Linear Interval Parametric Approach to Testing Pseudoconvexity

Milan Hladík* Lubomir V. Kolev Iwona Skalna

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

LeGO 2018
XIV Global Optimization Workshop, Leiden
September 18–21, 2018

Pseudoconvexity

Motivation

Convexity has many nice properties in the context of optimization.
What about its generalizations?

Definition

Let $f : \mathbb{R}^n \to \mathbb{R}$ be twice differentiable and $\mathcal{S} \subset \mathbb{R}^n$ an open convex set. Then $f(x)$ is *pseudoconvex* on $\mathcal{S}$ if for every $x, y \in \mathcal{S}$ we have

$$\nabla f(x)^T (y - x) \geq 0 \quad \Rightarrow \quad f(y) \geq f(x).$$

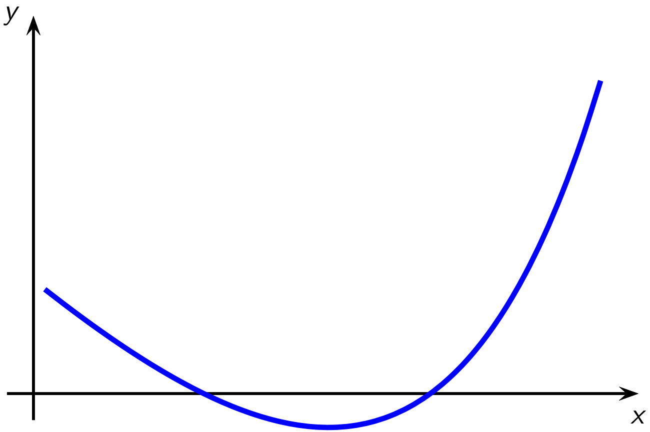
Key Properties

Minimizing pseudoconvex objective functions on convex feasible sets,

- each stationary point is a global minimum,
- each local minimum is a global minimum,
- the optimal solution set is convex.

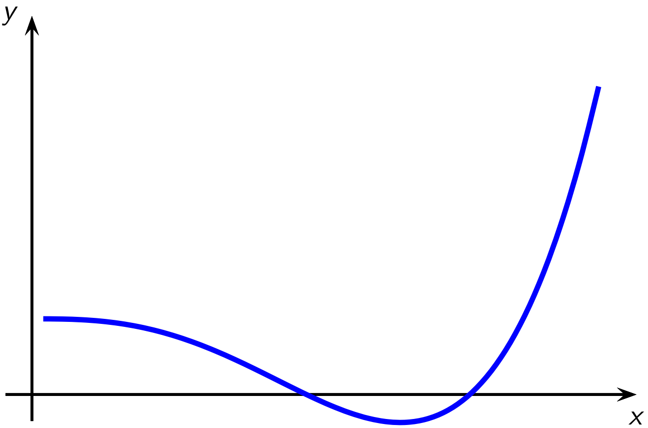
Illustration

Convex function



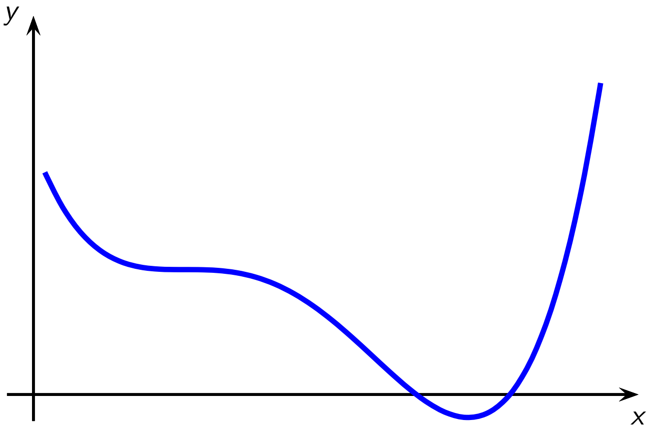
Illustration

Pseudoconvex function



Illustration

Quasiconvex function



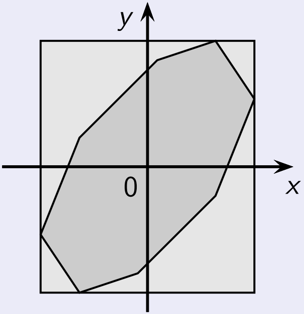
Problem Formulation

Problem formulation

Given an affine form

$$\boldsymbol{x}(\boldsymbol{p}) := \left\{ \sum_{k=1}^{K} x^{(k)} p_k + x_c,\ p \in \boldsymbol{p} \right\}.$$

Geometrically it is a zonotope



The question

Is a differentiable function $f : \mathbb{R}^n \to \mathbb{R}$ pseudoconvex on $\boldsymbol{x}(\boldsymbol{p})$?

Problem Formulation

Theorem (Ahmadi et al., 2013)

Deciding pseudoconvexity is NP-hard on a class of quartic polynomials.

Aim

Therefore we will be content with cheap sufficient conditions.

Technical Tools

Interval analysis

Interval arithmetic:

$$
\begin{aligned}
\boldsymbol{x}+\boldsymbol{y} &= [\underline{x}+\underline{y}, \overline{x}+\overline{y}],\\
\boldsymbol{x}-\boldsymbol{y} &= [\underline{x}-\overline{y}, \overline{x}-\underline{y}],\\
\boldsymbol{x}\boldsymbol{y} &= [\min(\underline{x}\underline{y}, \underline{x}\overline{y}, \overline{x}\underline{y}, \overline{x}\overline{y}), \max(\underline{x}\underline{y}, \underline{x}\overline{y}, \overline{x}\underline{y}, \overline{x}\overline{y})],\\
\boldsymbol{x}/\boldsymbol{y} &= [\min(\underline{x}/\underline{y}, \underline{x}/\overline{y}, \overline{x}/\underline{y}, \overline{x}/\overline{y}), \max(\underline{x}/\underline{y}, \underline{x}/\overline{y}, \overline{x}/\underline{y}, \overline{x}/\overline{y})], \quad 0 \notin \boldsymbol{y}.
\end{aligned}
$$

Evaluation of functions and their derivatives,. . .

For interval matrix $\boldsymbol{A}$: regularity, eigenvalues, det, positive semidef., . . .

Technical Tools

Affine arithmetic (reduced/revised version)

Given two affine forms

$$\boldsymbol{x}(p) := \textstyle\sum_{k=1}^{K} x_k p_k + \boldsymbol{x}_0 = x^T p + \boldsymbol{x}_0,$$
$$\boldsymbol{y}(p) := \textstyle\sum_{k=1}^{K} y_k p_k + \boldsymbol{y}_0 = y^T p + \boldsymbol{y}_0,$$

where $p \in \boldsymbol{p}$. For any $\alpha, \beta \in \mathbb{R}$ we have

$$\boldsymbol{x}(p) + \boldsymbol{y}(p) = (x + y)^T p + (\boldsymbol{x}_0 + \boldsymbol{y}_0),$$
$$\alpha \boldsymbol{x}(p) = (\alpha x)^T p + (\alpha \boldsymbol{x}_0).$$

Nonlinear operations have to be approximated. Multiplication usually reads

$$\boldsymbol{x}(p) \cdot \boldsymbol{y}(p) = ((y_0)_c x + (x_0)_c y)^T p + \boldsymbol{z},$$

where $\boldsymbol{z} = [z_c - z_\Delta, z_c + z_\Delta]$ encloses the accumulative error with

$$z_c = x_c y_c + \tfrac{1}{2} x^T y,$$
$$z_\Delta = |x_c| y_\Delta + |y_c| x_\Delta + \left(|x|^T e + x_\Delta\right)\left(|y|^T e + y_\Delta\right) - \tfrac{1}{2} |x|^T |y|.$$

Method Based on Mereau and Paquet

Theorem (Mereau and Paquet, 1974)

The function $f(x)$ is pseudoconvex on set $\mathcal{S}$ if there is $\alpha \geq 0$ such that

$$\nabla^2 f(x) + \alpha \nabla f(x) \nabla f(x)^T \tag{$\star$}$$

is positive semidefinite for all $x \in \mathcal{S}$.

Proposition

We have that $(\star)$ is positive semidefinite if and only if

$$\begin{pmatrix} -\frac{1}{\alpha} & \nabla f(x)^T \\ \nabla f(x) & \nabla^2 f(x) \end{pmatrix}$$

has at most one simple negative eigenvalue.

Method Based on Mereau and Paquet

Theorem (Mereau and Paquet, 1974)

The function $f(x)$ is pseudoconvex on set $\mathcal{S}$ if there is $\alpha \geq 0$ such that

$$\nabla^2 f(x) + \alpha \nabla f(x) \nabla f(x)^T \tag{$\star$}$$

is positive semidefinite for all $x \in \mathcal{S}$.

Method

- Enclose $\nabla f(\boldsymbol{x}(\boldsymbol{p})) \subseteq \boldsymbol{g}(\boldsymbol{p})$, $\nabla^2 f(\boldsymbol{x}(\boldsymbol{p})) \subseteq \boldsymbol{H}(\boldsymbol{p})$
- Denote

$$\boldsymbol{D}'(\boldsymbol{p}) := \begin{pmatrix} -\frac{1}{\alpha} & \boldsymbol{g}(\boldsymbol{p})^T \\ \boldsymbol{g}(\boldsymbol{p}) & \boldsymbol{H}(\boldsymbol{p}) \end{pmatrix} \supseteq \begin{pmatrix} -\frac{1}{\alpha} & \nabla f(x)^T \\ \nabla f(x) & \nabla^2 f(x) \end{pmatrix}$$

- Check that the second smallest eigenvalue of the matrices $\boldsymbol{D}'(\boldsymbol{p})$ stays nonnegative.

Method Based on Crouzeix

Theorem (Crouzeix, 1998)

Function $f(x)$ is pseudoconvex on $\mathcal{S}$ if for each $x \in \mathcal{S}$ and every $y \neq 0$ such that $\nabla f(x)^T y = 0$ we have $y^T \nabla^2 f(x) y > 0$.

Equivalently, by Crouzeix (1998),

$$D(x) := \begin{pmatrix} 0 & \nabla f(x)^T \\ \nabla f(x) & \nabla^2 f(x) \end{pmatrix}.$$

has n positive eigenvalues on $\mathcal{S}$.

Method

- Compute
$$\boldsymbol{D}(\boldsymbol{p}) := \begin{pmatrix} 0 & \boldsymbol{g}(\boldsymbol{p})^T \\ \boldsymbol{g}(\boldsymbol{p}) & \boldsymbol{H}(\boldsymbol{p}) \end{pmatrix} \supseteq \begin{pmatrix} 0 & \nabla f(x)^T \\ \nabla f(x) & \nabla^2 f(x) \end{pmatrix}$$
- Compute an enclosure $\boldsymbol{\lambda}_2$ for the second smallest eigenvalue of $\boldsymbol{D}(\boldsymbol{p})$ and check that $\underline{\lambda}_2 > 0$.

Method Based on Crouzeix and Ferland

Theorem (Crouzeix and Ferland, 1982)

Function $f(x)$ is pseudoconvex on $\mathcal{S}$ if for each $x \in \mathcal{S}$ either $\nabla^2 f(x)$ is positive semidefinite, or $\nabla^2 f(x)$ has one simple negative eigenvalue and there is $b \in \mathbb{R}^n$ such that $\nabla^2 f(x) b = \nabla f(x)$ and $\nabla f(x)^T b < 0$.

Preliminaries

- Enclose $\nabla f(\boldsymbol{x}(\boldsymbol{p})) \subseteq \boldsymbol{g}(\boldsymbol{p})$, $\nabla^2 f(\boldsymbol{x}(\boldsymbol{p})) \subseteq \boldsymbol{H}(\boldsymbol{p})$
- Condition

$$\exists b : \;\; Hb = g, \;\; g^T b < 0$$

is equivalent to $g^T H^{-1} g < 0$ for each $g \in \boldsymbol{g}(\boldsymbol{p})$ and $H \in \boldsymbol{H}(\boldsymbol{p})$.

The method checks that

– every matrix in $\boldsymbol{H}(\boldsymbol{p})$ has at most one simple negative eigenvalue,

– we have $g^T H^{-1} g < 0$ for every $g \in \boldsymbol{g}(p)$, $H \in \boldsymbol{H}(p)$ and $p \in \boldsymbol{p}$.

Method Based on Crouzeix and Ferland

The method checks that

– every matrix in $\boldsymbol{H}(\boldsymbol{p})$ has at most one simple negative eigenvalue, (1)

– we have $g^T H^{-1} g < 0$ for every $g \in \boldsymbol{g}(p)$, $H \in \boldsymbol{H}(p)$ and $p \in \boldsymbol{p}$. (2)

Proposition

- *It is NP-hard to check* (1)*.*
- *It is NP-hard to check* (2) *even for functions of type* $f(x) = \frac{1}{2}x^T Ax$*.*

Two tests for (1)

- Compute an enclosing interval $\boldsymbol{\lambda}_2$ for the second smallest eigenvalue of the matrices in $\boldsymbol{H}(\boldsymbol{p})$ and then check whether $\underline{\lambda}_2 \geq 0$.
- Check for $\lambda_2(H_c) > 0$ and regularity of $\boldsymbol{H}(\boldsymbol{p})$.

Method Based on Crouzeix and Ferland

The method checks that

– every matrix in $\boldsymbol{H}(\boldsymbol{p})$ has at most one simple negative eigenvalue, (1)

– we have $g^T H^{-1} g < 0$ for every $g \in \boldsymbol{g}(p)$, $H \in \boldsymbol{H}(p)$ and $p \in \boldsymbol{p}$. (2)

Proposition

- *It is NP-hard to check* (1)*.*
- *It is NP-hard to check* (2) *even for functions of type* $f(x) = \frac{1}{2}x^T Ax$*.*

Two tests for (2)

- Let $\boldsymbol{y}(\boldsymbol{p})$ be an affine form enclosure of the solution set of $\boldsymbol{H}(\boldsymbol{p})y = \boldsymbol{g}(\boldsymbol{p})$. Check that $\boldsymbol{g}(p)^T \boldsymbol{y}(p) < 0$.
- Check that $\det(D_c) < 0$ and $\boldsymbol{D}(\boldsymbol{p})$ is regular.

Numerical Experiments

Example (Random choices of $\boldsymbol{H}(\boldsymbol{p})$ and $\boldsymbol{g}(\boldsymbol{p})$)

$n =$ dimension, $d =$ radius of intervals, $K =$ number of parameters.

Success rate (in %):

			standard interval approach					parametric approach						
n	K	d	MP1	MP2	F	CF	C	MPa	MPb	CF1a	CF1b	CF2a	CF2b	C
5	5	0.1	21.0	35.0	**36.0**	**36.0**	**36.0**	36.0	**40.0**	0.0	0.0	**40.0**	18.0	36.0
10	5	0.1	7.0	30.0	**37.0**	**37.0**	27.0	27.0	**58.0**	0.0	0.0	53.0	28.0	27.0
5	5	0.3	0.0	18.0	**30.0**	**30.0**	20.0	20.0	**39.0**	0.0	0.0	**39.0**	14.0	20.0
10	5	0.3	0.0	5.0	9.0	**10.0**	3.0	3.0	**52.0**	0.0	0.0	39.0	20.0	3.0
5	10	0.1	7.0	18.0	18.0	**24.0**	19.0	19.0	**30.0**	0.0	0.0	**30.0**	9.0	19.0
10	10	0.1	0.0	19.0	**24.0**	**24.0**	13.0	13.0	**61.0**	0.0	0.0	51.0	29.0	13.0
15	10	0.1	0.0	5.0	**7.0**	**7.0**	2.0	2.0	**61.0**	0.0	0.0	36.0	23.0	2.0
5	10	0.3	0.0	4.0	12.0	**18.0**	8.0	8.0	**31.0**	0.0	0.0	**31.0**	14.0	8.0
10	10	0.3	0.0	0.0	0.0	0.0	**1.0**	1.0	**20.0**	0.0	0.0	11.0	4.0	1.0
15	10	0.3	0.0	0.0	0.0	0.0	0.0	0.0	**8.0**	0.0	0.0	2.0	1.0	0.0
5	10	0.5	0.0	0.0	0.0	**4.0**	0.0	0.0	**13.0**	0.0	0.0	**13.0**	4.0	0.0
10	10	0.5	0.0	0.0	0.0	0.0	0.0	0.0	**1.0**	0.0	0.0	**1.0**	**1.0**	0.0

Parametric methods approx. 7 times slower.

Numerical Experiments

Example (Benchmark data (success rate in %))

function	n	div	conv	standard interval approach MP1	MP2	F	CF	C	parametric approach MPa	MPb	CF1a	CF1b	CF2a	CF2b	C
mhw4d	5	5	0.0	0	0	0	0	0	0	**100**	0	0	0	**100**	0
mhw4d	5	8	0.0	0	0	0	0	0	0	**100**	0	0	**100**	**100**	0
Colville	4	5	0.0	0	0	0	0	0	0	0	0	0	0	0	0
Colville	4	10	0.0	0	0	0.2	**0.3**	0	0	**6.6**	0	0	0	**6.6**	0
Rosenbrock	4	5	0.0	0	0	0.2	**0.3**	0	0	0	0	0	0	0	0
Rosenbrock	4	10	0.0	0	0	0.2	**0.3**	0	0	**13.4**	0	0	6.8	**13.4**	0
Rosenbrock	4	15	0.0	0	0	**2.3**	1.8	0	0	**35.5**	0	0	19.6	35.5	0
G&P	2	10	0.0	0	0	0	0	0	0	0	0	0	0	0	0
G&P	2	50	0.4	0	0	0	0	**0.1**	1.9	**2.8**	0	0	**2.8**	2.6	1.9
G&P	2	100	1.4	0	0	0	0	**1.1**	3.4	**7.1**	0	0	**7.1**	4.7	3.4
f5_Messine	3	10	0.0	0	0	0	0	0	0	0	0	0	0	0	0
f5_Messine	3	30	0.0	0	0	0.1	0.2	**1.5**	**3.9**	15.7	10.0	13.9	10.5	**18.8**	4.0
f5_Messine	3	40	0.0	0	0	3.2	4.4	**4.6**	**6.7**	17.9	14.2	20.6	14.3	**24.8**	7.2
6hum_camel	2	10	0.0	0	0	0	0	0	0	0	0	0	0	0	0
6hum_camel	2	50	31.1	4.1	11.3	32.1	**36.9**	26.6	35.0	**56.2**	2.1	2.1	**56.2**	53.2	35.0
6hum_camel	2	100	43.1	10.9	16.7	45.2	**50.8**	40.9	48.6	**62.2**	5.2	5.2	**62.2**	60.1	48.6

Conclusion

Summary

- Compared to direct interval approach, parametric methods are significantly more efficient, but slightly slower.

Next steps?

- Pseudoconvex envelopes instead of convex envelopes?
- Some variation of the αBB method with pseudoconvex functions?

References

J. Crouzeix and J. A. Ferland.
Criteria for quasi-convexity and pseudo-convexity: Relationships and comparisons.
Math. Program., 23(1):193–205, 1982.

J.-P. Crouzeix.
Characterizations of Generalized Convexity and Generalized Monotonicity, A Survey.
In J.-P. Crouzeix et al., eds., *Generalized Convexity, Generalized Monotonicity: Recent Results*, pages 237–256. Springer, 1998.

J. A. Ferland.
Mathematical programming problems with quasi-convex objective functions.
Math. Program., 3(1):296–301, 1972.

P. Mereau and J.-G. Paquet.
Second order conditions for pseudo-convex functions.
SIAM J. Appl. Math., 27:131–137, 1974.