkó. A comparison of complete global optimization solvers. *Math. Program.*, 103(2):335–356, 2005. 167

J. von Neumann and O. Morgenstern. *Theory of Games and Economic Behavior. With an introduction by Harold Kuhn and an afterword by Ariel Rubinstein*. Princeton University Press, Princeton, NJ, 2007. 178

K. Nickel. On the Newton method in interval analysis. MRC technical summary report 1136, Mathematics Research Center, University of Wisconsin, 1971. 149

K. Nickel, editor. *Interval Mathematics*, volume 29 of *LNCS*, Cham, 1975. Springer. 138

K. Nickel. Die Überschätzung des Wertebereichs einer Funktion in der Intervallrechnung mit Anwendungen auf lineare Gleichungssysteme. *Comput.*, 18(1):15–36, 1977. 63

K. Nickel. Arithmetic of complex sets. *Comput.*, 24:97–105, 1980. 135, 136, 138

S. Ning and R. B. Kearfott. A comparison of some methods for solving linear interval equations. *SIAM J. Numer. Anal.*, 34(4):1289–1305, 1997. 57

J. Nocedal and S. J. Wright. *Numerical Optimization*. Springer Series in Operations Research and Financial Engineering. Springer, New York, 2nd edition, 2006. 15, 160

A. S. Noghabi, H. R. Mashhadi, and J. Sadeh. Optimal coordination of directional overcurrent relays considering different network topologies using interval linear programming. *IEEE Trans. Power Deliv.*, 25(3):1348–1354, 2010.

J. Novotná, M. Hladík, and T. Masařík. Duality gap in interval linear programming. *J. Optim. Theory Appl.*, 184(2):565–580, 2020. ISSN 1573-2878. doi: 10.1007/s10957-019-01610-y. URL `https://doi.org/10.1007/s10957-019-01610-y`. 173

D. Oelschlägel and H. Süße. Interval analytic treatment of convex programming problems. *Comput.*, 24 (2):213–225, 1980. 176

D. Oetomo, D. Daney, B. Shirinzadeh, and J.-P. Merlet. An interval-based method for workspace analysis of planar flexure-jointed mechanism. *J Mech. Des.*, 131(1):0110141–01101411, 2009. 114, 160, 186

W. Oettli. On the solution set of a linear system with inaccurate coefficients. *J. SIAM Numer. Anal.*, 2 (1):115–118, 1965. 41

W. Oettli and W. Prager. Compatibility of approximate solution of linear equations with given error bounds for coefficients and right-hand sides. *Numer. Math.*, 6:405–409, 1964. 39

C. Oliveira and C. H. Antunes. Multiple objective linear programming models with interval coefficients – an illustrated overview. *Eur. J. Oper. Res.*, 181(3):1434–1463, 2007.

C. Oliveira and C. H. Antunes. An interactive method of tackling uncertainty in interval multiple objective linear programming. *J. Math. Sci.*, 161(6):854–866, 2009.

M. Padberg. *Linear Optimization and Extensions*. Springer, Berlin, 2nd edition, 1999. 15, 18

C. H. Papadimitriou. *Computational Complexity*. Addison Wesley Longman, 1994. 19

K. B. Petersen and M. S. Pedersen. *The Matrix Cookbook*. Technical University of Denmark, 2012. URL `http://www2.imm.dtu.dk/pubdb/pubs/3274-full.html`. Version 20121115. 14

D. B. Petkovski. Reduced conservatism in stability analysis of interval matrices. *J. Optim. Theory Appl.*, 70(3):597–606, 1991. 109

J. D. Pintér. *Global Optimization in Action*, volume 6 of *NOIA*. Kluwer, Dordrecht, 1996. 161

S. Poljak and J. Rohn. Radius of nonsingularity. Technical Report KAM Series (88-117), Department of Applied Mathematics, Charles University, Prague, 1988. 65

S. Poljak and J. Rohn. Checking robust nonsingularity is NP-hard. *Math. Control Signals Syst.*, 6(1): 1–9, 1993. 65

E. D. Popova. On the solution of parametrised linear systems. In W. Krämer and J. W. von Gudenberg, editors, *Scientific Computing, Validated Numerics, Interval Methods*, pages 127–138. Kluwer, 2001. 79

E. D. Popova. Parametric interval linear solver. *Numer. Algorithms*, 37(1-4):345–356, 2004. 79

E. D. Popova. Webcomputing service framework. *Int. J. Inf. Theor. Appl.*, 13(3):246–254, 2006a. 43, 79

E. D. Popova. Computer-assisted proofs in solving linear parametric problems. In *12th GAMM - IMACS International Symposium on Scientific Computi ng, Computer Arithmetic and Validated Numerics, SCAN 2006, Conference Post -Proceedings, September 26-29, Duisburg, Germany*, page 35. IEEE Computer Society Press, 2006b. 79

E. D. Popova. Improved solution enclosures for over- and underdetermined interval linear systems. In I. Lirkov, S. Margenov, and J. Waśniewski, editors, *Large-Scale Scientific Computing*, volume 3743 of *LNCS*, pages 305–312, 2006c. 72

E. D. Popova. Solving linear systems whose input data are rational functions of interval parameters. In T. Boyanov et al., editor, *Numerical Methods and Applications. NMA 2006*, volume 4310 of *LNCS*, pages 345–352, Berlin, Heidelberg, 2007. Springer. 73

E. D. Popova. Explicit characterization of a class of parametric solution sets. *Comptes Rendus de L'Academie Bulgare des Sciences*, 62(10):1207–1216, 2009. 74

E. D. Popova. Solvability of parametric interval linear systems of equations and inequalities. *SIAM J. Matrix Anal. Appl.*, 36(2):615–633, 2015. 74, 89, 90

E. D. Popova and W. Krämer. Inner and outer bounds for the solution set of parametric linear systems. *J. Comput. Appl. Math.*, 199(2):310–316, 2007. 79

L. Pronzato and E. Walter. Minimum-volume ellipsoids containing compact sets : Application to parameter bounding. *Automatica*, 30(11):1731–1739, 1994. 139

J. D. Pryce and G. F. Corliss. Interval arithmetic with containment sets. *Comput.*, 78(3):251–276, 2006.

Z. Qiu, S. Chen, and I. Elishakoff. Bounds of eigenvalues for structures with an interval description of uncertain-but-non-random parameters. *Chaos Soliton. Fract.*, 7(3):425–434, 1996. 109

Z. Qiu, P. C. Müller, and A. Frommer. An approximate method for the standard interval eigenvalue problem of real non-symmetric interval matrices. *Commun. Numer. Methods Eng.*, 17(4):239–251, 2001. 109

M. Rada, M. Hladík, and E. Garajová. Testing weak optimality of a given solution in interval linear programming revisited: NP-hardness proof, algorithm and some polynomially-solvable cases. *Optim. Lett.*, 13(4):875–890, June 2019. ISSN 1862-4480. doi: 10.1007/s11590-018-1289-z. URL `http://link.springer.com/article/10.1007/s11590-018-1289-z`. 173

L. B. Rall. *Automatic Differentiation: Techniques and Applications*, volume 120 of *LNCS*. Springer, New York, 1981. 132

S. Ratschan. Search heuristics for box decomposition methods. *J. Glob. Optim.*, 24(1):35–49, 2002. 164

S. Ratschan. Efficient solving of quantified inequality constraints over the real numbers. *ACM Trans. Comput. Logic*, 7(4):723–748, 2006. 155

S. Ratschan et al. RSolver. Software Package, 2004. URL `http://rsolver.sourceforge.net`. 155

H. Ratschek. Die Subdistributivität der Intervallarithmetik. *ZAMM, Z. Angew. Math. Mech.*, 51(3): 189–192, 1971. in German. 28

H. Ratschek. Inclusion functions and global optimization. *Math. Program.*, 33(3):300–317, 1985. 164, 165

H. Ratschek and J. Rokne. Nonuniform variable precision bisecting. In C. Brezinski and U. Kulisch, editors, *Computational and applied Mathematics I*, pages 419–428, Amsterdam, 1992. Elsevier. 164

H. Ratschek and J. Rokne. *Geometric Computations with Interval and New Robust Methods. Applications in Computer Graphics, GIS and Computational Geometry.* Horwood Publishing, Chichester, 2003. 114, 129, 151, 187

H. Ratschek and J. Rokne. *New Computer Methods for Global Optimization.* Wiley, Chichester, 2007. 167

H. Ratschek and J. G. Rokne. Formulas for the width of interval products. *Reliab. Comput.*, 1(1):9–14, 1995. 193

D. Ratz. Box-splitting strategies for the interval Gauss–Seidel step in a global optimization method. *Comput.*, 53(3-4):337–353, 1994. 163

D. Ratz. Inclusion isotone extended interval arithmetic. Technical Report 5/1996, Institut für Angewandte Mathematik, Karlsruhe, 1996. URL `http://digbib.ubka.uni-karlsruhe.de/volltexte/67997`. 31, 32

D. Ratz. *Automatic Slope Computation and Its Application in Nonsmooth Global Optimization.* Shaker, Aachen, 1998. professorial dissertation. 126

D. Ratz. Nonsmooth global optimization. In U. Kulisch et al., editor, *Perspectives on Enclosure Methods*, pages 277–337. Springer, Wien, 2001. 123, 125, 159, 163, 167

S. Ray and P. S. V. Nataraj. A matrix method for efficient computation of Bernstein coefficients. *Reliab. Comput.*, 17:40–71, 2012. URL `http://interval.louisiana.edu/reliable-computing-journal/volume-17/reliable-computing-17-pp-40-71.pdf`. 129

K. Reichmann. Abbruch beim Intervall-Gauß-Algorithmus. *Comput.*, 22(4):355–361, 1979. in German. 47, 81

G. Rex and J. Rohn. A note on checking regularity of interval matrices. *Linear Multilinear Algebra*, 39 (3):259–262, 1995. 68

G. Rex and J. Rohn. Sufficient conditions for regularity and singularity of interval matrices. *SIAM J. Matrix Anal. Appl.*, 20(2):437–445, 1998. 67, 70

F. N. Ris. *Interval Snalysis and Spplications to Linear Algebra.* D. phil. thesis, University of Oxford, 1972. 52

F. N. Ris. Tools for the analysis of interval arithmetic. In K. Nickel, editor, *Interval Mathematics*, volume 29 of *LNCS*, pages 75–98, Berlin, Heidelberg, 1975. Springer. 193

J. Rohn. Soustavy lineárních rovnic s intervalově zadanými koeficienty (Systems of linear equations with inexact data). *Ekonom.-Mat. Obzor*, 12:311–315, 1976. in Czech. 170

J. Rohn. Input-output planning with inexact data. Freiburger Intervall-Berichte 78/9, Albert-Ludwigs-Universität, Freiburg, 1978. 84

J. Rohn. Strong solvability of interval linear programming problems. *Comput.*, 26:79–82, 1981. 92

J. Rohn. Interval linear systems. Freiburger Intervall-Berichte 84/7, Albert-Ludwigs-Universität, Freiburg, 1984. 43, 170

J. Rohn. Miscellaneous results on linear interval systems. Freiburger Intervall-Berichte 85/9, Albert-Ludwigs-Universität, Freiburg, 1985. 40, 87, 113

J. Rohn. Inner solutions of linear interval systems. In K. Nickel, editor, *Interval mathematics 1985, Proceedings of the International Symposium*, volume 212 of *LNCS*, pages 157–158. Springer, Berlin, 1986. 83

J. Rohn. Inverse-positive interval matrices. *ZAMM, Z. Angew. Math. Mech.*, 67:T492–T493, 1987. 63

J. Rohn. Systems of linear interval equations. *Linear Algebra Appl.*, 126(C):39–78, 1989a. 62, 65, 66, 69, 115

J. Rohn. On sensitivity of the optimal value of a linear program. *Ekonom.-Mat. Obzor*, 25(1):105–107, 1989b. 178, 186

J. Rohn. Cheap and tight bounds: The recent result by E. Hansen can be made more efficient. *Interval Comput.*, 4:13–21, 1993a. 57, 115

J. Rohn. Interval matrices: Singularity and real eigenvalues. *SIAM J. Matrix Anal. Appl.*, 14(1):82–91, 1993b. 66, 108, 109

J. Rohn. Inverse interval matrix. *SIAM J. Numer. Anal.*, 30(3):864–870, 1993c. 65, 115

J. Rohn. Stability of the optimal basis of a linear program under uncertainty. *Oper. Res. Lett.*, 13(1): 9–12, 1993d. 176

J. Rohn. Checking positive definiteness or stability of symmetric interval matrices is NP-hard. *Commentat. Math. Univ. Carol.*, 35(4):795–797, 1994a. 110

J. Rohn. Positive definiteness and stability of interval matrices. *SIAM J. Matrix Anal. Appl.*, 15(1): 175–184, 1994b. 109, 110

J. Rohn. Enclosing solutions of linear interval equations is NP-hard. *Comput.*, 53(3-4):365–368, 1994c. 69

J. Rohn. NP-hardness results for some linear and quadratic problems. Technical Report 619, Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague, January 1995. URL `http://hdl.handle.net/11104/0122691`. 86, 171, 172

J. Rohn. An algorithm for checking stability of symmetric interval matrices. *IEEE Trans. Autom. Control*, 41(1):133–136, 1996. 109

J. Rohn. Complexity of some linear problems with interval data. *Reliab. Comput.*, 3(3):315–323, 1997. 171, 172

J. Rohn. Bounds on eigenvalues of interval matrices. *ZAMM, Z. Angew. Math. Mech.*, 78(Suppl. 3): S1049–S1050, 1998a. 106

J. Rohn. Linear programming with inexact data is NP-hard. *ZAMM, Z. Angew. Math. Mech.*, 78 (Supplement 3):S1051–S1052, 1998b. 95

J. Rohn. Solvability of systems of linear interval equations. *SIAM J. Matrix Anal. Appl.*, 25(1):237–245, 2003. 93

J. Rohn. A method for handling dependent data in interval linear systems. Technical Report 911, Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague, 2004. URL `http://hdl.handle.net/11104/0010645`. 79

J. Rohn. Solvability of systems of interval linear equations and inequalities. In M. Fiedler et al., editor, *Linear Optimization Problems with Inexact Data*, chapter 2, pages 35–77. Springer, New York, 2006a. 19, 62, 92, 97

J. Rohn. Interval linear programming. In M. Fiedler et al., editor, *Linear Optimization Problems with Inexact Data*, chapter 3, pages 79–100. Springer, New York, 2006b. 169, 171

J. Rohn. Forty necessary and sufficient conditions for regularity of interval matrices: A survey. *Electron. J. Linear Algebra*, 18:500–512, 2009. 66

J. Rohn. An improvement of the Bauer-Skeel bounds. Technical Report V-1065, Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague, 2010. URL `http://hdl.handle.net/11104/0182727`. 57

J. Rohn. Explicit inverse of an interval matrix with unit midpoint. *Electron. J. Linear Algebra*, 22: 138–150, 2011. 115

J. Rohn. A handbook of results on interval linear problems. Technical Report 1163, Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague, 2012a. URL `http://hdl.handle.net/11104/0212095`. 62, 63, 65, 66, 103, 112, 115

J. Rohn. A manual of results on interval linear problems. Technical Report 1164, Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague, 2012b. URL `http://hdl.handle.net/11104/0212115`. 40, 43, 44, 62, 63, 65, 67, 70, 109, 178, 186

J. Rohn. On Rump's characterization of P-matrices. *Optim. Lett.*, 6(5):1017–1020, 2012c. 66

J. Rohn. A Farkas-type theorem for interval linear inequalities. *Optim. Lett.*, 8(4):1591–1598, 2014. 94

J. Rohn. Decision making under uncertainty: A mathematician's view of planned economy. Technical report V-1269, Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague, 2019a. URL `http://hdl.handle.net/11104/0297666`. in Czech. 84

J. Rohn. VERSOFT: Verification software in MATLAB / INTLAB, version 10, 2019b. URL `http://uivtx.cs.cas.cz/~rohn/matlab/`. 36

J. Rohn and A. Deif. On the range of eigenvalues of an interval matrix. *Comput.*, 47(3-4):373–377, 1992. 109

J. Rohn and R. Farhadsefat. Inverse interval matrix: A survey. *Electron. J. Linear Algebra*, 22:704–719, 2011. 115

J. Rohn and V. Kreinovich. Computing exact componentwise bounds on solutions of linear systems with interval data is NP-hard. *SIAM J. Matrix Anal. Appl.*, 16(2):415–420, 1995. 69

J. Rohn and J. Kreslová. Linear interval inequalities. *Linear Multilinear Algebra*, 38(1-2):79–82, 1994. 96

J. Rohn and G. Rex. Interval P-matrices. *SIAM J. Matrix Anal. Appl.*, 17(4):1020–1024, 1996. 115

J. Rokne. A note on the Bernstein algorithm for bounds for interval polynomials. *Comput.*, 21(2):159–170, 1979. 129

J. Rokne and P. Lancaster. Complex interval arithmetic. *Commun. ACM*, 14:111–112, 1971. 135

C. Roos, T. Terlaky, and J.-P. Vial. *Interior Point Methods for Linear Optimization*. Springer, New York, 2nd edition, 2006. 18

W. Rudin. *Real and Complex Analysis*. McGraw-Hill, New York, 3rd edition, 1987. 14

S. M. Rump. *Kleine Fehlerschranken bei Matrixproblemen*. PhD thesis, Universität Karlsruhe, Karlsruhe, 1980. in German. 55

S. M. Rump. Solving algebraic problems with high accuracy. In U. Kulisch and W. Miranker, editors, *A new approach to scientific computation, Proc. Symp.*, pages 51–120. Academic Press, New York, 1983. 55, 67, 72, 186

S. M. Rump. Rigorous sensitivity analysis for systems of linear and nonlinear equations. *Math. Comput.*, 54(190):721–736, 1990. 52, 71

S. M. Rump. On the solution of interval linear systems. *Comput.*, 47(3):337–353, Sep 1992. 55

S. M. Rump. Verification methods for dense and sparse systems of equations. In J. Herzberger, editor, *Topics in Validated Computations*, Studies in Computational Mathematics, pages 63–136, Amsterdam, 1994. Elsevier. Proceedings of the IMACS-GAMM International Workshop on Validated Computations, University of Oldenburg. 21, 78, 79, 150

S. M. Rump. Almost sharp bounds for the componentwise distance to the nearest singular matrix. *Linear Multilinear Algebra*, 42(2):93–107, 1997a. 68

S. M. Rump. Bounds for the componentwise distance to the nearest singular matrix. *SIAM J. Matrix Anal. Appl.*, 18(1):83–103, 1997b. 68, 70, 189

S. M. Rump. A note on epsilon-inflation. *Reliab. Comput.*, 4(4):371–375, 1998. 186

S. M. Rump. INTLAB – INTerval LABoratory. In T. Csendes, editor, *Developments in Reliable Computing*, pages 77–104. Kluwer Academic Publishers, Dordrecht, 1999a. `http://www.ti3.tu-harburg.de/rump/`. 36, 138

S. M. Rump. Fast and parallel interval arithmetic. *BIT Numer. Math.*, 39(3):534–554, 1999b. 138

S. M. Rump. Verification methods: Rigorous results using floating-point arithmetic. *Acta Numer.*, 19: 287–449, 2010. 21, 37, 54, 71, 79, 150, 152, 183, 184, 185

S. M. Rump. Accurate solution of dense linear systems, part I: Algorithms in rounding to nearest. *J. Comput. Appl. Math.*, 242:157–184, 2013a. 184

S. M. Rump. Accurate solution of dense linear systems, part II: Algorithms using directed rounding. *J. Comput. Appl. Math.*, 242:185–212, 2013b. 184

T. Rzeżuchowski and J. Wąsowski. Characterization of AE solution sets of parametric linear systems based on the techniques of convex sets. *Linear Algebra Appl.*, 533, 2017. 84

J. F. Sanjuan-Estrada, L. G. Casado, and I. Garcia. Reliable algorithms for ray intersection in computer graphics based on interval arithmetic. In *XVI Brazilian Symposium on Computer Graphics and Image Processing, SIBGRAPI 2003*, pages 35–42, 2003. 187

H. Schichl and A. Neumaier. Exclusion regions for systems of equations. *SIAM J. Numer. Anal.*, 42(1): 383–408, 2004. 126, 150

H. Schichl, F. Domes, T. Montanher, and K. Kofler. Interval unions. *BIT*, 57:531–556, 2017. 139

S. Schirra. Robustness and precision issues in geometric computation. In J.-R. Sack and J. Urrutia, editors, *Handbook of Computational Geometry*, chapter 14, pages 597–632. Elsevier, North-holland, 2000. 187

A. Schrijver. *Theory of Linear and Integer Programming. Repr.* Wiley, Chichester, 1998. 15, 18, 19, 76

H. Schwandt. *Schnelle fast global konvergente Verfahren für die Fünf-Punkt-Diskretisierung der Poissongleichung mit Dirichletschen Randbedingungen auf Rechteckgebieten.* PhD thesis, Technische Universität Berlin, 1981. 150

T. Sederberg and D. B. Buehler. Offsets of polynomial Bézier curves: Hermite approximation with error bounds. In T. Lyche and L. L. Schumaker, editors, *Mathematical Methods in Computer Aided Geometric Design II*, pages 549–558. Academic Press, 1992. 129

T. W. Sederberg and R. T. Farouki. Approximation by interval Bézier curves. *IEEE Comput. Graph. Appl.*, 12:87–95, 1992. URL `https://api.semanticscholar.org/CorpusID:17180427`. 129

N. P. Seif, M. A. Hassanein, and A. S. Deif. Inverse problem of the interval linear system of equations. *Comput.*, 63(2):185–200, 1999.

P. Serafini. Linear programming with variable matrix entries. *Oper. Res. Lett.*, 33(2):165–170, 2005.

I. A. Sharaya. Quantifier-free descriptions for quantifier solutions to interval linear systems of relations. *Trudy Instituta Matematiki i Mekhaniki UrO RAN*, 20(2):311–323, 2014a. English version at `https://arxiv.org/abs/1802.09199`. 100

I. A. Sharaya. IntLinInc2D and IntLinInc3D, 2014b. URL `http://interval.ict.nsc.ru/Programing/`. Release 01.09.2014. 43

S. P. Shary. On controlled solution set of interval algebraic systems. *Interval Comput.*, 4(6):66–75, 1992. 83

S. P. Shary. Solving the linear interval tolerance problem. *Math. Comput. Simul.*, 39(1-2):53–85, 1995a. 84

S. P. Shary. Linear static systems under interval uncertainty: Efficient algorithms to solve control and stabilization problems. *Computational Technologies*, 13:64–80, 1995b. in Russian. 82

S. P. Shary. Linear static systems under interval uncertainty: Algorithms to solve control and stabilization problems. *Reliab. Comput.*, 1(Suppl.):181–184, 1995c. Extended Abstracts of APIC'95: International Workshop on Applications of Interval Computations, El Paso, Febr. 23–25. 134

S. P. Shary. A new technique in systems analysis under interval uncertainty and ambiguity. *Reliab. Comput.*, 8(5):321–418, 2002. 82, 83, 84, 133, 134

S. P. Shary. On full-rank interval matrices. *Numer. Anal. Appl.*, 7(3):241–254, 2014. 112

S. P. Shary and B. Moradi. Solving interval linear least squares problems by PPS-methods. *Numer. Algorithms*, 87(1):41–75, 2021. 72

I. Skalna. A method for outer interval solution of systems of linear equations depending linearly on interval parameters. *Reliab. Comput.*, 12(2):107–120, 2006. 77, 78, 79

I. Skalna. On checking the monotonicity of parametric interval solution of linear structural systems. *LNCS*, 4967:1400–1409, 2008. 79

I. Skalna. Enclosure for the solution set of parametric linear systems with non-affine dependencies. In R. Wyrzykowski et al., editor, *Parallel Processing and Applied Mathematics*, volume 7204 of *LNCS*, pages 513–522. Springer, 2012. 73

I. Skalna. *Parametric Interval Algebraic Systems*, volume 766 of *Studies in Computational Intelligence*. Springer, Cham, 2018. 77, 79

I. Skalna and M. Hladík. A new method for computing a p-solution to parametric interval linear systems with affine-linear and nonlinear dependencies. *BIT Numer. Math.*, 57(4):1109–1136, 2017. ISSN 1572-9125. doi: 10.1007/s10543-017-0679-4. URL `https://link.springer.com/article/10.1007/s10543-017-0679-4`. 79

I. Skalna and M. Hladík. Direct and iterative methods for interval parametric algebraic systems producing parametric solutions. *Numer. Linear Algebra Appl.*, 26(3):e2229:1–e2229:24, 2019. ISSN 1099-1506. doi: 10.1002/nla.2229. URL `https://onlinelibrary.wiley.com/doi/abs/10.1002/nla.2229`. 79, 115

S. Skelboe. Computation of rational interval functions. *BIT Numer. Math.*, 14(1):87–95, 1974. 32, 118, 167

A. Skjäl, T. Westerlund, R. Misener, and C. A. Floudas. A generalization of the classical αBB convex underestimation via diagonal and nondiagonal quadratic terms. *J. Optim. Theory Appl.*, 154(2):462–490, 2012. 166

S. Smale. Mathematical problems for the next century. *Math. Intell.*, 20(2):7–15, 1998. 34

A. P. Smith. Fast construction of constant bound functions for sparse polynomials. *J. Glob. Optim.*, 43 (2-3):445–458, 2009. 131

A. P. Smith, J. Garloff, and H. Werkle. Verified solution for a statically determinate truss structure with uncertain node locations. *J. Civ. Eng. Architect.*, 4(11):1–10, 2010. 79

L. B. Smith. Interval arithmetic determinant evaluation and its use in testing for a chebyshev system. *Commun. ACM*, 12(2):89–93, 1969. 113

J. M. Snyder. *Generative Modeling for Computer Graphics and CAD: Symbolic Shape Design Using Interval Analysis*. Academic Press, San Diego, CA, USA, 1992. 187

O. Spaniol. Die Distributivität in der Intervallarithmetik. *Comput.*, 5(1):6–16, 1970. in German. 28

V. Stahl. *Interval Methods for Bounding the Range of Polynomials and Solving Systems of Nonlinear Equations*. Dissertation, Johannes Kepler University Linz, Austria, 1995. 127, 128, 129

N. F. Stewart. Guaranteed error bounds for the initial value problem using polytope arithmetic. *BIT*, 11:328–337, 1971.

J. Stoer and R. Bulirsch. *Introduction to Numerical Analysis*, volume 12 of *Texts in Applied Mathematics*. Springer, New York, 3rd edition, 2002. 13

N. Stolte and A. Kaufman. Parallel spatial enumeration of implicit surfaces using interval arithmetic for octree generation and its direct visualisation. In *Implicit Surfaces'98. The Third International Workshop on Implicit Surfaces*, pages 81–87, Seattle, 1998. 24

K. G. Suffern and E. D. Fackerell. Interval methods in computer graphics. *Comput. Graph.*, 15(3): 331–340, 1991. 187

T. Sunaga. Theory of an interval algebra and its application to numerical analysis. *RAAG Memoirs*, 2: 29–46, 1958. 33, 151

A. Tarski. *A Decision Method for Elementary Algebra and Geometry.* University of California Press, Berkeley, 2nd edition, 1951. 143

M. Tawarmalani and N. V. Sahinidis. A polyhedral branch-and-cut approach to global optimization. *Math. Program.*, 103(2B):225–249, 2005. 165, 168

G. Trombettoni, Y. Papegay, G. Chabert, and O. Pourtallier. A box-consistency contractor based on extremal functions. In D. Cohen, editor, *Principles and Practice of Constraint Programming – CP 2010*, volume 6308 of *LNCS*, pages 491–498. Springer, 2010. 157

G. Trombettoni, I. Araya, B. Neveu, and G. Chabert. Inner regions and interval linearizations for global optimization. In W. Burgard and D. Roth, editors, *Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence, AAAI 2011, San Francisco, California, USA*. AAAI Press, 2011. 164, 180

S. Vajda. *Mathematical Programming.* Addison-Wesley, Reading Mass., USA, 1961. 85, 96, 173

L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM Rev.*, 38(1):49–95, 1996. 105

A. A. Vatolin. On the problems of linear programming with integral coefficients. *USSR Comput. Math. Math. Phys.*, 4(11-12):18–23, 1984. 100

S. A. Vavasis. *Nonlinear Optimization: Complexity Issues.* Oxford University Press, New York, 1991. 111

C. Viegas, D. Daney, M. Tavakoli, and A. T. de Almeida. Performance analysis and design of parallel kinematic machines using interval analysis. *Mech. Mach. Theory*, 115:218–236, 2017. 160

W. V. Walter. FORTRAN-XSC: A portable Fortran 90 module library for accurate and reliable scientific computing. In R. Albrecht et al., editor, *Validation Numerics: Theory and Applications*, volume 9 of *Computing Suppl. 9*, pages 265–285. Springer, Vienna, 1993. 37

K. Wang, A. N. Michel, and D. Liu. Necessary and sufficient conditions for the hurwitz and schur stability of interval matrices. *IEEE Trans. Autom. Control*, 39(6):1251 –1255, 1994. 109

S. Wang, K. Wang, L. Wu, and C. Wang. Polar affine arithmetic: Optimal affine approximation and operation development for computation in polar form under uncertainty. *ACM Trans. Math. Softw.*, 45(1):6:1–6:29, 2019. 138

M. Warmus. Calculus of approximations. *Bull. Acad. Pol. Sci., Cl. III*, 4:253–259, 1956. 33

M. E. Wieser, N. Holden, T. B. Coplen, J. K. Böhlke, M. Berglund, W. A. Brand, P. De Bièvre, M. Gröning, R. D. Loss, J. Meija, T. Hirata, T. Prohaska, R. Schoenberg, G. O'Connor, T. Walczyk, S. Yoneda, and X.-K. Zhu. Atomic weights of the elements 2011 (IUPAC technical report). *Pure Appl. Chem*, 85(5):1047–1078, 2013. 24

Y. Xiao and R. Unbehauen. Robust hurwitz and schur stability test for interval matrices. In *Proceedings of the 39th IEEE Conference on Decision and Control*, volume 5, pages 4209–4214, Sydney, Australia, 2000. 109

A. G. Yakovlev. Computer arithmetic of multiintervals. *Problems of Cybernetics*, 125:66–81, 1986. Scientific Council on the Complex Problem "Cybernetics" of the Academy of Sciences of the USSR. 139

R. C. Young. The algebra of many-valued quantities. *Math. Ann.*, 104(1):260–290, 1931. 33

F. Zapata, V. Kreinovich, C. Joslyn, and E. Hogan. Orders on intervals over partially ordered sets: extending Allen's algebra and interval graph results. *Soft Comput.*, 17(8):1379–1391, 2013. 25

F. Zhou, G. H. Huang, G.-X. Chen, and H.-C. Guo. Enhanced-interval linear programming. *Eur. J. Oper. Res.*, 199(2):323–333, 2009.

W. Zhu. Unsolvability of some optimization problems. *Appl. Math. Comput.*, 174(2):921–926, 2006. 143

M. Zimmer, W. Krämer, and E. D. Popova. Solvers for the verified solution of parametric linear systems. *Comput.*, 94(2-4):109–123, 2012. 79