Introduction to Interval Computation and Numerical Verification

part II.

Milan Hladík

Faculty of Mathematics and Physics,
Charles University in Prague, Czech Republic
http://kam.mff.cuni.cz/~hladik/

Seminář numerické analýzy a zimní škola – SNA'17
Ostrava, 30. ledna – 3. února 2017

Outline

Next Section

Mean value form

Theorem

Let $f : \mathbb{R}^n \mapsto \mathbb{R}$, $\boldsymbol{x} \in \mathbb{IR}^n$ and $a \in \boldsymbol{x}$. Then

$$f(\boldsymbol{x}) \subseteq f(a) + \nabla f(\boldsymbol{x})^T(\boldsymbol{x} - a),$$

Proof.

By the mean value theorem, for any $x \in \boldsymbol{x}$ there is $c \in \boldsymbol{x}$ such that

$$f(x) = f(a) + \nabla f(c)^T(x - a) \in f(a) + \nabla f(\boldsymbol{x})^T(\boldsymbol{x} - a). \qquad \square$$

Improvements

- successive mean value form

$$\begin{aligned} f(\boldsymbol{x}) \subseteq f(a) &+ f'_{x_1}(\boldsymbol{x}_1, a_2, \ldots, a_n)(\boldsymbol{x}_1 - a_1) \\ &+ f'_{x_2}(\boldsymbol{x}_1, \boldsymbol{x}_2, a_3 \ldots, a_n)(\boldsymbol{x}_2 - a_2) + \ldots \\ &+ f'_{x_n}(\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{n-1}, \boldsymbol{x}_n)(\boldsymbol{x}_n - a_n). \end{aligned}$$

- replace derivatives by slopes

Slopes

Slope form enclosure

$$f(\boldsymbol{x}) \subseteq f(a) + S(\boldsymbol{x}, a)(\boldsymbol{x} - a),$$

where $a \in \boldsymbol{x}$ and

$$S(x,a) := \begin{cases} \frac{f(x)-f(a)}{x-a} & \text{if } x \neq a, \\ f'(x) & \text{otherwise.} \end{cases}$$

Remarks

- Slopes can be replaced by derivatives, but slopes are tighter.
- Slopes can be computed in a similar way as derivatives.

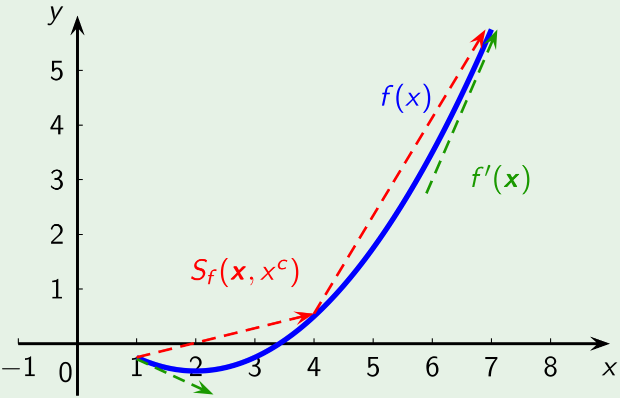
function	its slope $S(x,a)$
x	1
$f(x) \pm g(x)$	$S_f(x,a) \pm S_g(x,a)$
$f(x) \cdot g(x)$	$S_f(x,a)g(a) + f(x)S_g(x,a)$
$e^{f(x)}$	$e^{f(x)} S_f(x,a)$

Slopes

Example

$$f(x) = \tfrac{1}{4}x^2 - x + \tfrac{1}{2}, \qquad \boldsymbol{x} = [1, 7].$$

$$f'(\boldsymbol{x}) = [-\tfrac{1}{2}, \tfrac{5}{2}], \qquad S_f(\boldsymbol{x}, x^c) = [\tfrac{1}{4}, \tfrac{7}{4}].$$



Notice: Slopes cannot be used for monotonicity checking.

Next Section

Nonlinear Equations

Problem statement

Find all solutions to

$$f_j(x_1, \ldots, x_n) = 0, \quad j = 1, \ldots, j^*$$

inside the box $\boldsymbol{x}^0 \in \mathbb{IR}^n$.

Theorem (Zhu, 2005)

For a polynomial $p(x_1, \ldots, x_n)$*, there is no algorithm solving*

$$p(x_1, \ldots, x_n)^2 + \sum_{i=1}^{n} \sin^2(\pi x_i) = 0.$$

Proof.

From Matiyasevich's theorem solving the 10th Hilbert problem. □

Remark

Using the arithmetical operations only, the problem is decidable by Tarski's theorem (1951).

Interval Newton method

Classical Newton method

. . . is an iterative method

$$x^{k+1} := x^k - \nabla f(x^k)^{-1} f(x^k), \quad k = 0, \dots$$

Cons

- Can miss some solutions
- Not verified (Are we really close to the true solution?)

Interval Newton method – Stupid intervalization

$$\boldsymbol{x}^{k+1} := \boldsymbol{x}^k - \nabla f(\boldsymbol{x}^k)^{-1} f(\boldsymbol{x}^k), \quad k = 0, \dots$$

Interval Newton method – Good intervalization

$$\begin{aligned} N(x^k, \boldsymbol{x}^k) &:= x^k - \nabla f(\boldsymbol{x}^k)^{-1} f(x^k), \\ \boldsymbol{x}^{k+1} &:= \boldsymbol{x}^k \cap N(x^k, \boldsymbol{x}^k), \qquad k = 0, \dots \end{aligned}$$

Interval Newton method

Theorem (Moore, 1966)

If $x, x^0 \in \boldsymbol{x}$ and $f(x) = 0$, then $x \in N(x^0, \boldsymbol{x})$.

Proof.

By the Mean value theorem,

$$f_i(x) - f_i(x^0) = \nabla f_i(c_i)^T (x - x^0), \quad \forall i = 1, \dots, n.$$

If x is a root, we have

$$-f_i(x^0) = \nabla f_i(c_i)^T (x - x^0).$$

Define $A \in \mathbb{R}^{n \times n}$ such that its ith row is equal to $\nabla f_i(c_i)^T$. Hence

$$-f(x^0) = A(x - x^0),$$

from which

$$x = x^0 - A^{-1} f(x^0) \in x^0 - \nabla f(\boldsymbol{x})^{-1} f(x^0).$$

Notice, that this does not mean that there is $c \in \boldsymbol{x}$ such that

$$-f(x^0) = \nabla f(c)(x - x^0).$$

□

Interval Newton method

Theorem (Nickel, 1971)

If $\emptyset \neq N(x^0, \boldsymbol{x}) \subseteq \boldsymbol{x}$*, then there is a unique root in* $\boldsymbol{x}$ *and* $\nabla f(\boldsymbol{x})$ *is regular.*

Proof.

"Regularity." Easy.

"Existence." By Brouwer's fixed-point theorem.

[Any continuous mapping of a compact convex set into itself has a fixed point.]

"Uniqueness." If there are two roots $y_1 \neq y_2$ in $\boldsymbol{x}$, then by the Mean value theorem,

$$f(y_1) - f(y_2) = A(y_1 - y_2)$$

for some $A \in \nabla f(\boldsymbol{x})$;. Since $f(y_1) = f(y_2) = 0$, we get

$$A(y_1 - y_2) = 0$$

and by the nonsingularity of A, the roots are identical. □

Interval Newton method

Practical implementation

Instead of

$$N(x^k, \boldsymbol{x}^k) := x^k - \nabla f(\boldsymbol{x}^k)^{-1} f(x^k)$$

let $N(x^k, \boldsymbol{x}^k)$ be an enclosure of the solution set (with respect to x) of

$$\nabla f(\boldsymbol{x})(x - x^0) = -f(x^0).$$

Extended interval arithmetic

So far

$$\frac{[12, 15]}{[-2, 3]} = (-\infty, \infty).$$

Now,

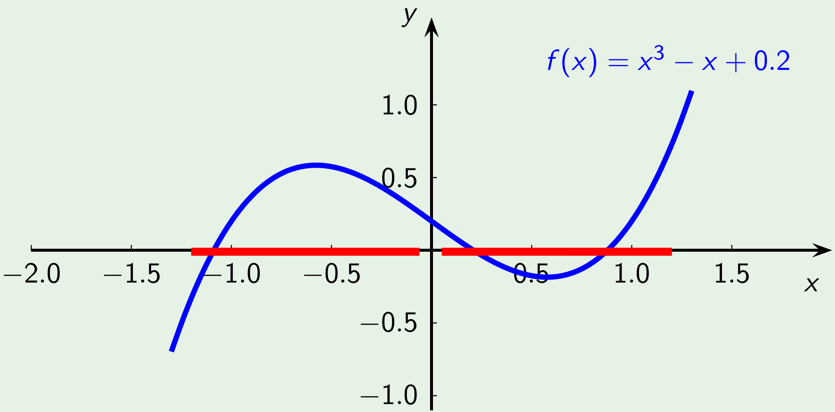
$$\boldsymbol{a}/\boldsymbol{b} := \{a/b : a \in \boldsymbol{a}, 0 \neq b \in \boldsymbol{b}\}.$$

So,

$$\frac{[12, 15]}{[-2, 3]} = (-\infty, -6] \cup [4, \infty).$$

Interval Newton method

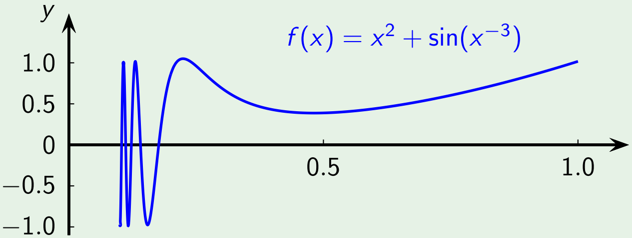
Example



In six iterations precision 10^{-11} (quadratic convergence).

Interval Newton method

Example (Moore, 1993)



All 318 roots of in the interval $[0.1, 1]$ found with accuracy 10^{-10}.
The left most root is contained in $[0.10003280626, 0.10003280628]$.

Summary

- $N(x^0, \boldsymbol{x})$ contains all solutions in $\boldsymbol{x}$
- If $\boldsymbol{x} \cap N(x^0, \boldsymbol{x}) = \emptyset$, then there is no root in $\boldsymbol{x}$
- If $\emptyset \neq N(x^0, \boldsymbol{x}) \subseteq \boldsymbol{x}$, then there is a unique root in $\boldsymbol{x}$

Krawczyk method

Krawczyk operator

Let $x^0 \in \boldsymbol{x}$ and $C \in \mathbb{R}^{n \times n}$, usually $C \approx \nabla f(x^0)^{-1}$. Then

$$K(\boldsymbol{x}) := x^0 - Cf(x^0) + (I_n - C\nabla f(\boldsymbol{x}))(\boldsymbol{x} - x^0).$$

Theorem

Any root of $f(x)$ in $\boldsymbol{x}$ is included in $K(\boldsymbol{x})$.

Proof.

If x^1 is a root of $f(x)$, then it is a fixed point of

$$g(x) := x - Cf(x).$$

By the mean value theorem,

$$g(x^1) \in g(x^0) + \nabla g(\boldsymbol{x})(x^1 - x^0),$$

whence

$$\begin{aligned} x^1 \in g(\boldsymbol{x}) &\subseteq g(x^0) + \nabla g(\boldsymbol{x})(\boldsymbol{x} - x^0) \\ &= x^0 - Cf(x^0) + (I_n - C\nabla f(\boldsymbol{x}))(\boldsymbol{x} - x^0). \end{aligned}$$

□

Krawczyk method

Theorem

If $K(\boldsymbol{x}) \subseteq \boldsymbol{x}$, then there is a root in $\boldsymbol{x}$.

Proof.

Recall

$$g(x) := x - Cf(x).$$

By the proof of the previous Theorem, $K(\boldsymbol{x}) \subseteq \boldsymbol{x}$ implies

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

Thus, there is a fixed point $x^0 \in \boldsymbol{x}$ of $g(x)$,

$$g(x^0) = x^0 - Cf(x^0) = x^0,$$

so x^0 is a root of $f(x)$. $\square$

Krawczyk method

Theorem (Kahan, 1968)

If $K(\boldsymbol{x}) \subseteq \mathit{int}\,\boldsymbol{x}$, then there is a unique root in $\boldsymbol{x}$ and $\nabla f(\boldsymbol{x})$ is regular.

Recall Theorem from "ε-inflation" (for solving $\boldsymbol{A}x = \boldsymbol{b}$)

Let $\boldsymbol{x} \in \mathbb{IR}^n$ and $C \in \mathbb{R}^{n \times n}$. If

$$K(\boldsymbol{x}) = C\boldsymbol{b} + (I_n - C\boldsymbol{A})\boldsymbol{x} \subseteq \mathit{int}\,\boldsymbol{x},$$

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

Proof.

The inclusion $K(\boldsymbol{x}) \subseteq \mathit{int}\,\boldsymbol{x}$ reads

$$-Cf(x^0) + (I_n - C\nabla f(\boldsymbol{x}))(\boldsymbol{x} - x^0) \subseteq \mathit{int}\,(\boldsymbol{x} - x^0)$$

Apply the above Theorem for

$$\boldsymbol{b} := -f(x^0), \quad \boldsymbol{A} := \nabla f(\boldsymbol{x}), \quad \boldsymbol{x} := \boldsymbol{x} - x^0$$

We have that $\nabla f(\boldsymbol{x})$ is regular, which implies uniqueness. $\square$

More general constraints

Constraints

- equations $h_i(x) = 0,\ i = 1, \dots, I$
- inequalities $g_j(x) \leq 0,\ j = 1, \dots, J$
- may be others, but not considered here
  ($\neq$, quantifications, logical operators, lexicographic orderings, ...)

Problem

Denote by Σ the set of solutions in an initial box $\boldsymbol{x}^0 \in \mathbb{IR}^n$?

Problem: How to describe Σ?

Subpavings

Split $\boldsymbol{x}$ into a union of three sets of boxes such that

- the first set has boxes provably containing no solution
- the second set has boxes that provably consist of only solutions
- the third set has boxes which may or may not contain a solution

Subpaving Example

Example

$$x^2 + y^2 \leq 16,$$
$$x^2 + y^2 \geq 9$$

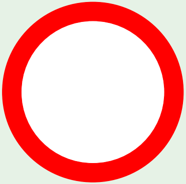


Figure: Exact solution set

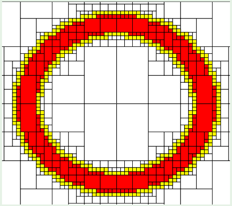


Figure: Subpaving approximation

Subpaving Example

Example

$$(x-1)^2 + (y-2)^2 \leq \frac{1}{7},$$
$$(x^2 + y^2 - 9)\left(\tfrac{1}{3}x - y^2\right) \geq \frac{1}{2}$$

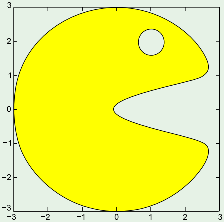


Figure: Exact solution set

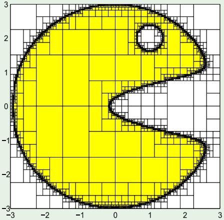


Figure: Subpaving approximation

Subpaving Algorithm

Branch & Bound approach

- divide $\boldsymbol{x}^0$ recursively into sub-boxes,
- remove sub-boxes with provably no solutions
- contract sub-boxes

Some simple tests

- Test for $\boldsymbol{x} \subseteq \Sigma$:
  - no equations and $\overline{g}_j(\boldsymbol{x}) \leq 0 \ \forall j$
- Test for $\boldsymbol{x} \cap \Sigma = \emptyset$:
  - $0 \notin h_i(\boldsymbol{x})$ for some i
  - $\underline{g}_j(\boldsymbol{x}) > 0$ for some j

Also very important

- Which box to choose (data structure fo $\mathcal{L}$)?
- How to divide the box? (which coordinate, which place, how many sub-boxex)

A Simple Contractor – Constraint Propagation

Example

Consider the constraint

$$x + yz = 7, \quad x \in [0,3],\ y \in [3,5],\ z \in [2,4].$$

- Express x

$$x = 7 - yz \in 7 - [3,5][2,4] = [-13,1].$$

  Thus, the domain for 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 for 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 for z is $[2,4] \cap [\frac{12}{7}, \frac{7}{3}] = [2, \frac{7}{3}]$.

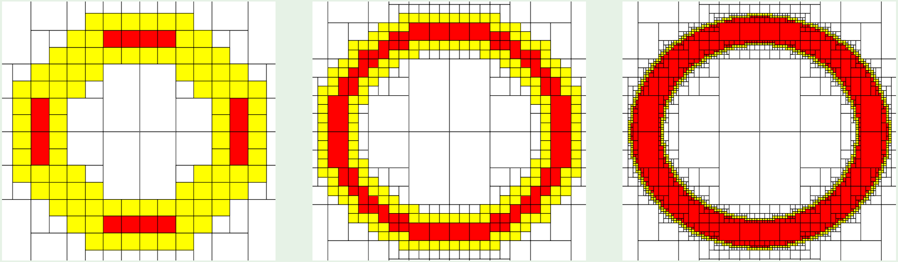
No further propagation needed as each variable appears just once.

Other Techniques

Other techniques

- Various kinds of consistencies (2B, 3B,...), shaving,...

Example (thanks to Elif Garajová)



$\varepsilon = 1.0$
time: 0.952 s

$\varepsilon = 0.5$
time: 2.224 s

$\varepsilon = 0.125$
time: 9.966 s

Software

Free constraint solving software

- *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.emn.fr/z-info/ibex`
- *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/info/perso/permanents/granvil/realpaver`
- *RSolver* (by Stefan Ratschan),
  solver for quantified constraints over the real numbers,
  implemented in the programming language OCaml,
  `http://rsolver.sourceforge.net/`

References

G. Alefeld and J. Herzberger.
Introduction to Interval Computations.
Academic Press, New York, 1983.

F. Benhamou and L. Granvilliers.
Continuous and interval constraints.
In *Handbook of Constraint Programming*, chap. 16. Elsevier, 2006.

G. Chabert and L. Jaulin.
Contractor programming.
Artif. Intell., 173(11):1079-1100, 2009.

F. Goualard and C. Jermann.
A reinforcement learning approach to interval constraint propagation.
Constraints, 13(1):206–226, 2008.

L. Jaulin, M. Kieffer, O. Didrit, and É. Walter.
Applied Interval Analysis.
Springer, London, 2001.

Next Section

Introduction

Rigorous computation

What and why?

Can we obtain rigorous numerical results by using floating-point arithmetic?

Yes, by extending to interval arithmetic. Direct usage is however not effective!

Example (Amplification factor for the interval Gaussian elimination)

$n = 20$	$n = 50$	$n = 100$	$n = 170$
10^2	10^5	10^{10}	10^{16}

Advise

Postpone interval computation to the very end.

Verification

Verification

Compute a solution by floating-point arithmetic, and then to verify that the result is correct or determine rigorous distance to a true solution.

Typically, we can prove uniqueness (=the problem is well posed). Therefore, verifying singularity of a matrix cannot be performed!

What we will do

As an example, we show a verification method for the problem of finding a root of a function $f : \mathbb{R}^n \to \mathbb{R}^n$.

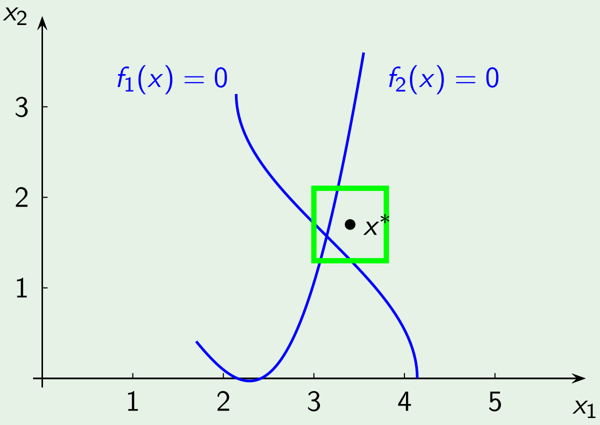
Problem statement

Given $x^* \in \mathbb{R}^n$ a numerically computed (=approximate) solution of $f(x) = 0$, find a small interval $0 \in \boldsymbol{y} \in \mathbb{IR}^n$ such that the true solution lies in $x^* + \boldsymbol{y}$.

Illustration of Verification

Example

Illustration of the verification of x^* to be a solution of $f(x) = 0$.



Ingredients

Brouwer fixed-point theorem

Let U be a convex compact set in $\mathbb{R}^n$ and $g\colon U \to U$ a continuous function. Then there is a fixed point, i.e., $\exists x \in U : g(x) = x$.

Observation

Finding a root of $f(x)$ is equivalent to finding a fixed-point of the function $g(y) \equiv y - C \cdot f(x^* + y)$, where C is any nonsingular matrix of order n.

Perron theory of nonnegative matrices

- If $|A| \leq B$, then $\rho(A) \leq \rho(B)$.
  ($\leq$ is meant entrywise and $\rho(\cdot)$ is the spectral radius)
- If $A \geq 0$, $x > 0$ and $Ax < \alpha x$, then $\rho(A) < \alpha$.

Lemma

If $\boldsymbol{z} + \boldsymbol{R}\boldsymbol{y} \subseteq \mathrm{int}\, \boldsymbol{y}$, then $\rho(R) < 1$ for every $R \in \boldsymbol{R}$.

Proof. $|R| y^{\Delta} < y^{\Delta}$, whence by Perron theory $\rho(R) < 1$. $\square$

Cooking

Theorem

Suppose $0 \in \boldsymbol{y}$*. Now if*

$$-C \cdot f(x^*) + (I - C \cdot \nabla f(x^* + \boldsymbol{y})) \cdot \boldsymbol{y} \subseteq int\, \boldsymbol{y},$$

then:

- C *and every matrix in* $\nabla f(x^* + \boldsymbol{y})$ *are nonsingular, and*
- *there is a unique root of* $f(x)$ *in* $x^* + \boldsymbol{y}$*.*

Proof.

By the mean value theorem,

$$f(x^* + y) \in f(x^*) + \nabla f(x^* + \boldsymbol{y})y.$$

By the assumptions, the function

$$g(y) = y - C \cdot f(x^* + y) \in -C \cdot f(x^*) + (I - C \cdot \nabla f(x^* + \boldsymbol{y}))\boldsymbol{y} \subseteq int\, \boldsymbol{y}$$

has a fixed point, which shows “existence”.

By Lemma, C and $\nabla f(x^* + \boldsymbol{y})$ are nonsingular, implying “uniqueness”. $\square$

Cooking

Implementation

- take $C \approx \nabla f(x^*)^{-1}$ (numerically computed inverse),
- take $\boldsymbol{y} := C \cdot f(x^*)$ and repeat inflation

$$\boldsymbol{y} := \Big(-C \cdot f(x^*) + (I - C \cdot \nabla f(x^* + \boldsymbol{y})) \cdot \boldsymbol{y} \Big) \cdot [0.9, 1.1] + 10^{-20}[-1, 1]$$

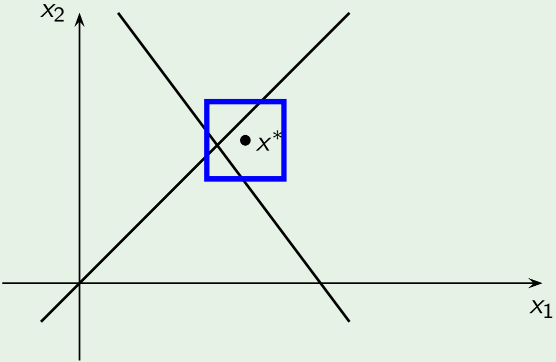
until the assumption of Theorem are satisfied.

Verification of a Linear System of Equations

Problem formulation

Given a real system $Ax = b$ and x^* approximate solution, find $\boldsymbol{y} \in \mathbb{IR}^n$ such that $A^{-1}b \in x^* + \boldsymbol{y}$.

Example



Verification of a Linear System of Equations

Given the system $Ax = b$ and an approximate solution x^*.

Theorem

Suppose $0 \in \boldsymbol{y}$*. Now if*

$$C(b - Ax^*) + (I - CA)\boldsymbol{y} \subseteq \mathit{int}\, \boldsymbol{y},$$

then:

- C *and* A *are nonsingular,*
- *there is a unique solution of* $Ax = b$ *in* $x^* + \boldsymbol{y}$*.*

Proof.

Use the previous result with $f(x) = Ax - b$. □

Implementation

- take $C \approx A^{-1}$ (numerically computed inverse),

Verification of a Linear System of Equations

ε-inflation method (Caprani and Madsen, 1978, Rump, 1980)

Repeat inflating $\boldsymbol{y} := [0.9, 1.1]\boldsymbol{x} + 10^{-20}[-1, 1]$ and updating

$$\boldsymbol{x} := C(b - Ax^*) + (I - CA)\boldsymbol{y}$$

until $\boldsymbol{x} \subseteq \mathit{int}\, \boldsymbol{y}$.

Then, $\Sigma \subseteq x^* + \boldsymbol{x}$.

Results

- Verification is about 7 times slower than solving the original problem (for random instances of dimension 100 to 2000).

Verification of a Linear System of Equations

Example

Let A be the Hilbert matrix of size 10 (i.e., $a_{ij} = \frac{1}{i+j-1}$), and $b := Ae$.

Then $Ax = b$ has the solution $x = e = (1, \ldots, 1)^T$.

Approximate solution by Matlab:	Enclosing interval by ε-inflation method (2 iterations):
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]

Verification of a Linear System of Equations

Challenge

- verification for large systems
  (one cannot use preconditioning by the inverse matrix)

References

S.M. Rump.
Verification methods: Rigorous results using floating-point arithmetic.
Acta Numerica, 19:187–449, 2010.

Next Section

Tolerable Solutions

Motivation

So far, existentially quantified interval systems

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

Now, incorporate universal quantification as well!

Definition (Tolerable solutions)

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 $Ax = b$.

In other words,

$$\forall A \in \boldsymbol{A} \, \exists b \in \boldsymbol{b} : Ax = b.$$

Equivalent characterizations

- $\boldsymbol{A}x \subseteq \boldsymbol{b}$,
- $|A^c x - b^c| \leq -A^\Delta |x| + b^\Delta$.

Tolerable Solutions

Theorem (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.

"$\Leftarrow$" Let $A \in \boldsymbol{A}$. Then

$$Ax = Ax_1 - Ax_2 \leq \overline{A}x_1 - \underline{A}x_2 \leq \overline{b},$$
$$Ax = Ax_1 - Ax_2 \geq \underline{A}x_1 - \overline{A}x_2 \geq \underline{b}$$

Thus, $Ax \in \boldsymbol{b}$ and $Ax = b$ for some $b \in \boldsymbol{b}$.

"$\Rightarrow$" Let $x \in \mathbb{R}^n$ be a tolerable solution. Define $x_1 := \max\{x, 0\}$ and $x_2 := \max\{-x, 0\}$ the positive and negative part of x, respectively. Then $x = x_1 - x_2$, $|x| = x_1 + x_2$, and $|A^c x - b^c| \leq -A^\Delta |x| + b^\Delta$ draws

$$A^c(x_1 - x_2) - b^c \leq -A^\Delta(x_1 + x_2) + b^\Delta,$$
$$-A^c(x_1 - x_2) + b^c \leq -A^\Delta(x_1 + x_2) + b^\Delta.$$

$\square$

Tolerable Solutions – Application

Example (Leontief's Input–Output Model of Economics)

- economy with n sectors (e.g., agriculture, industry, transportation, etc.),
- sector i produces a single commodity of amount x_i,
- production of each unit of the jth commodity will require a_{ij} (amount) of the ith commodity
- 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

$$x = Ax + d.$$

The solution $x = (I_n - A)^{-1}d = \sum_{k=0}^{\infty} A^k d$ is nonnegative if $\rho(A) < 1$.

Question: Exists x such that for any $A \in \boldsymbol{A}$ there is $d \in \boldsymbol{d}$: $(I_n - A)x = d$?

AE Solutions

Quantified system $\boldsymbol{A}x = \boldsymbol{b}$

- each interval parameter $\boldsymbol{a}_{ij}$ and $\boldsymbol{b}_i$ is quantified by $\forall$ or $\exists$
- the universally quantified parameters are denoted by $\boldsymbol{A}^{\forall}$, $\boldsymbol{b}^{\forall}$,
- the existentially quantified parameters are denoted by $\boldsymbol{A}^{\exists}$, $\boldsymbol{b}^{\exists}$
- the system reads $(\boldsymbol{A}^{\forall} + \boldsymbol{A}^{\exists})x = \boldsymbol{b}^{\forall} + \boldsymbol{b}^{\exists}$

Definition (AE solution set)

$$\Sigma_{AE} := \big\{x \in \mathbb{R}^n : \\ \forall A^{\forall} \in \boldsymbol{A}^{\forall}\, \forall b^{\forall} \in \boldsymbol{b}^{\forall}\, \exists A^{\exists} \in \boldsymbol{A}^{\exists}\, \exists b^{\exists} \in \boldsymbol{b}^{\exists} : (A^{\forall} + A^{\exists})x = b^{\forall} + b^{\exists}\big\}.$$

AE Solutions

Theorem (Shary, 1995)

$$\Sigma_{AE} = \{x \in \mathbb{R}^n : \boldsymbol{A}^{\forall}x - \boldsymbol{b}^{\forall} \subseteq \boldsymbol{b}^{\exists} - \boldsymbol{A}^{\exists}x\}. \tag{1}$$

Proof.

$$\begin{aligned}\Sigma_{AE} &= \{x \in \mathbb{R}^n : \forall A^{\forall} \in \boldsymbol{A}^{\forall}\, \forall b^{\forall} \in \boldsymbol{b}^{\forall}\, \exists A^{\exists} \in \boldsymbol{A}^{\exists}\, \exists b^{\exists} \in \boldsymbol{b}^{\exists} : A^{\forall}x - b^{\forall} = b^{\exists} - A^{\exists}x\} \\ &= \{x \in \mathbb{R}^n : \forall A^{\forall} \in \boldsymbol{A}^{\forall}\, \forall b^{\forall} \in \boldsymbol{b}^{\forall} : A^{\forall}x - b^{\forall} \in \boldsymbol{b}^{\exists} - \boldsymbol{A}^{\exists}x\} \\ &= \{x \in \mathbb{R}^n : \boldsymbol{A}^{\forall}x - \boldsymbol{b}^{\forall} \subseteq \boldsymbol{b}^{\exists} - \boldsymbol{A}^{\exists}x\}.\end{aligned}$$

□

Theorem (Rohn, 1996)

$$\Sigma_{AE} = \{x \in \mathbb{R}^n : |A^c x - b^c| \le ((\boldsymbol{A}^{\exists})^{\Delta} - (\boldsymbol{A}^{\forall})^{\Delta})|x| + (\boldsymbol{b}^{\exists})^{\Delta} - (\boldsymbol{b}^{\forall})^{\Delta}\}.$$

Proof.

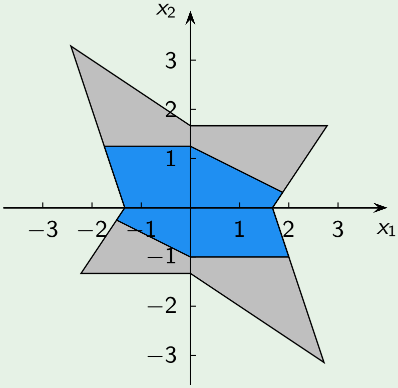
Using (1) and the fact $\boldsymbol{p} \subseteq \boldsymbol{q} \;\Leftrightarrow\; |p^c - q^c| \le q^{\Delta} - p^{\Delta}$, we get

$$\begin{aligned}|(\boldsymbol{A}^{\forall}x - \boldsymbol{b}^{\forall})^c - (\boldsymbol{b}^{\exists} - \boldsymbol{A}^{\exists}x)^c| &\le (\boldsymbol{A}^{\exists}x - \boldsymbol{b}^{\exists})^{\Delta} - (\boldsymbol{b}^{\forall} - \boldsymbol{A}^{\forall}x)^{\Delta} \\ &= (\boldsymbol{A}^{\exists})^{\Delta}|x| + \boldsymbol{b}^{\exists\Delta} - (\boldsymbol{A}^{\forall})^{\Delta}x| - \boldsymbol{b}^{\forall\Delta}.\end{aligned}$$

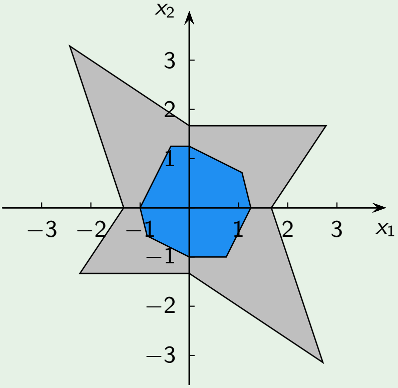
□

Example

$$\begin{pmatrix} [3,4]^{\exists} & [-2,1]^{\exists} \\ [0,2]^{\forall} & [3,4]^{\forall} \end{pmatrix} x = \begin{pmatrix} [-4,5]^{\exists} \\ [-4,5]^{\exists} \end{pmatrix}. \qquad \begin{pmatrix} [3,4]^{\forall} & [-2,1]^{\forall} \\ [0,2]^{\forall} & [3,4]^{\forall} \end{pmatrix} x = \begin{pmatrix} [-4,5]^{\exists} \\ [-4,5]^{\exists} \end{pmatrix}.$$



AE solution set.



Tolerable solution set.

Next Section

Eigenvalues of Symmetric Interval Matrices

A symmetric interval matrix

$$\boldsymbol{A}^S := \{A \in \boldsymbol{A} : A = A^T\}.$$

Without loss of generality assume that $\underline{A} = \underline{A}^T$, $\overline{A} = \overline{A}^T$, and $\boldsymbol{A}^S \neq \emptyset$.

Eigenvalues of a symmetric interval matrix

Eigenvalues of a symmetric $A \in \mathbb{R}^{n \times n}$: $\lambda_1(A) \geq \cdots \geq \lambda_n(A)$.
Eigenvalue sets of $\boldsymbol{A}^S$ are compact intervals

$$\boldsymbol{\lambda}_i(\boldsymbol{A}^S) := \left\{\lambda_i(A) : A \in \boldsymbol{A}^S\right\}, \quad i = 1, \ldots, n.$$

Theorem

Checking whether $0 \in \boldsymbol{\lambda}_i(\boldsymbol{A}^S)$ for some $i = 1, \ldots, n$ is NP-hard.

Proof.

$\boldsymbol{A}$ is singular iff $\boldsymbol{M}^S := \begin{pmatrix} 0 & \boldsymbol{A} \\ \boldsymbol{A}^T & 0 \end{pmatrix}^S$ is singular (has a zero eigenvalue). □

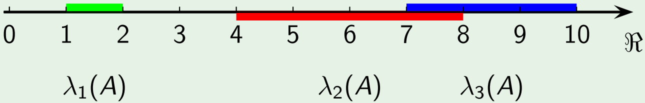
Eigenvalues – An Example

Example

Let

$$A \in \boldsymbol{A} = \begin{pmatrix} [1,2] & 0 & 0 \\ 0 & [7,8] & 0 \\ 0 & 0 & [4,10] \end{pmatrix}$$

What are the eigenvalue sets?
We have $\boldsymbol{\lambda}_1(\boldsymbol{A}^S) = [7,10]$, $\boldsymbol{\lambda}_2(\boldsymbol{A}^S) = [4,8]$ and $\boldsymbol{\lambda}_3(\boldsymbol{A}^S) = [1,2]$.



Eigenvalue sets are compact intervals. They may intersect or equal.

Eigenvalues – Some Exact Bounds

Theorem (Hertz, 1992)

We have

$$\overline{\lambda}_1(\boldsymbol{A}^S) = \max_{z\in\{\pm1\}^n} \lambda_1(A^c + \operatorname{diag}(z)A^\Delta \operatorname{diag}(z)),$$

$$\underline{\lambda}_n(\boldsymbol{A}^S) = \min_{z\in\{\pm1\}^n} \lambda_n(A^c - \operatorname{diag}(z)A^\Delta \operatorname{diag}(z)).$$

Proof.

"Upper bound." By contradiction suppose that there is $A \in \boldsymbol{A}^S$ such that

$$\lambda_1(A) > \max_{z\in\{\pm1\}^n} \lambda_1(A_z), \quad \left[\text{where } A_z \equiv A^c + \operatorname{diag}(z)A^\Delta \operatorname{diag}(z)\right]$$

Thus $Ax = \lambda_1(A)x$ for some x with $\|x\|_2 = 1$.
Put $z^* := \operatorname{sgn}(x)$, and by the Rayleigh–Ritz Theorem we have

$$\begin{aligned} \lambda_1(A) &= x^T A x \le x^T A_{z^*} x \\ &\le \max_{y:\|y\|_2=1} y^T A_{z^*} y = \lambda_1(A_{z^*}). \end{aligned}$$

□

Eigenvalues – Some Other Exact Bounds

Theorem

$\underline{\lambda}_1(\boldsymbol{A}^S)$ and $\overline{\lambda}_n(\boldsymbol{A}^S)$ are polynomially computable by semidefinite programming with arbitrary precision.

Proof.

We have

$$\overline{\lambda}_n(\boldsymbol{A}^S) = \max \alpha \ \text{ subject to } \ A - \alpha I_n \text{ is positive semidefinite}, \ A \in \boldsymbol{A}^S.$$

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 semidefinite programming problem reads

$$\overline{\lambda}_n(\boldsymbol{A}^S) = \max \alpha \ \text{ subject to } \ M(A, \alpha) \text{ is positive semidefinite.}$$

□

Eigenvalues – Enclosures

Theorem

We have

$$\boldsymbol{\lambda}_i(\boldsymbol{A}^S) \subseteq [\lambda_i(A^c) - \rho(A^\Delta), \lambda_i(A^c) + \rho(A^\Delta)], \quad i = 1, \ldots, n.$$

Proof.

Recall for any $A, B \in \mathbb{R}^{n \times n}$,

$$|A| \leq B \quad \Rightarrow \quad \rho(A) \leq \rho(|A|) \leq \rho(B),$$

and for A, B symmetric (Weyl's Theorem)

$$\lambda_i(A) + \lambda_n(B) \leq \lambda_i(A + B) \leq \lambda_i(A) + \lambda_1(B), \quad i = 1, \ldots, n.$$

Let $A \in \boldsymbol{A}^S$, so $|A - A^c| \leq A^\Delta$. Then

$$\begin{aligned}\lambda_i(A) = \lambda_i(A^c + (A - A^c)) &\leq \lambda_i(A^c) + \lambda_1(A - A^c)\\ &\leq \lambda_i(A^c) + \rho(|A - A^c|) \leq \lambda_i(A^c) + \rho(A^\Delta).\end{aligned}$$

Similarly for the lower bound. $\square$

Eigenvalues – Easy Cases

Theorem

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*
$$\overline{\lambda}_1(\boldsymbol{A}^S) = \lambda_1(\overline{A}), \quad \underline{\lambda}_n(\boldsymbol{A}^S) = \lambda_n(\underline{A}).$$

Proof.

1. For the sake of simplicity suppose $A^c \geq 0$. Then $\forall A \in \boldsymbol{A}^S$ we have $|A| \leq \overline{A}$, whence
$$\lambda_1(A) = \rho(A) \leq \rho(\overline{A}) = \lambda_1(\overline{A}).$$
2. By Hertz's theorem,
$$\overline{\lambda}_1(\boldsymbol{A}^S) = \max_{z \in \{\pm 1\}^n} \lambda_1(A^c + \operatorname{diag}(z) A^\Delta \operatorname{diag}(z)),$$
$$= \lambda_1(A^c + A^\Delta) = \lambda_1(\overline{A}).$$

$\square$

Positive Semidefiniteness

$\boldsymbol{A}^S$ is *positive semidefinite* if every $A \in \boldsymbol{A}^S$ is positive semidefinite.

Theorem

The following are equivalent

1. $\boldsymbol{A}^S$ *is positive semidefinite,*
2. $A_z \equiv A^c - \operatorname{diag}(z)A^\Delta \operatorname{diag}(z)$ *is positive semidefinite* $\forall z \in \{\pm 1\}^n$*,*
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 $A_z \in \boldsymbol{A}^S$.

"(2) $\Rightarrow$ (3)" Let $x \in \mathbb{R}^n$ and put $z := \operatorname{sgn}(x)$. Now,

$$x^T A^c x - |x|^T A^\Delta |x| = x^T A^c x - x^T \operatorname{diag}(z) A^\Delta \operatorname{diag}(z) x = x^T A_z x \geq 0.$$

"(3) $\Rightarrow$ (1)" Let $A \in \boldsymbol{A}^S$ and $x \in \mathbb{R}^n$. Now,

$$\begin{aligned} x^T A x = x^T A^c x + x^T (A - A^c) x &\geq x^T A^c x - |x^T (A - A^c) x| \\ &\geq x^T A^c x - |x|^T A^\Delta |x| \geq 0. \end{aligned}$$

□

Positive Definiteness

$\boldsymbol{A}^S$ is *positive definite* if every $A \in \boldsymbol{A}^S$ is positive definite.

Theorem

The following are equivalent

1. $\boldsymbol{A}^S$ *is positive definite,*
2. $A_z \equiv A^c - \operatorname{diag}(z)A^\Delta \operatorname{diag}(z)$ *is positive definite for each* $z \in \{\pm 1\}^n$*,*
3. $x^T A^c x - |x|^T A^\Delta |x| > 0$ *for each* $0 \neq x \in \mathbb{R}^n$*,*
4. A^c *is positive definite and* $\boldsymbol{A}$ *is regular.*

Proof.

"(1) $\Leftrightarrow$ (2) $\Leftrightarrow$ (3)" analogously.
"(1) $\Rightarrow$ (4)" If there are $A \in \boldsymbol{A}$ and $x \neq 0$ such that $Ax = 0$, then

$$0 = x^T A x = x^T \tfrac{1}{2}(A + A^T)x,$$

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 $\lambda_i(A^c) > 0\ \forall i$, and regularity of $\boldsymbol{A}$ implies $\boldsymbol{\lambda}_i(\boldsymbol{A}^S) > 0\ \forall i$. $\square$

Complexity

Theorem (Nemirovskii, 1993)

Checking positive semidefiniteness of $\boldsymbol{A}^S$ is co-NP-hard.

Theorem (Rohn, 1994)

Checking positive definiteness of $\boldsymbol{A}^S$ is co-NP-hard.

Theorem (Jaulin and Henrion, 2005)

Checking whether there is a positive definite matrix in $\boldsymbol{A}^S$ is a polynomial time problem.

Proof.

There is a positive definite matrix in $\boldsymbol{A}^S$ iff $\overline{\lambda}_n(\boldsymbol{A}^S) > 0$.
So we can check it by semidefinite programming. $\square$

Sufficient Conditions

Theorem

1. $\boldsymbol{A}^S$ *is positive semidefinite if* $\lambda_n(A^c) \geq \rho(A^\Delta)$.
2. $\boldsymbol{A}^S$ *is positive definite if* $\lambda_n(A^c) > \rho(A^\Delta)$.
3. $\boldsymbol{A}^S$ *is positive definite if* A^c *is positive definite and* $\rho(|(A^c)^{-1}|A^\Delta) < 1$.

Proof.

1. $\boldsymbol{A}^S$ is positive semidefinite iff $\underline{\lambda}_n(\boldsymbol{A}^S) \geq 0$.
   Now, employ the smallest eigenvalue set enclosure
   $$\boldsymbol{\lambda}_n(\boldsymbol{A}^S) \subseteq [\lambda_n(A^c) - \rho(A^\Delta), \lambda_n(A^c) + \rho(A^\Delta)].$$
2. Analogous.
3. Use Beeck's sufficient condition for regularity of $\boldsymbol{A}$. $\square$

Application: Convexity Testing

Theorem

A function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is convex on $\boldsymbol{x} \in \mathbb{IR}^n$ iff its Hessian $\nabla^2 f(x)$ is positive semidefinite $\forall x \in$ int $\boldsymbol{x}$.

Corollary

A function $f : \mathbb{R}^n \mapsto \mathbb{R}$ is convex on $\boldsymbol{x} \in \mathbb{IR}^n$ if $\nabla^2 f(\boldsymbol{x})$ is positive semidefinite.

Application: Convexity Testing

Example

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 reads

$$\nabla^2 f(x, y, z) = \begin{pmatrix} 6x + 4y & 4x - z & -y \\ 4x - z & 16 & -x + 6z \\ -y & -x + 6z & 6y \end{pmatrix}$$

Evaluation the Hessian matrix by interval arithmetic results in

$$\nabla^2 f(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}) \subseteq \begin{pmatrix} [16, 26] & [7, 12] & -[1, 2] \\ [7, 12] & 16 & [-3, 4] \\ -[1, 2] & [-3, 4] & [6, 12] \end{pmatrix}$$

Now, both sufficient conditions for positive definiteness succeed. Thus, we can conclude that f si convex on the interval domain.

References

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.

M. Hladík, D. Daney, and E. P. Tsigaridas.
Characterizing and approximating eigenvalue sets of symmetric interval matrices.
Comput. Math. Appl., 62(8):3152–3163, 2011.

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.

J. Rohn.
Positive definiteness and stability of interval matrices.
SIAM J. Matrix Anal. Appl., 15(1):175–184, 1994.

J. Rohn.
A handbook of results on interval linear problems.
Tech. Rep. 1163, Acad. of Sci. of the Czech Republic, Prague, 2012.
http://uivtx.cs.cas.cz/~rohn/publist/!aahandbook.pdf

Next Section

Conclusion

Interval computation offers:

- nice theory, methods and applications
- many open problems
- interdisciplinarity