Sufficient Conditions for Pseudoconvexity by Using Linear Interval Parametric Techniques

Milan Hladík[1,a)], Lubomir Kolev[2,b)] and Iwona Skalna[3,c)]

[1] *Department of Applied Mathematics, Charles University, Malostranské nám. 25, 118 00, Prague, Czech Republic.*
[2] *Department of Theoretical Electrotechnics, Technical University of Sofia, Bulgaria.*
[3] *AGH University of Science and Technology, Poland.*

[a)] Corresponding author: milan.hladik@matfyz.cz
[b)] lkolev@tu-sofia.bg
[c)] skalna@agh.edu.pl

Abstract. The recent paper (DOI: 10.1007/s10898-017-0537-6) suggests various practical tests (sufficient conditions) for checking pseudoconvexity of a twice differentiable function on an interval domain. The tests were implemented using interval extensions of the gradient and the Hessian of the function considered. In this paper, we present an alternative approach which is based on the use of linear interval parametric enclosures of the gradient and the Hessian. It is shown that the new approach results in more efficient tests for checking pseudoconvexity.

INTRODUCTION

Pseudoconvexity of real functions is a generalized concept of convexity. Pseudoconvex objective functions have some nice properties in the context of optimization: On the convex feasible set, each stationary point is a global minimum point, each local minimum is a global minimum, and the optimal solution set is convex [1].

Recently, Hladík [2] proposed several tests for checking pseudoconvexity of such functions on interval domains based on various second-order characterizations of pseudoconvexity. The tests were implemented using interval extensions of the gradient and Hessian. The extensions were obtained using interval arithmetic.

In this paper, an alternative approach is suggested that is based on the use of the so-called affine (linear interval parametric) form of nonlinear functions [3, 4]. Accordingly, the gradients and Hessians are treated in the affine forms, too. We modify the pseudoconvexity tests such that they utilize the affine form structure more effectively.

Interval notation. An interval matrix is defined as

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

where $\underline{A}$ and $\overline{A}$, $\underline{A} \le \overline{A}$, are given matrices. Throughout the paper, inequalities and absolute values applied on vectors and matrices are understood entrywise. The midpoint and radius matrices corresponding to $\boldsymbol{A}$ are defined as $A_c := \frac{1}{2}(\underline{A} + \overline{A})$, and $A_\Delta := \frac{1}{2}(\overline{A} - \underline{A})$, respectively. The set of all interval $m \times n$ matrices is denoted by $\mathbb{IR}^{m\times n}$. Interval vectors are considered as one-column interval matrices. For an overview on interval arithmetic we refer the readers, e.g., to [5].

Affine forms. An affine form matrix generalizes the concept of an interval matrix. It is a more subtle parametric matrix enabling to represent linear dependencies and arithmetic operations. A parametric matrix in an affine form reads

$$\boldsymbol{A}(p) = \sum_{k=1}^{K} A^{(k)} p_k + \boldsymbol{A}, \quad p \in \boldsymbol{p},$$

where $A^{(1)}, \dots, A^{(K)} \in \mathbb{R}^{m\times n}$, $\boldsymbol{A} \in \mathbb{IR}^{m\times n}$ are known, and $p_1, \dots, p_K$ are parameters varying within given intervals $\boldsymbol{p}_1, \dots, \boldsymbol{p}_K$. The first term represents the linear structure, and the second term $\boldsymbol{A}$ represents the accumulative error resulting from approximation of non-affine operations. Without loss of generality we may assume that $\boldsymbol{p}_1 = \dots = \boldsymbol{p}_K = [-1, 1]$, which can be achieved by suitable scaling and shifting. The affine form represents the set of values

$$\boldsymbol{A}(\boldsymbol{p}) = \left\{\textstyle\sum_{k=1}^K A^{(k)} p_k + A;\ A \in \boldsymbol{A},\ p \in \boldsymbol{p}\right\}.$$

The smallest interval matrix enclosing $\boldsymbol{A}(\boldsymbol{p})$ is that obtained by interval evaluation

$$\boldsymbol{B} := [\boldsymbol{A}(\boldsymbol{p})] = \sum_{k=1}^{K} A^{(k)}[-1, 1] + \boldsymbol{A},$$

where $B_c = A_c$ and $B_\Delta = \sum_{k=1}^K |A^{(k)}| + A_\Delta$. Affine arithmetic is described, e.g., in [6, 7]. There are several variants known, we will employ the so called *revised affine form* arithmetic (see, e.g., [4, 7]).

Characterizations of pseudoconvexity. Let $f\colon \mathbb{R}^n \to \mathbb{R}$ 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

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

We recall some known [8, 9, 10, 1, 11] second order characterizations of pseudoconvexity that we will utilize for checking pseudoconvexity. The statements below are formulated for checking pseudoconvexity on convex set $\mathcal{S}$ with nonempty interior.

Theorem 1 (Mereau and Paquet, 1974)**.** *Function $f(x)$ is pseudoconvex on $\mathcal{S}$ if there exists $\alpha \geq 0$ such that augmented Hessian*

$$\nabla^2 f(x) + \alpha \nabla f(x) \nabla f(x)^T$$

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

Theorem 2 (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 $\nabla f(x)^T b < 0$.*

Theorem 3 (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 $y^T \nabla^2 f(x) y > 0$.*

Problem formulation. Throughout this paper,

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

is an affine form with nonempty interior. It is assumed that at the beginning accumulative error $x_\Delta = 0$, but this assumption does not restrict the generality of the presented results. Further, it is assumed that $f\colon \mathbb{R}^n \to \mathbb{R}$ is twice differentiable on an open set containing $\boldsymbol{x}(\boldsymbol{p})$. We address the question whether $f(x)$ is pseudoconvex on $\boldsymbol{x}(\boldsymbol{p})$, and we present several efficient sufficient conditions.

Preliminaries. Let $\boldsymbol{H}(\boldsymbol{p})$ and $\boldsymbol{g}(\boldsymbol{p})$ be affine enclosures of the Hessian matrices and gradients of $f(x)$ on $\boldsymbol{x}(\boldsymbol{p})$. That is,

$$H = \nabla^2 f(x(p)) \in \boldsymbol{H}(p) \quad \forall p \in \boldsymbol{p},$$
$$g = \nabla f(x(p)) \in \boldsymbol{g}(p) \quad \forall p \in \boldsymbol{p}.$$

Such affine interval enclosures can be computed by (revised) affine arithmetic using automatic or symbolic differentiation. We will also employ the so called border Hessian matrix

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

Its affine form enclosure then reads

$$\boldsymbol{D}(\boldsymbol{p}) := \begin{pmatrix} 0 & \boldsymbol{g}(\boldsymbol{p})^T \\ \boldsymbol{g}(\boldsymbol{p}) & \boldsymbol{H}(\boldsymbol{p}) \end{pmatrix} = \left\{ \begin{pmatrix} 0 & g^T \\ g & H \end{pmatrix};\ g \in \boldsymbol{g}(p),\ H \in \boldsymbol{H}(p),\ p \in \boldsymbol{p} \right\}.$$

AN INTERVAL PARAMETRIC APPROACH TO CHECKING PSEUDOCONVEXITY

Method based on Mereau and Paquet

By Theorem 1, we have to check positive semidefiniteness of

$$\nabla^2 f(x) + \alpha \nabla f(x) \nabla f(x)^T \tag{1}$$

with $x \in \boldsymbol{x}(\boldsymbol{p})$ and for some $\alpha > 0$. Computationally, this form is not so convenient; see [2]. Thus we rewrite it using a modified border Hessian matrix

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

Proposition 1. *We have that* (1) *is positive semidefinite if and only if $D'(x)$ has at most one simple negative eigenvalue.*

Notice that in the definition of $D'(x)$, we can replace the top left value of $-\frac{1}{\alpha}$ by any other negative constant.

Method MP. The resulting algorithm is as follows. Let $\boldsymbol{D'}(\boldsymbol{p})$ be a parametric form enclosure of the modified border Hessian matrices. Check that the second smallest eigenvalue of the matrices in $\boldsymbol{D'}(\boldsymbol{p})$ stays nonnegative.

Methods based on Crouzeix and Ferland

The condition from Theorem 2 that there is b such that $Hb = g$, $g^T b < 0$ is equivalent to $g^T H^{-1} g < 0$ for $g \in \boldsymbol{g}(\boldsymbol{p})$ and $H \in \boldsymbol{H}(\boldsymbol{p})$. In order to utilize Theorem 2, it is sufficient to check that

$$\text{– every matrix in } \boldsymbol{H}(\boldsymbol{p}) \text{ has at most one simple negative eigenvalue,} \tag{2}$$

$$\text{– we have } g^T H^{-1} g < 0 \text{ for every } g \in \boldsymbol{g}(p),\ H \in \boldsymbol{H}(p) \text{ and } p \in \boldsymbol{p}. \tag{3}$$

We show that each of the above conditions is hard in general. The first one is hard even for standard interval matrices, and the second one is hard even when restricted to quadratic polynomials.

Proposition 2. *It is NP-hard to check whether every matrix in $\boldsymbol{H}$ has at most one simple negative eigenvalue.*

Proposition 3. *Let $\boldsymbol{x} \in \mathbb{IR}^n$. Let $f(x) = \frac{1}{2} x^T A x$ with $A \in \mathbb{R}^{n \times n}$ nonsingular and having one simple negative eigenvalue. Then checking the condition $\nabla f(x)^T (\nabla^2 f(x))^{-1} \nabla f(x) < 0$ for every $x \in \boldsymbol{x}$ is NP-hard.*

Corollary 1. *Checking pseudoconvexity on $\boldsymbol{x} \in \mathbb{IR}^n$ is NP-hard for functions of the form $f(x) = \frac{1}{2} x^T A x$ with $A \in \mathbb{R}^{n \times n}$ nonsingular and having one simple negative eigenvalue.*

Method CF1. The method is based on verifying conditions (2) and (3). For the purpose of checking for (3), we consider the interval parametric system $\boldsymbol{H}(\boldsymbol{p}) y = \boldsymbol{g}(\boldsymbol{p})$. Its solution set is defined as

$$\{y \in \mathbb{R}^n;\ \exists p \in \boldsymbol{p},\ \exists H \in \boldsymbol{H}(p),\ \exists g \in \boldsymbol{g}(p) : Hy = g\}.$$

There are various methods for computing an interval enclosure of this solution set. However, we employ a solver computing an enclosure in the affine form [12, 4], so the subsequent evaluation will be more effective than by using simple interval enclosures. Now, let $\boldsymbol{y}(\boldsymbol{p})$ be an affine form enclosure of the solution set, and calculate $\boldsymbol{s}(p) := \boldsymbol{g}(p)^T \boldsymbol{y}(p)$. Eventually, we check that $\boldsymbol{s}(p) < 0$ for every $p \in \boldsymbol{p}$ simply by checking that the right end-point of the interval $[\boldsymbol{s}(\boldsymbol{p})]$ is negative.

Method CF2. As in Hladík [2], the condition (3) can be equivalently expressed as $\det(D) < 0$ for every $D \in \boldsymbol{D}(\boldsymbol{p})$. The determinant is negative for every realization if it is negative for one realization (e.g., the midpoint one) and $\boldsymbol{D}(\boldsymbol{p})$ is regular (i.e., it contains nonsingular matrices only). This gives us an alternative method to CF1.

Method based on Crouzeix

By Crouzeix [10], Theorem 3 can be equivalently formulated as that $D(x)$ has n positive eigenvalues. This leads to the following algorithm.

Method C. Let $\boldsymbol{D}(\boldsymbol{p})$ be a parametric form enclosure of the border Hessian matrices. Compute an enclosure $\boldsymbol{\lambda}$ for the second smallest eigenvalue of $\boldsymbol{D}(\boldsymbol{p})$. Eventually, verify that $\underline{\lambda} > 0$.

Provided $0 \notin \nabla f(x)$ for every feasible $x \in \boldsymbol{x}(\boldsymbol{p})$, that is, $0 \notin \boldsymbol{g}(\boldsymbol{p})$, then the above test can be slightly weakened and check only for $\underline{\lambda} \geq 0$; see Crouzeix [10]. Notice that the condition $0 \notin \boldsymbol{g}(\boldsymbol{p})$ can be checked by linear programming since $0 \in \boldsymbol{g}(\boldsymbol{p})$ takes the form of linear constraints in variables $p_1, \ldots, p_K$

$$\textstyle\sum_{k=1}^{K} g^{(k)} p_k + \underline{g} \leq 0 \leq \sum_{k=1}^{K} g^{(k)} p_k + \overline{g}, \quad p_k \in \boldsymbol{p}_k, \ k = 1, \ldots, K.$$

As a faster sufficient condition, we can use interval evaluation and check for $0 \notin [\boldsymbol{g}(\boldsymbol{p})]$.

CONCLUSION

We presented several methods for checking pseudoconvexity of a general differentiable function on an affine interval domain, that is, on a zonotope. To obtain more efficient methods, we employed affine forms and affine arithmetic and adapted the tests accordingly to utilize affine forms. We intend to carry out thorough numerical comparisons to see whether the novel parametric methods are more efficient than those based on standard interval arithmetic.

ACKNOWLEDGMENTS

M. Hladík was supported by Czech Science Foundation project CE-ITI (P202/12/G061).

REFERENCES

[1] N. Hadjisavvas, S. Komlósi, and S. Schaible, eds., *Handbook of generalized convexity and generalized monotonicity* (Springer, New York, 2005).
[2] M. Hladík, J. Glob. Optim. **71**, 443–455 (2018).
[3] L. Kolev and I. Skalna, Numer. Algorithms (2018), 10.1007/s11075-017-0418-6, in press.
[4] I. Skalna and M. Hladík, BIT Numer. Math. **57**, 1109–1136 (2017).
[5] R. E. Moore, R. B. Kearfott, and M. J. Cloud, *Introduction to Interval Analysis* (SIAM, 2009).
[6] L. de Figueiredo and J. Stolfi, *Self-Validated Numerical Methods and Applications*, Brazilian Mathematics Colloquium monograph (IMPA, Rio de Janeiro, Brazil, 1997).
[7] I. Skalna, *Parametric Interval Algebraic Systems* (Springer, 2018).
[8] J. Crouzeix, Math. Program. **18**, 349–352 (1980).
[9] J. Crouzeix and J. A. Ferland, Math. Program. **23**, 193–205 (1982).
[10] J.-P. Crouzeix, in *Generalized Convexity, Generalized Monotonicity: Recent Results*, edited by J.-P. Crouzeix, J.-E. Martinez-Legaz, and M. Volle (Springer, 1998), pp. 237–256.
[11] P. Mereau and J.-G. Paquet, SIAM J. Appl. Math. **27**, 131–137 (1974).
[12] L. V. Kolev, Appl. Math. Comput. **246**, 229–246 (2014).