An Extension of the αBB-type Underestimation to Linear Parametric Hessian Matrices

Milan Hladík

Department of Applied Mathematics,
Faculty of Mathematics and Physics,
Charles University in Prague,
Czech Republic

MAGO'14
XII Global Optimization Workshop, Málaga
September 1–4, 2014

Problem formulation

Convex underestimators

Let

1. $f : \mathbb{R}^n \mapsto \mathbb{R}$ be a twice-differentiable function,
2. $x_i \in \mathbf{x}_i = [\underline{x}_i, \overline{x}_i]$, $i = 1, \ldots, n$, interval domains for the variables.

Construct a function $g : \mathbb{R}^n \mapsto \mathbb{R}$ satisfying:

1. $f(x) \geq g(x)$ for every $x \in \mathbf{x}$,
2. $g(x)$ is convex on $x \in \mathbf{x}$.

Application

- Deterministic global optimization methods based on branch & bound scheme.
- Rigorous bound on objective and optimal values.
- Rigorous enclosures of constraints.

Forms of underestimators

Forms of underestimators

1. αBB method (Floudas et al., 1995–2013)

$$g(x) := f(x) - \sum_{i=1}^{n} \alpha_i(\overline{x}_i - x_i)(x_i - \underline{x}_i), \qquad (*)$$

2. non-diagonal αBB method (Akrotirianakis et al., 2004, Skjäl et al., 2012)

$$g(x) := f(x) - (\overline{x} - x)^T P(x - \underline{x}) + q(x),$$

   where $q(x)$ is a piecewise linear convex function.

3. γBB method (Akrotirianakis and Floudas, 2004)

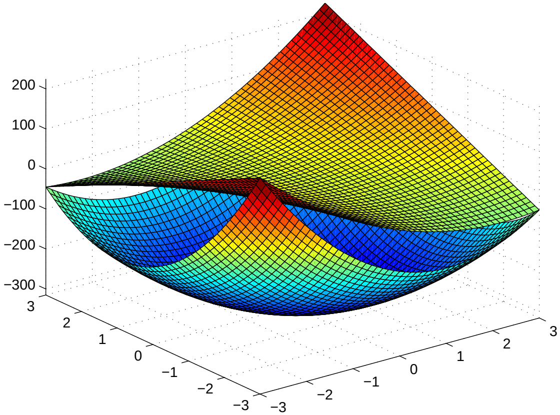
$$g(x) := f(x) - \sum_{i=1}^{n}(1 - e^{\gamma_i(\overline{x}_i - x_i)})(1 - e^{\gamma_i(x_i - \underline{x}_i)}) \qquad (\dagger)$$

Theorem (Floudas and Kreinovich, 2007)

Forms $(*)$ and $(\dagger)$ are the only shift-invariant forms of the gen. scheme

$$g(x) := f(x) - \sum_{i=1}^{n} \alpha_i(g_i(\overline{x}_i) - g_i(x_i))(g_i(x_i) - g_i(\underline{x}_i)).$$

Illustration



Function $f(x)$ and its convex underestimator $g(x)$.

Computation of α

Idea

Choose α large enough to ensure positive semidefinitness of the Hessian of

$$g(x) := f(x) - \sum_{i=1}^{n} \alpha_i(\overline{x}_i - x_i)(x_i - \underline{x}_i),$$

The Hessian of $g(x)$ reads

$$\nabla^2 g(x) = \nabla^2 f(x) + 2\,\mathrm{diag}(\alpha).$$

Interval Hessian matrix

Let $\mathbf{H}$ be an interval matrix enclosing the image of $\nabla^2 f(x)$ over $x \in \mathbf{x}$:

$$\frac{\partial^2}{\partial x_i \partial x_j} f(x) \in \mathbf{h}_{ij} = [\underline{h}_{ij}, \overline{h}_{ij}], \quad \forall x \in \mathbf{x}.$$

Remarks

- Checking positive semidefiniteness of each $H \in \mathbf{H}$ is co-NP-hard (Nemirovskii, 1993).
- Various enclosures for eigenvalues of $H \in \mathbf{H}$.
- Scaled Gerschgorin method enables to express α_i-s.

Computation of α

Scaled Gerschgorin method for α

$$\alpha_i := \max\left\{0, -\frac{1}{2}\left(\underline{h}_{ii} - \sum_{j\neq i} |\mathbf{h}_{ij}| d_j / d_i\right)\right\}, \quad i = 1, \dots, n,$$

where $|\mathbf{h}_{ij}| = \max\left\{|\underline{h}_{ij}|, |\overline{h}_{ij}|\right\}$.

- To reflect the range of the variable domains, use $d := \overline{x} - \underline{x}$.

Theorem (H., 2014)

The choice $d := \overline{x} - \underline{x}$ is optimal (i.e., it minimizes the maximum separation distance between $f(x)$ and $g(x)$) if

$$\underline{h}_{ii} d_i - \sum_{j\neq i} |\mathbf{h}_{ij}| d_j \leq 0, \quad \forall i = 1, \dots, n.$$

An idea for improvement

Consider more gentle enclosure of the Hessian matrices than $\mathbf{H}$.

Hessians in linear parametric forms

An idea for improvement

Consider an interval matrix in a linear parametric form

$$H(p) = \sum_{k=1}^{K} H^{(k)} p_k,$$

where $H^{(1)}, \ldots, H^{(K)} \in \mathbb{R}^{n \times n}$ are fixed matrices and $p_1 \in \mathbf{p}_1, \ldots, p_K \in \mathbf{p}_K$.

Assumption

$$\forall x \in \mathbf{x},\ \exists p \in \mathbf{p} :\ \nabla^2 f(x) = H(p).$$

First (stupid) idea

Evaluate

$$\mathbf{H} := H(\mathbf{p}) = \sum_{k=1}^{K} H^{(k)} \mathbf{p}_k$$

and apply the standard method.

Hessians in linear parametric forms

Second idea

Fix $i \in \{1,\dots,n\}$. Then α_i should satisfy

$$-2\alpha_i \le H(p)_{ii} - \sum_{j\ne i} |H(p)_{ij}| d_j/d_i, \quad \forall p \in \mathbf{p}, \forall i = 1,\dots,n. \qquad (*)$$

Define

$$J^+ := \{j \ne i : \underline{H(\mathbf{p})}_{ij} \ge 0\},$$
$$J^- := \{j \notin J^+ \cup \{i\} : \overline{H(\mathbf{p})}_{ij} \le 0\},$$
$$J^0 := \{1,\dots,n\} \setminus (J^+ \cup J^- \cup \{i\}).$$

For $j \in J^0$, find the best upper estimation

$$|H(p)_{ij}| \le \gamma_{ij} H(p)_{ij} + \beta_{ij},$$

Then $(*)$ reads

$$-2\alpha_i \le H(p)_{ii} - \sum_{j\in J^+} H(p)_{ij}\frac{d_j}{d_i} + \sum_{j\in J^-} H(p)_{ij}\frac{d_j}{d_i} - \sum_{j\in J^0}(\gamma_{ij}H(p)_{ij} + \beta_{ij})\frac{d_j}{d_i}$$
$$= \sum_{k=1}^{K}\left(H_{ii}^{(k)} - \sum_{j\in J^+} H_{ij}^{(k)}\frac{d_j}{d_i} + \sum_{j\in J^-} H_{ij}^{(k)}\frac{d_j}{d_i} - \sum_{j\in J^0}\gamma_{ij}H_{ij}^{(k)}\frac{d_j}{d_i}\right)p_k$$
$$- \sum_{j\in J^0}\beta_{ij}\frac{d_j}{d_i}.$$

Hessians in linear parametric forms

Computation of α

Take

$$\alpha_i := \max\left\{0, -\tfrac{1}{2}\underline{h}_i\right\},$$

where

$$\mathbf{h}_i := \sum_{k=1}^{K}\left(H_{ii}^{(k)} - \sum_{j\in J^+} H_{ij}^{(k)}\frac{d_j}{d_i} + \sum_{j\in J^-} H_{ij}^{(k)}\frac{d_j}{d_i} - \sum_{j\in J^0}\gamma_{ij}H_{ij}^{(k)}\frac{d_j}{d_i}\right)\mathbf{p}_k$$
$$- \sum_{j\in J^0}\beta_{ij}\frac{d_j}{d_i}.$$

Proposition

The second idea is never worse than the first idea.

Example

A class of functions with parametric Hessians:

$$f(x) = \sum_{\ell=1}^{L} c_\ell x_{i_\ell} x_{j_\ell} x_{k_\ell},$$

where $i_\ell, j_\ell, k_\ell \in \{0,\ldots,n\}$ and $x_0 = 1$.

Hessians in linear parametric forms

Example

Let

$$f(x) = x_1^3 + x_1^2 x_2 - 2x_1x_2x_3 + 3x_2x_3^2 - 3x_2^2,$$

where $x \in \mathbf{x} = [0,3] \times [0,3] \times [0,3]$. The Hessian reads

$$\begin{aligned}
\nabla^2 f(x) &= \begin{pmatrix} 6x_1 + 2x_2 & 2x_1 - 2x_3 & -2x_2 \\ 2x_1 - 2x_3 & -6 & -2x_1 + 6x_3 \\ -2x_2 & -2x_1 + 6x_3 & 6x_2 \end{pmatrix} \\
&= \begin{pmatrix} 6 & 2 & 0 \\ 2 & 0 & -2 \\ 0 & -2 & 0 \end{pmatrix} x_1 + \begin{pmatrix} 2 & 0 & -2 \\ 0 & 0 & 0 \\ -2 & 0 & 6 \end{pmatrix} x_2 \\
&+ \begin{pmatrix} 0 & -2 & 0 \\ -2 & 0 & 6 \\ 0 & 6 & 0 \end{pmatrix} x_3 + \begin{pmatrix} 0 & 0 & 0 \\ 0 & -6 & 0 \\ 0 & 0 & 0 \end{pmatrix}.
\end{aligned}$$

- The classical lower bound: -74.1507
- The parametric approach lower bound: -61.5826

Interval computations

Aim

Compute an interval enclosure of the image $f(\mathbf{x}) := \{f(x) : x \in \mathbf{x}\}$.

Interval arithmetic

$$\mathbf{a} + \mathbf{b} = [\underline{a} + \underline{b}, \overline{a} + \overline{b}], \quad \mathbf{a} - \mathbf{b} = [\underline{a} - \overline{b}, \overline{a} - \underline{b}],$$
$$\mathbf{ab} = [\min(\underline{a}\underline{b}, \underline{a}\overline{b}, \overline{a}\underline{b}, \overline{a}\overline{b}), \max(\underline{a}\underline{b}, \underline{a}\overline{b}, \overline{a}\underline{b}, \overline{a}\overline{b})],$$
$$\mathbf{a}/\mathbf{b} = [\min(\underline{a}/\underline{b}, \underline{a}/\overline{b}, \overline{a}/\underline{b}, \overline{a}/\overline{b}), \max(\underline{a}/\underline{b}, \underline{a}/\overline{b}, \overline{a}/\underline{b}, \overline{a}/\overline{b})], \quad 0 \notin \mathbf{b}.$$

Basic functions

$\exp(\mathbf{x}) = [\exp(\underline{x}), \exp(\overline{x})]$, $\mathbf{x}^2 = \ldots$, $\sin(\mathbf{x}) = \ldots$

Natural interval extension

Given an expression for $f(x)$, compute an enclosure by interval arithmetic and basic function evaluation.

Other improvements

Monotonicity checking, interval refinements, ...

Slopes

Slope form enclosure

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

where $a \in \mathbf{x}$ 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

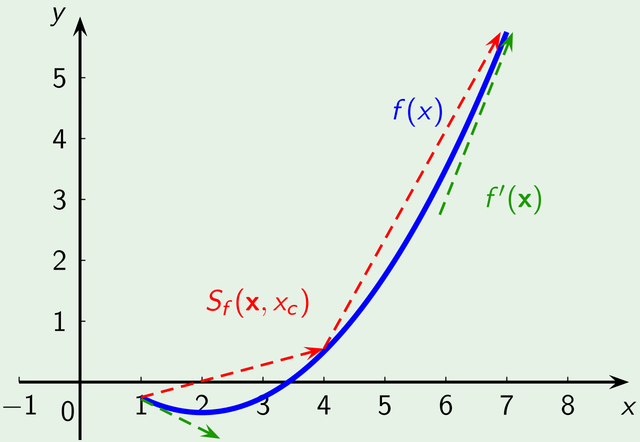
- Slopes can be computed in a similar way as derivatives.
- Slopes can be replaced by derivatives, but slopes gives tighter intervals.

Slopes

Example

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

$$f'(\mathbf{x}) = [-\frac{1}{2}, \frac{5}{2}], \qquad S_f(\mathbf{x}, x_c) = [\frac{1}{4}, \frac{7}{4}].$$



Parametric Hessians by slope expansion

Preliminaries

Denote

$$f_{ij}(x) := \frac{\partial^2}{\partial x_i \partial x_j} f(x).$$

Let us have a slope expansion of $f_{ij}(x)$ on $\mathbf{x}$

$$f_{ij}(x) \in f_{ij}(a) + \mathbf{S}_{ij}(\mathbf{x}, a)^T (x - a).$$

Computation of α

$$\alpha_i := \max\left\{0, -\tfrac{1}{2}\underline{h}_i\right\},$$

where

$$\mathbf{h}_i := \left(\mathbf{S}_{ii}(\mathbf{x}, a) - \sum_{j \in J^+} \mathbf{S}_{ij}(\mathbf{x}, a)\frac{d_j}{d_i} + \sum_{j \in J^-} \mathbf{S}_{ij}(\mathbf{x}, a)\frac{d_j}{d_i} - \sum_{j \in J^0} \gamma_{ij}\mathbf{S}_{ij}(\mathbf{x}, a)\frac{d_j}{d_i}\right)^T (\mathbf{x} - a)$$
$$+ f_{ii}(a) - \sum_{j \in J^+} f_{ij}(a)\frac{d_j}{d_i} + \sum_{j \in J^-} f_{ij}(a)\frac{d_j}{d_i} - \sum_{j \in J^0} (\gamma_{ij} f_{ij}(a) + \beta_{ij})\frac{d_j}{d_i}.$$

Parametric Hessians by slope expansion

Choice of the center $a \in \mathbf{x}$

- Quality of the resulting enclosures depends on a.
- Possible choices: x_c, $\underline{x}$, or which one?

Example (Gounaris and Floudas, 2008)

Let

$$f(x) = \left(x_2 - \frac{5.1}{4\pi^2}x_1^2 + \frac{5}{\pi}x_1 - 6\right)^2 + 10\left(1 - \frac{1}{8\pi}\right)\cos(x_1) + 10$$

with $\mathbf{x} = [7, 10] \times [0, 5]$.

- The optimal value: 0.3979
- The classical lower bound: -13.8967
- The parametric approach lower bound with $a = x_c$: -10.9777
- The parametric approach lower bound with $a = \underline{x}$: -27.1285

Conclusion

Conclusion

- Parametric forms of Hessian enclosures enable more delicate construction of convex underestimators
- A parametric form can be obtained by slope expansion.

Open problems

- Choice of the center of slope expansion $a \in \mathbf{x}$