New inclusion functions in interval global optimization of engineering structures

Andrzej Pownuk
Chair of Theoretical Mechanics
Faculty of Civil Engineering
Silesian University of Technology

Pownuk Andrzej, URL: http://zeus.polsl.gliwice.pl/~pownuk
Global optimization

• Gradient methods
• Stochastic methods
• Special methods (e.g. linear programming)
• Analytical methods
• **Branch and bound methods**
Problem with gradient methods

Many local minima

\[ f(x) = x^2 \cdot \left( 1 - 0.5 \cdot \cos \left( \frac{1}{x} \right) \right) \]
Problem with stochastic methods

\[ y = x^2 \]

Convergence criteria

Pownuk Andrzej, URL: http://zeus.polsl.gliwice.pl/~pownuk
Branch and bound methods
Contents

Properties of interval global optimization

Interval arithmetic

Acceleration devices

New inclusion functions

Examples of applications

Conclusions
Properties of interval global optimization

The algorithm guarantees that all stationary global solutions (in the initial interval) have been found.

The bounds on the solution(s) are guaranteed to be correct.

Error from all sources are accounted for.

The algorithm can solve the global optimization problem when the objective function is nondifferentiable or even not continuous.
Software

GlobSol - SUN Microsystems and Marquette University

Numerica - Pascal Van Hentenryck, Laurent Michel, i Yves Deville

UniCalc - Russian Research Institute of Artificial Intelligence

GLOBOPT - Arnold Neumaier, University of Vienna

Pownuk Andrzej, URL: http://zeus.polsl.gliwice.pl/~pownuk
Interval arithmetic

\[ \bar{x} = [x^-, x^+] = \{x \in R : x^- \leq x \leq x^+ \} \]

Interval operations

\[ \bar{x} \circ \bar{y} = \{x \circ y : x \in \bar{x}, y \in \bar{y} \} \]

for example

\[ \bar{x} + \bar{y} = [x^- + y^-, x^+ + y^+] \]
Interval extension

\[ f(x) = x^2 - x \]

\[ \hat{f}(x) = \bar{x} \cdot \bar{x} - \bar{x} \]

Interval function

\[ f(\bar{x}) = \{ f(\bar{x}) : x \in \bar{x} \} \]
Interval extension

\[ \hat{f}([-1, 2]) = [-1, 2] \cdot [-1, 2] - [-1, 2] = [-4, 5] \]

\[ [-1, 2] \cdot [-1, 2] = [-2, 4] \]

\[ [-2, 4] - [-1, 2] = [-2, 4] + [-2, 1] = [-4, 5] \]

\[ f([-1, 2]) = \{ f(x) : x \in [-1, 2] \} = [-\frac{1}{4}, 2] \]
\[
\hat{f}([-1, 2]) = [-1, 2] \cdot [-1, 2] - [-1, 2] = [-4, 5]
\]
\[
f([-1, 2]) = \{ f(x) : x \in [-1, 2] \} = [-\frac{1}{4}, 2]
\]
\[
f([-1, 2]) \subseteq \hat{f}([-1, 2])
\]

**Fundamental property of interval arithmetic**

\[
f(\overline{x}) \subseteq \hat{f}(\overline{x})
\]

Pownuk Andrzej, URL: http://zeus.polsl.gliwice.pl/~pownuk
\[ f(x) = x^2 - x \]

\[
\min_{x \in [-1, 2]} f(x) = -\frac{1}{4} \quad \max_{x \in [-1, 2]} f(x) = 2
\]

\[
\min \hat{f}([-1, 2]) = -4 \quad \max \hat{f}([-1, 2]) = 5
\]

\[
\min \hat{f}(\bar{x}) \leq \min_{x \in \bar{x}} f(x) \leq \max_{x \in \bar{x}} f(x) \leq \max \hat{f}(\bar{x})
\]
Basic algorithm

\[ \hat{f}(\bar{x}_1) < \hat{f}(\bar{x}_2) \]

Pownuk Andrzej, URL: http://zeus.polsl.gliwice.pl/~pownuk
Example

\[ f(x) = x^2 - x, \quad \bar{x} = [-10, 10] \]

\[ x_{opt} = \frac{1}{2}, \quad f_{opt} = -\frac{1}{4} \]

\[ L = \{\bar{x}\} \]

Iteration 1

<table>
<thead>
<tr>
<th>\bar{x}</th>
<th>[-10, 10]</th>
</tr>
</thead>
<tbody>
<tr>
<td>\hat{f}(\bar{x})</td>
<td>[-110, 110]</td>
</tr>
<tr>
<td>f(mid(\bar{x}))</td>
<td>0</td>
</tr>
</tbody>
</table>
\[ L = \{ \bar{x}_1, \ldots, \bar{x}_n \} \]
\[ f^- = \min \{ \inf \hat{f}(\bar{x}_1), \ldots, \inf \hat{f}(\bar{x}_n) \} \]
\[ f^+ = \max \{ \sup \hat{f}(\bar{x}_1), \ldots, \sup \hat{f}(\bar{x}_n) \} \]
\[ f_{opt} = \min \{ f(\text{mid}(\bar{x}_1)), \ldots, f(\text{mid}(\bar{x}_n)) \} \]
\[ f^- \leq f_{opt} \leq f^+ \]
\[ x^- = \min \{ \inf \bar{x}_1, \ldots, \inf \bar{x}_n \} \]
\[ x^+ = \max \{ \sup \bar{x}_1, \ldots, \sup \bar{x}_n \} \]
\[ x^- \leq x_{opt} \leq x^+ \]

Pownuk Andrzej, URL: http://zeus.polsl.gliwice.pl/~pownuk
Iteration 1

<table>
<thead>
<tr>
<th></th>
<th>[-10, 10]</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\bar{x}$</td>
<td>$[-10, 10]$</td>
</tr>
<tr>
<td>$\hat{f}(\bar{x})$</td>
<td>$[-110, 110]$</td>
</tr>
<tr>
<td>$f(mid(\bar{x}))$</td>
<td>0</td>
</tr>
</tbody>
</table>

\[
f^- \leq f_{opt} \leq f^+
\]

\[
f^- = -110 \quad f_{opt} = 0 \quad f^+ = 110
\]

\[
x^- \leq x_{opt} \leq x^+
\]

\[
x^- = -10 \quad x_{opt} = 0 \quad x^+ = 10
\]

\[
f^+ - f^- = 220
\]

\[
x^+ - x^- = 20
\]
### Iteration 2

<table>
<thead>
<tr>
<th>$\bar{x}$</th>
<th>$\hat{f}(\bar{x})$</th>
<th>$f(mid(\bar{x}))$</th>
</tr>
</thead>
<tbody>
<tr>
<td>[-10, 0]</td>
<td>[0, 110]</td>
<td>30</td>
</tr>
<tr>
<td>[0, 10]</td>
<td>[-10, 100]</td>
<td>20</td>
</tr>
</tbody>
</table>

\[
\begin{align*}
  f^- &= -10 & f_{opt} &= 0 & f^+ &= 110 \\
  x^- &= -10 & x_{opt} &= 0 & x^+ &= 10 \\
  f^+ - f^- &= 120 \\
  x^+ - x^- &= 20
\end{align*}
\]

Pownuk Andrzej, URL: http://zeus.polsl.gliwice.pl/~pownuk
### Iteration 3

<table>
<thead>
<tr>
<th>$\bar{x}$</th>
<th>$\hat{f}(\bar{x})$</th>
<th>$f(mid(\bar{x}))$</th>
</tr>
</thead>
<tbody>
<tr>
<td>[-10, -5]</td>
<td>[30, 110]</td>
<td>63.75</td>
</tr>
<tr>
<td>[-5, 0]</td>
<td>[0, 30]</td>
<td>8.75</td>
</tr>
<tr>
<td>[0, 5]</td>
<td>[-5, 25]</td>
<td>3.75</td>
</tr>
<tr>
<td>[5, 10]</td>
<td>[15, 95]</td>
<td>48.75</td>
</tr>
</tbody>
</table>

\[
\begin{align*}
    f^- &= -5, \quad f_{opt} = 0, \quad f^+ = 30 \\
    x^- &= -5, \quad x_{opt} = 0, \quad x^+ = 5 \\
    f^+ - f^- &= 35 \\
    x^+ - x^- &= 10
\end{align*}
\]

Pownuk Andrzej, URL: http://zeus.polsl.gliwice.pl/~pownuk
### Iteration 4

<table>
<thead>
<tr>
<th>$\bar{x}$</th>
<th>$\hat{f}(\bar{x})$</th>
<th>$f(mid(\bar{x}))$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$[-5, -2.5]$</td>
<td>$[8.75, 30]$</td>
<td>17.81</td>
</tr>
<tr>
<td>$[-2.5, 0]$</td>
<td>$[0, 8.75]$</td>
<td>2.81</td>
</tr>
<tr>
<td>$[0, 2.5]$</td>
<td>$[-2.5, 6.25]$</td>
<td>0.31</td>
</tr>
<tr>
<td>$[2.5, 5]$</td>
<td>$[1.25, 22.5]$</td>
<td>10.31</td>
</tr>
</tbody>
</table>

\[
f^- = -2.5, \quad f_{opt} = 0, \quad f^+ = 8.75 \]
\[
x^- = -2.5, \quad x_{opt} = 0, \quad x^+ = 2.5 \]
\[
f^+ - f^- = 11.25 \]
\[
x^+ - x^- = 5 \]

Pownuk Andrzej, URL: http://zeus.polsl.gliwice.pl/~pownuk
Iteration 5

<table>
<thead>
<tr>
<th>$\bar{x}$</th>
<th>$\hat{f}(\bar{x})$</th>
<th>$f(mid(\bar{x}))$</th>
</tr>
</thead>
<tbody>
<tr>
<td>[-2.5, -1.25]</td>
<td>[2.81, 8.75]</td>
<td>5.39</td>
</tr>
<tr>
<td>[-1.25, 0]</td>
<td>[0, 2.81]</td>
<td>1.02</td>
</tr>
<tr>
<td>[0, 1.25]</td>
<td>[-1.25, 1.56]</td>
<td>-0.23</td>
</tr>
<tr>
<td>[1.25, 2.5]</td>
<td>[-0.93, 5]</td>
<td>1.64</td>
</tr>
</tbody>
</table>

$f^- = -1.25, f_{opt} = -0.23, f^+ = 5$

$x^- = 0, x_{opt} = 0.625, x^+ = 2.5$

$f^+ - f^- = 6.25$

$x^+ - x^- = 2.5$

Pownuk Andrzej, URL: http://zeus.polsl.gliwice.pl/~pownuk
<table>
<thead>
<tr>
<th>$\bar{x}$</th>
<th>$\hat{f}(\bar{x})$</th>
<th>$f(\text{mid}(\bar{x}))$</th>
</tr>
</thead>
<tbody>
<tr>
<td>[0, 0.625]</td>
<td>[-0.625, 0.39]</td>
<td>-0.214</td>
</tr>
<tr>
<td>[0.625, 1.25]</td>
<td>[-0.859, 0.93]</td>
<td>-0.059</td>
</tr>
<tr>
<td>[1.25, 1.875]</td>
<td>[-0.31, 2.26]</td>
<td>0.878</td>
</tr>
<tr>
<td>[1.875, 2.5]</td>
<td>[1.01, 4.37]</td>
<td>1.64</td>
</tr>
</tbody>
</table>

\[
f^- = -0.859, \quad f_{opt} = -0.23, \quad f^+ = 2.26
\]

\[
x^- = 0, \quad x_{opt} = 0.625, \quad x^+ = 1.875
\]

\[
f^+ - f^- = 3.119
\]

\[
x^+ - x^- = 1.875
\]

Pownuk Andrzej, URL: http://zeus.polsl.gliwice.pl/~pownuk
Basic algorithm

**Step 0** Set \([y] = [x]\) and \(y = \hat{f}([x])^-\). Initialize the list \(L = ([y], y))\) and the cut-off level \(z = \hat{f}([x])^+\).

**Step 1** Choose a coordinate direction \(k \in \{1, 2, ..., n\}\).

**Step 2** Bisect \([y]\) in direction \(k\): \([y] = [v_1] \cup [v_2]\).

**Step 3** Calculate \(\hat{f}([v_1])\) and \(\hat{f}([v_2])\) and set \(v_i = \hat{f}([v_i])^-\) for \(i = 1, 2\) and \(z = \min\{z, \hat{f}([v_1])^+, \hat{f}([v_2])^+\}\).

**Step 4** Remove \((y, y)\) from the list \(L\).

**Step 5** Cutoff test: discard the pair \(([v_i], v_i)\) if \(v_i > z\) (where \(i = 1, 2\)).

**Step 6** Add any remaining pair(s) to the list \(L\). If the list becomes empty then STOP.

**Step 7** Denote the pair with the smallest second element by \([y], y\) .

**Step 8** If the width of \(\hat{f}([y])\) is less than \(\varepsilon\), then print \(\hat{f}([y])\) and STOP.

**Step 9** Go to step 1.
Acceleration devices

Monotonicity test

\[ 0 \not\in \frac{\partial \hat{f}(\bar{x})}{\partial x_j} \implies \forall x \in \bar{x}, \frac{\partial f(x)}{\partial x_i} \neq 0 \]

and function \( y = f(\ldots x_i, \ldots) \) is monotone.
Acceleration devices

Monotonicity test

If $\frac{df}{dx} > 0$, then $y^- = f(x^-)$, $y^+ = f(x^+)$

If $\frac{df}{dx} < 0$, then $y^- = f(x^+)$, $y^+ = f(x^-)$

it can be written

$$y^- = f(x^{-\text{sign}\left(\frac{df}{dx}\right)}), \quad y^+ = f(x^{\text{sign}\left(\frac{df}{dx}\right)})$$
Acceleration devices

Midpoint test

If $f(mid(x_1)) < \inf \hat{f}(x_2)$
then $\forall x \in x_2$, $f(mid(x_1)) < f(x)$
and $x_{opt} \notin x_2$

$f(x_{opt}) = \inf_{x: x \in x} f(x)$

The interval $x_2$ can be neglected in a future calculation

Pownuk Andrzej, URL: http://zeus.polsl.gliwice.pl/~pownuk
Acceleration devices

A nonconvexity check

If the function $f$ has unconstrained minimum at $x^*$, then $f$ must be convex in some neighborhood of $x^*$.

Hence, the Hessian $H$ of $f$ must be positive semidefinite at $x^*$.

A necessary condition for this is that the diagonal elements $H_{ii}$ $(i=1,..., n)$ be nonnegative.

Pownuk Andrzej, URL: http://zeus.polsl.gliwice.pl/~pownuk
Acceleration devices

A nonconvexity check

Consider an interval $\mathbf{x}$.

If $\hat{H}_{ii}(\mathbf{x})^+ < 0$ for some $i = 1,\ldots, n$
then $H_{ii}(\mathbf{x}) < 0$ for all $\mathbf{x} \in \mathbf{x}$.
Hence, $\mathbf{H}$ cannot be positive semidefinite
for any point in $\mathbf{x}$.

Therefore, $f$ cannot have a stationary
minimum in $\mathbf{x}$ and $\mathbf{x}$ can be deleted.
The interval Newton method

1) Let $x^* \in \bar{x}_0$
2) $x_n = \text{mid}(\bar{x}_n)$
3) $N(x_n, \bar{x}_n) = x_n - \frac{f(x_n)}{\hat{f}'(\bar{x}_n)}$
4) $\bar{x}_{n+1} = \bar{x}_n \cap N(x_n, \bar{x}_n)$
5) Go to 1

Using this method we can find all solutions of algebraic equations
Acceleration devices

The interval Newton method

If \( \forall x \in \bar{x}, \text{grad } f(x) \neq 0, \)
then \( x_{opt} \notin \bar{x} \)

The interval \( \bar{x} \) can be neglected in a future calculation

Pownuk Andrzej, URL: http://zeus.polsl.gliwice.pl/~pownuk
Acceleration devices

Finding a function value as small as possible

If \( f(x^*) < \inf \hat{f}(\bar{x}_2) \) where \( x^* \in \bar{x} \)

then \( \forall x \in \bar{x}_2, f(x^*) < f(x) \)

Midpoint test gives better results if the number \( f(x^*) \) is as small as possible.

The point \( x^* \) can be found using any optimization method.

Pownuk Andrzej, URL: http://zeus.polsl.gliwice.pl/~pownuk
Acceleration devices

Use a good inclusion function

In a calculation it is better to use an inclusion function rather than the natural interval extension

Fundamental property of interval arithmetic

\[ f(x) \subseteq \hat{f}(x) \]

or

\[ w(f(x)) \leq w(\hat{f}(x)) \]

where \( w(\bar{x}) = x^+ - x^- \)

Interval extension gives overestimated results.

Pownuk Andrzej, URL: http://zeus.polsl.gliwice.pl/~pownuk
New inclusion function

$$0 \notin \frac{\partial \hat{f}(\bar{x})}{\partial x_i}$$

First order monotonicity test

$$-\text{sign}\left(\frac{\partial \hat{f}(\bar{x})}{\partial x}\right), \quad \text{sign}\left(\frac{\partial \hat{f}(\bar{x})}{\partial x}\right)$$

$$y^- = f(x - f(x), \quad y^+ = f(x + f(x))$$

$$f(\bar{x}) = y = [y^-, y^+]$$

Exact result

Pownuk Andrzej, URL: http://zeus.polsl.gliwice.pl/~pownuk
Second order monotonicity test

\[ 0 \in \frac{\partial \hat{f}(x)}{\partial x_i} \quad 0 \notin \frac{\partial^2 \hat{f}(x)}{\partial x_i \partial x_j} \]

\[ y_1^- = \frac{\partial f(x)}{\partial x} (x - \text{sign} \left( \frac{\partial^2 \hat{f}(x)}{\partial x^2} \right)), \quad y_1^+ = \frac{\partial f(x)}{\partial x} (x + \text{sign} \left( \frac{\partial^2 \hat{f}(x)}{\partial x^2} \right)) \]

\[ \bar{y}_1 = [y_1^-, y_1^+] \text{ (exact result)} \]

If \( 0 \notin \bar{y}_{1i} \), then

\[ y^- = f(x^{-\text{sign}(y_1)}), \quad y^+ = f(x^{\text{sign}(y_1)}) \]

Pownuk Andrzej, URL: http://zeus.polsl.gliwice.pl/~pownuk
N-th order monotonicity test

\[
\begin{align*}
\bar{y}_N^- &= \inf \frac{\partial^N \hat{f}(\bar{x})}{\partial \bar{x}^N} , \quad \bar{y}_N^+ &= \sup \frac{\partial^N \hat{f}(\bar{x})}{\partial \bar{x}^N} \\
\bar{y}_N &= [\bar{y}_N^-, \bar{y}_N^+] \\
\end{align*}
\]

\[
\begin{align*}
\bar{y}_k^- &= \frac{\partial^k f}{\partial \bar{x}^k} (\bar{x} - \text{sign}(\bar{y}_{k+1})) , \quad \bar{y}_k^+ &= \frac{\partial^k f}{\partial \bar{x}^k} (\bar{x} + \text{sign}(\bar{y}_{k+1})) \\
\end{align*}
\]

If \(0 \not\in \bar{y}_k\), then

the method gives exact result

Pownuk Andrzej, URL: http://zeus.polsl.gliwice.pl/~pownuk
Taylor series expansion

\[ f(x) \approx f_N(x) = \sum_{i=1}^{N} \frac{1}{i!} \frac{\partial^i f(x_0)}{\partial x^i} (x - x_0)^k \]

To calculate the coefficients \( \frac{\partial^k f(x_0)}{\partial x^k} \) we can apply sensitivity analysis.

This method is more efficient than pure interval method.

This method cannot get results with guaranteed accuracy.

Pownuk Andrzej, URL: http://zeus.polsl.gliwice.pl/~pownuk
\[ y = f(x_0) + f'(x_0) \cdot (x - x_0) \]

\[ y = f(x) \]

\[ \bar{x} \subseteq \bar{x}_M \quad \text{and} \quad 2 \cdot w(\bar{x}) = w(\bar{x}_M) \]

Pownuk Andrzej, URL: http://zeus.polsl.gliwice.pl/~pownuk
Shape optimisation of truss
Shape optimisation of truss

Initial shape
Final shape

P=1000 [kN]

P=100 [kN]

Pownuk Andrzej, URL: http://zeus.polsl.gliwice.pl/~pownuk
Objective function

\[ f(x) = \sum_{e=1}^{Ne} A_e L_e = \frac{1}{\sigma_0} \sum_{e=1}^{Ne} \left| N_e \right| \cdot L_e \]

Constraints

\[ \sigma_0 - \sigma_e \geq 0 \quad \text{(stress)} \]

\[ \left| N_e \right| \leq P_{st} \quad \text{(stability)} \]
Conclusions

The algorithm guarantees that all stationary global solutions (in the initial interval) have been found.

The bounds on the solution(s) are guaranteed to be correct.

Error from all sources are accounted for.

The algorithm can solve the global optimization problem when the objective function is nondifferentiable or even not continuous.
Conclusions

In order to check monotonicity of the objective function sensitivity analysis can be applied.

If we apply the Taylor expansion then the calculation are more efficient but we lose guaranteed accuracy of the results.

Presented algorithm is more efficient than method based on the natural interval extension.