Recognizing Pseudoconvexity of a Function on an Interval Domain

Milan Hladík

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

ODAM 2017
Olomoucian Days of Applied Mathematics, Olomouc
May 31 – June 2, 2017
Pseudoconvexity

Motivation
Con vexity 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 $S \subset \mathbb{R}^n$ an open convex set. Then $f(x)$ is pseudoconvex on $S$ if for every $x, y \in S$ we have

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

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.
Convex function
Pseudoconvex function
Quasiconvex function
Pseudoconvexity

Problem Formulation

Given a box $\mathbf{x} = [\underline{x}, \overline{x}]$ in $\mathbb{R}^n$ and differentiable $f : \mathbb{R}^n \rightarrow \mathbb{R}$. The question: Is $f(\mathbf{x})$ pseudoconvex on $\mathbf{x}$?

Why testing pseudoconvexity on a box?

In global optimization, feasible sub-domains have often the form of boxes. Verifying pseudoconvexity can help to process a given box (for example, by local search).

Theorem (Ahmadi et al., 2013)

Deciding pseudoconvexity is NP-hard on a class of quartic polynomials.
### Theorem (Mereau and Paquet, 1974)

The function $f(x)$ is pseudoconvex on $x$ if there is $\alpha \geq 0$ such that

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

is positive semidefinite for all $x \in x$.

Denote

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

and by $D(x)_r$ we denote the principal leading submatrix of size $r$.

### Theorem (Ferland, 1972)

The function $f(x)$ is pseudoconvex on $x$ if $\det(D(x)_r) < 0$ for every $r = 2, \ldots, n + 1$ and for all $x \in x$. 
Pseudoconvexity Characterizations

**Theorem (Crouzeix and Ferland, 1982)**

The function \( f(x) \) is pseudoconvex on \( x \) if for each \( x \in x \) 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 (Crouzeix, 1998)**

The function \( f(x) \) is pseudoconvex on \( x \) if for each \( x \in x \) the matrix \( D(x) \) is nonsingular and has exactly one simple negative eigenvalue.

**Theorem (Crouzeix, 1998)**

The function \( f(x) \) is pseudoconvex on \( x \) if for each \( x \in x \) and every \( y \neq 0 \) such that \( \nabla f(x)^T y = 0 \) we have \( y^T \nabla^2 f(x)y > 0 \).
Interval Methods for Testing Pseudoconvexity

Interval Enclosures

Let $H \in \mathbb{IR}^{n \times n}$ (interval matrix) and $g \in \mathbb{IR}^n$ (interval vector) such that

\[ \nabla^2 f(x) \in H \quad \forall x \in \mathbf{x}, \]
\[ \nabla f(x) \in g \quad \forall x \in \mathbf{x}. \]

- Such interval enclosures of the Hessian matrix and the gradient can be computed, e.g., by interval arithmetic using automatic differentiation.
- If every $H \in H$ is positive semidefinite, then $f(x)$ is convex and we are done. Therefore, we focus on problems such that not every $H \in H$ is positive semidefinite.

We will use the symmetric interval matrix

\[ D := \begin{pmatrix} 0 & g^T \\ g & H \end{pmatrix}. \]
Mereau and Paquet suggest to verify positive semidefiniteness of matrices
\[ M_{\alpha}(H, g) := H + \alpha gg^T, \quad H \in H, \ g \in g \]
for a suitable \( \alpha \geq 0 \).

**Direct Evaluation (MP1)**

By interval arithmetic and for a suitable \( \alpha \geq 0 \) evaluate
\[ M(\alpha) := H + \alpha gg^T. \]

Then check whether \( M(\alpha) \) is positive semidefinite.

Problems:
- Choice of \( \alpha \).
- Checking positive semidefiniteness of interval matrices is co-NP-hard.
- This approach does not utilize the structure of \( M_{\alpha}(x) \).

Sufficient condition is: \( \lambda_n(M(\alpha)_c) \geq \rho(M(\alpha)_\Delta) \).
Methods Based on Mereau and Paquet

Theorem

We have that $M_\alpha(H, g)$ is positive semidefinite for all $H \in \mathbf{H}$ and $g \in \mathbf{g}$ if

$$x^T(H_c + \alpha g_c g_c^T)x - |x|^T H_\Delta |x| - 2\alpha |g_c^T x||g_\Delta^T x| \geq 0, \quad \forall x \in \mathbb{R}^n$$

Theorem

We have that $M_\alpha(H, g)$ is positive semidefinite for all $H \in \mathbf{H}$ and $g \in \mathbf{g}$ if

$$H_c - \text{diag}(z)H_\Delta \text{diag}(z) + \alpha (g_c g_c^T - g_c g_\Delta^T \text{diag}(z) - \text{diag}(z) g_\Delta g_c^T)$$

is positive semidefinite for every $z \in \{\pm 1\}^n$.

Structure-Oriented Method (MP2)

Based on the above exponential formula.
Ferland suggests to check that for each symmetric $D \in \mathcal{D}$ and for each $r = 2, \ldots, n + 1$ we have $\det(D_r) < 0$.

**Theorem**

*It is co-NP-hard to check whether $\det(D) < 0$ for every symmetric $D \in \mathcal{D}$.*

**The Method (F)**

Check

$$\det((D_r)_c) < 0 \quad \text{and} \quad \rho(||(D_r)_c^{-1}||(D_r)_\Delta) < 1$$

for each $r = 2, \ldots, n + 1$. 

For $H$ symmetric, the condition that there is $b$ such that $Hb = g$, $g^T b < 0$ is equivalent to

$$\det(D) = \det \begin{pmatrix} 0 & g^T \\ g & H \end{pmatrix} < 0.$$ 

This gives us an equivalent condition:

**Theorem**

The function $f(x)$ is pseudoconvex on $x$ if for each symmetric $D \in D$ we have $\det(D) < 0$, and each symmetric $H \in H$ is nonsingular and has at most one simple negative eigenvalue.

**The Method (CF)**

The function $f(x)$ is pseudoconvex on $x$ if

$$\det(D_c) < 0, \quad \rho(|D_c^{-1}|D_\Delta) < 1, \quad \text{and} \quad 0 < \lambda_{n-1}(H_c) - \rho(H_\Delta).$$
Ferland suggests to check that the $n$th largest eigenvalue of every symmetric matrix $D \in D$ is positive.

**Theorem**

Checking that the $n$th largest eigenvalue of every symmetric matrix $D \in D$ is positive is a co-NP-hard problem even on the class of problems with $g = 0$, $H_c$ symmetric positive definite and entrywise nonnegative, and $H_\Delta$ consisting of ones.

**The Method (C)**

The function $f(x)$ is pseudoconvex on $x$ if $0 \not\in g$ and $\lambda_n(D_c) > \rho(D_\Delta)$. 
Numerical Experiments

Example (Random choices of $H$ and $g$)

$n =$dimension, $d =$radius of $H$ and $g$, 
$H := H - \gamma I_n$ minimally to fail positive semidefiniteness.

<table>
<thead>
<tr>
<th>$n$</th>
<th>$d$</th>
<th>MP1</th>
<th>MP2</th>
<th>F</th>
<th>CF</th>
<th>C</th>
<th>MP1</th>
<th>MP2</th>
<th>F</th>
<th>CF</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>1</td>
<td>0</td>
<td>21.2</td>
<td>35.7</td>
<td>40.7</td>
<td>43.5</td>
<td>1.12</td>
<td>9.32</td>
<td>2.14</td>
<td>0.835</td>
<td>0.644</td>
</tr>
<tr>
<td>10</td>
<td>1</td>
<td>0</td>
<td>3.2</td>
<td>9.4</td>
<td>11.0</td>
<td>29.3</td>
<td>0.889</td>
<td>49.8</td>
<td>3.71</td>
<td>0.831</td>
<td>0.669</td>
</tr>
<tr>
<td>15</td>
<td>1</td>
<td>0</td>
<td>0.3</td>
<td>1.0</td>
<td>1.3</td>
<td>20.3</td>
<td>0.958</td>
<td>427</td>
<td>5.34</td>
<td>0.860</td>
<td>0.694</td>
</tr>
<tr>
<td>20</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>11.8</td>
<td>1.32</td>
<td>3085</td>
<td>7.43</td>
<td>1.20</td>
<td>0.775</td>
</tr>
<tr>
<td>5</td>
<td>0.1</td>
<td>47</td>
<td>52</td>
<td>66.8</td>
<td>67.7</td>
<td>65.4</td>
<td>0.978</td>
<td>6.45</td>
<td>2.24</td>
<td>0.814</td>
<td>0.629</td>
</tr>
<tr>
<td>10</td>
<td>0.1</td>
<td>37</td>
<td>50.3</td>
<td>61</td>
<td>62</td>
<td>56.1</td>
<td>3.88</td>
<td>193</td>
<td>4.38</td>
<td>0.936</td>
<td>0.662</td>
</tr>
<tr>
<td>15</td>
<td>0.1</td>
<td>26.7</td>
<td>45.7</td>
<td>54.6</td>
<td>55.5</td>
<td>41.8</td>
<td>109</td>
<td>5814</td>
<td>6.61</td>
<td>0.973</td>
<td>0.681</td>
</tr>
<tr>
<td>20</td>
<td>0.1</td>
<td>25</td>
<td>51</td>
<td>57</td>
<td>57</td>
<td>41</td>
<td>6689</td>
<td>280048</td>
<td>11.1</td>
<td>1.25</td>
<td>0.793</td>
</tr>
</tbody>
</table>

The winners: Crouzeix and Ferland (CF) and Crouzeix (C)

Open problems: choice of $\alpha$ in (MP1–2), improve (CF) and (C)
References


