Convergence-Order Analysis of Lower Bounding Schemes for Constrained Global Optimization Problems

Rohit Kannan and Paul I. Barton

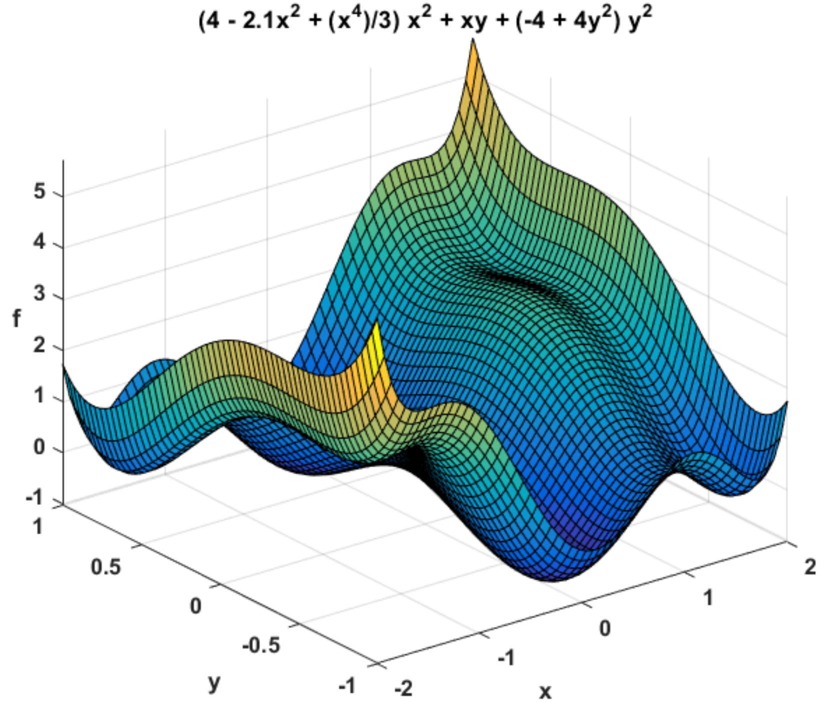
Process Systems Engineering Laboratory
Department of Chemical Engineering
Massachusetts Institute of Technology

August 10, 2016

Motivation

Clustering in Unconstrained Optimization

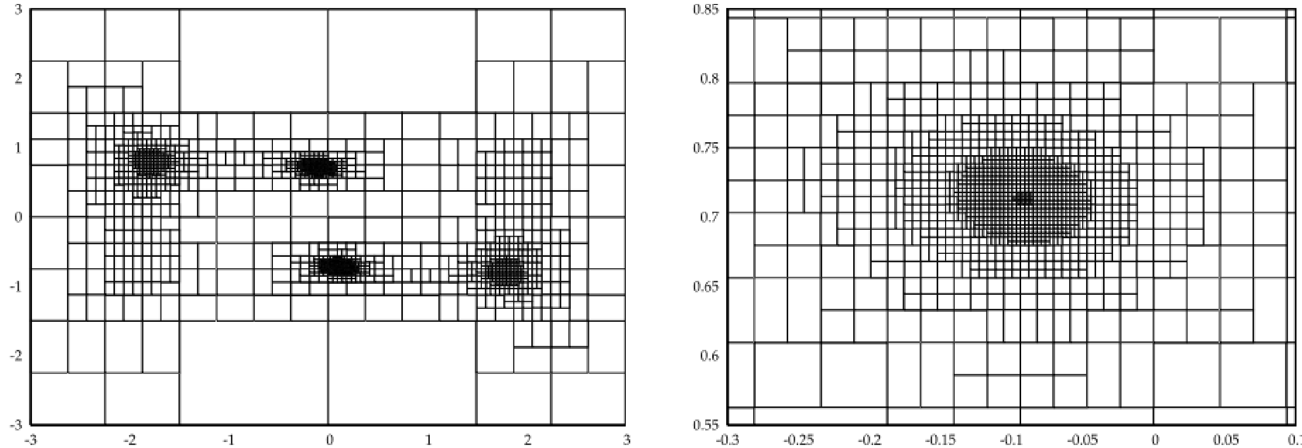
- Unconstrained minimization of



Motivation

Clustering in Unconstrained Optimization

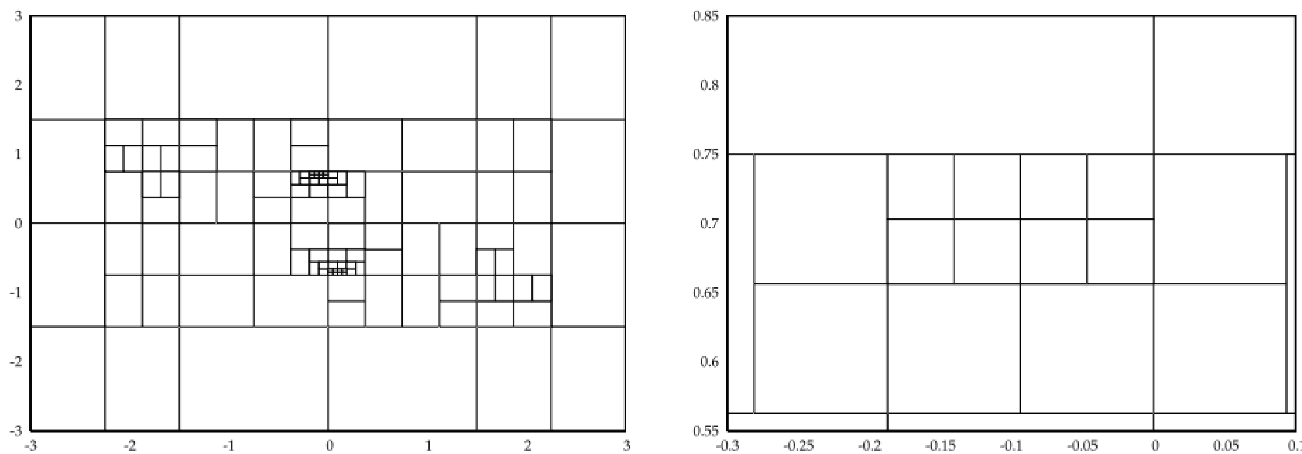
Natural Interval Extension



(a) Full domain

(b) Subset in vicinity of minimum

Centered Form



(c) Full domain

(d) Subset in vicinity of minimum

Wechsung, A., Ph.D. Thesis, MIT, 2014.

Motivation

Clustering in Constrained Optimization

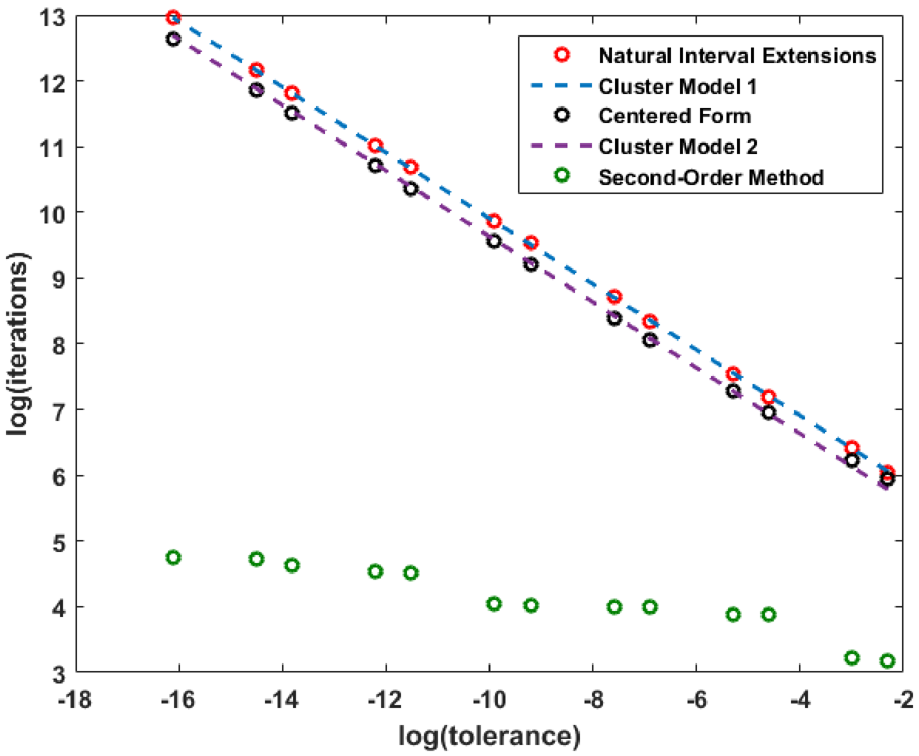
$$\min_{x,y} \; y^2 - 12x - 7y$$
$$\text{s.t.} \;\; y + 2x^4 - 2 = 0,$$
$$x \in [0,2], \; y \in [0,3].$$

Floudas, C. et al., Springer, 1999.

Motivation

Clustering in Constrained Optimization

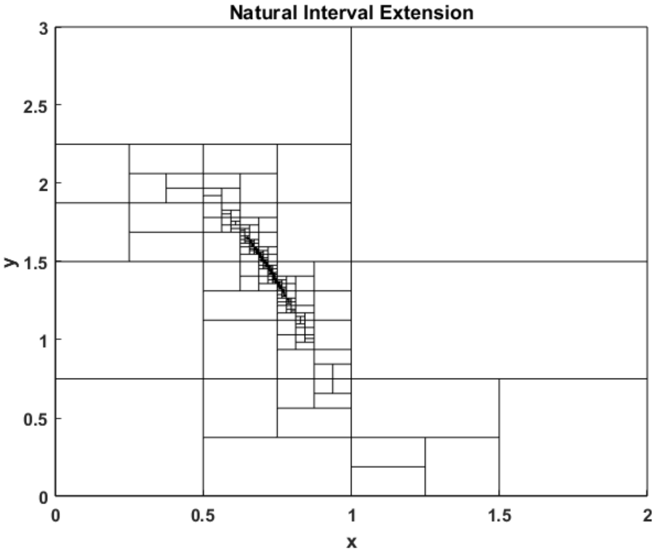
$$\min_{x,y}\ y^2 - 12x - 7y$$
$$\text{s.t.}\ \ y + 2x^4 - 2 = 0,$$
$$x \in [0,2],\ y \in [0,3].$$

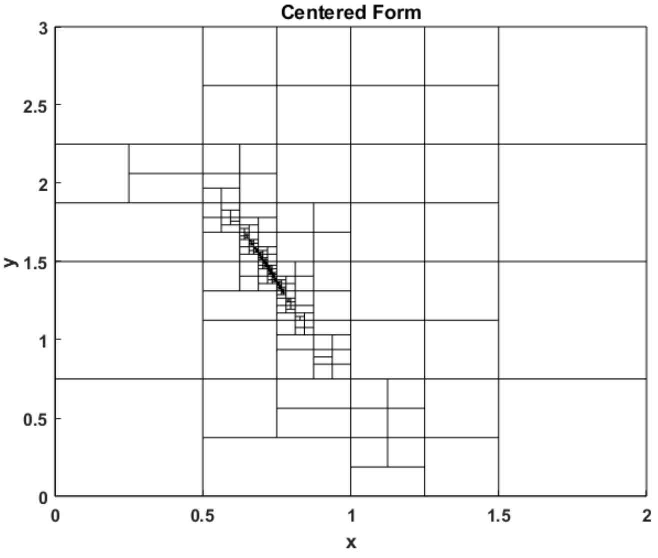


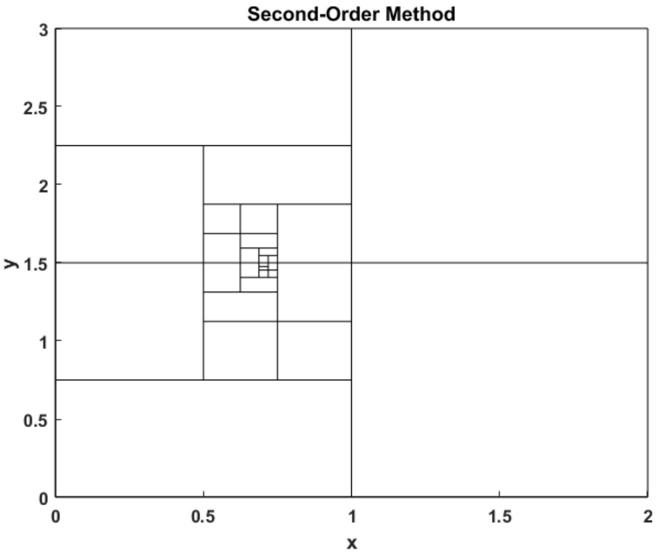
Floudas, C. et al., Springer, 1999.

Motivation

Clustering in Constrained Optimization

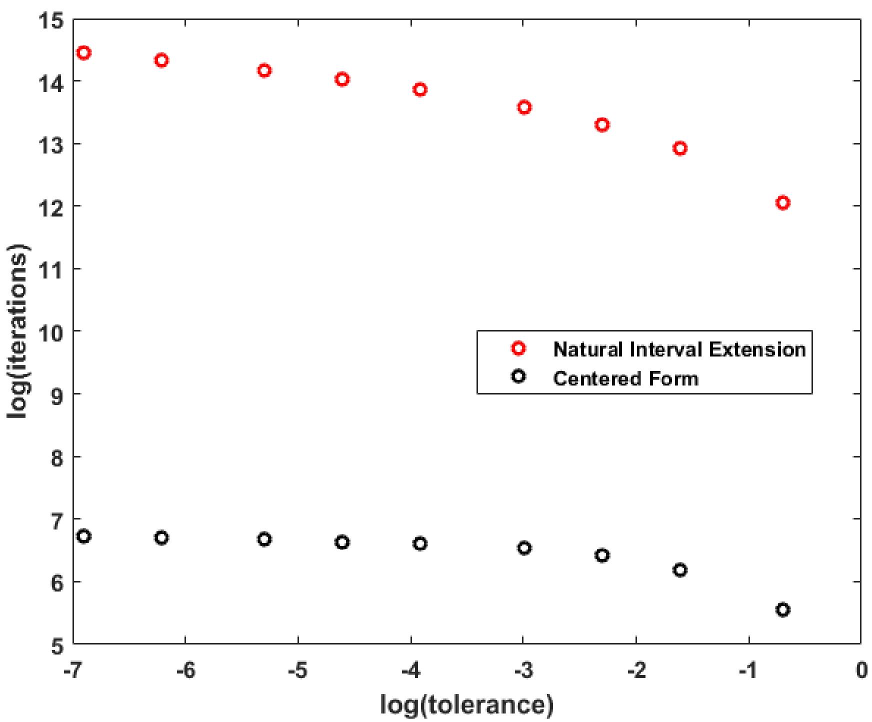






Motivation
Clustering in Constrained Optimization

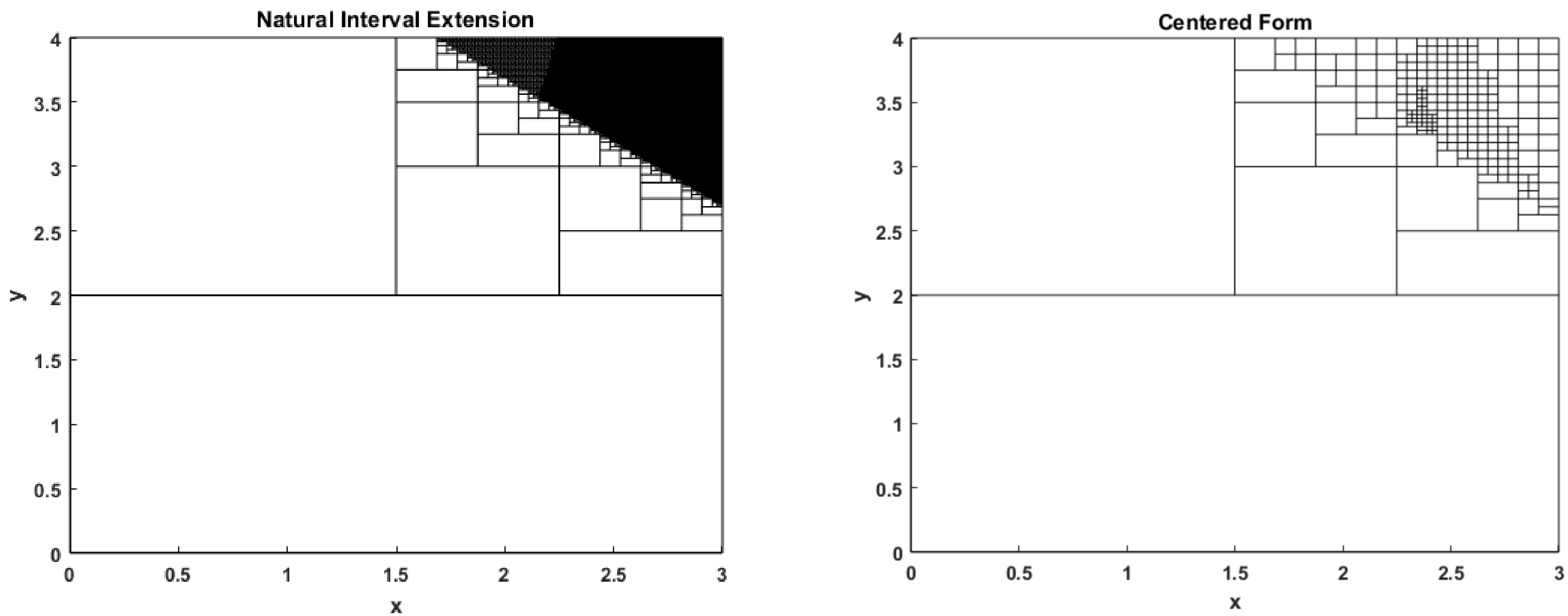
$$\min_{x,y} \; -x-y$$
$$\text{s.t. } \; y \le 2+2x^4-8x^3+8x^2,$$
$$y \le 4x^4-32x^3+88x^2-96x+36,$$
$$x\in[0,3],\; y\in[0,4].$$



Floudas, C. et al., Springer, 1999.

Motivation

Clustering in Constrained Optimization

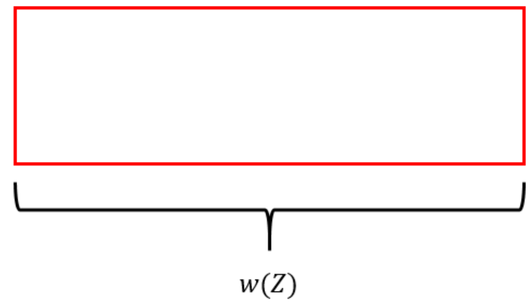


Definitions

- Width of an interval

Let $Z = [z_1^{\mathrm{L}}, z_1^{\mathrm{U}}] \times \cdots \times [z_n^{\mathrm{L}}, z_n^{\mathrm{U}}] \in \mathbb{IR}^n$.

The width of Z is given by $w(Z) = \max\limits_{i=1,\cdots,n}(z_i^{\mathrm{U}} - z_i^{\mathrm{L}})$.



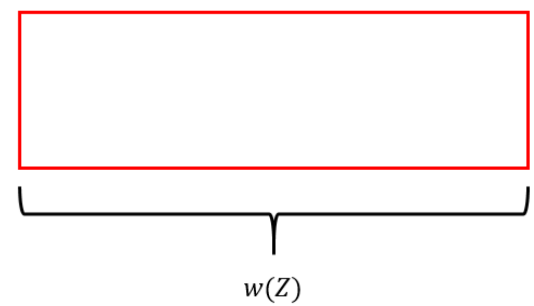
Bompadre, A. et al., J. Global Optim., 2012.

Definitions

- **Width of an interval**

Let $Z = [z_1^{\mathrm{L}}, z_1^{\mathrm{U}}] \times \cdots \times [z_n^{\mathrm{L}}, z_n^{\mathrm{U}}] \in \mathbb{IR}^n$.

The width of Z is given by $w(Z) = \max\limits_{i=1,\cdots,n} (z_i^{\mathrm{U}} - z_i^{\mathrm{L}})$.



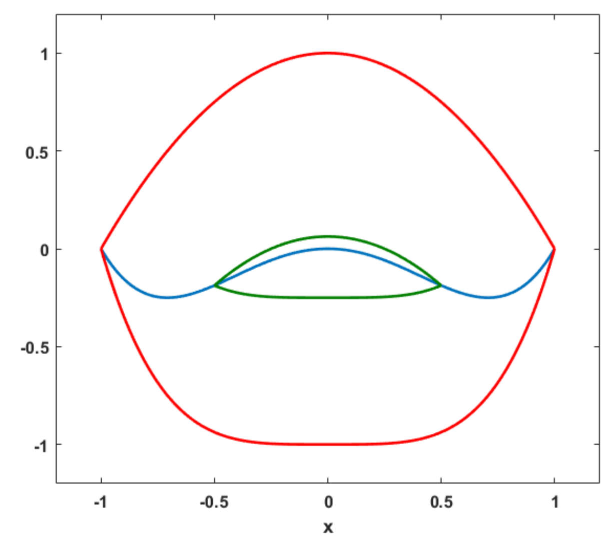
- **Schemes of relaxations**

Nonempty, bounded set $X \subset \mathbb{R}^n$, function $h: X \to \mathbb{R}$.

For each interval $Z \in \mathbb{I}X$, define convex relaxation $h_Z^{\mathrm{cv}}: Z \to \mathbb{R}$, concave relaxation $h_Z^{\mathrm{cc}}: Z \to \mathbb{R}$.

$\left(h_Z^{\mathrm{cv}}\right)\Big|_{Z \in \mathbb{I}X}$ defines a scheme of convex relaxations of h in X.

$\left(h_Z^{\mathrm{cc}}\right)\Big|_{Z \in \mathbb{I}X}$ defines a scheme of concave relaxations of h in X.



Bompadre, A. et al., J. Global Optim., 2012.

Definitions

- Hausdorff metric

Suppose $X = [x^{\mathrm{L}}, x^{\mathrm{U}}], Y = [y^{\mathrm{L}}, y^{\mathrm{U}}] \in \mathbb{IR}$ are two intervals.

Hausdorff metric $q(X, Y) := \max\left\{\left|x^{\mathrm{L}} - y^{\mathrm{L}}\right|, \left|x^{\mathrm{U}} - y^{\mathrm{U}}\right|\right\}$.

Definitions

- ## Hausdorff metric

Suppose $X=[x^{\mathrm{L}},x^{\mathrm{U}}], Y=[y^{\mathrm{L}},y^{\mathrm{U}}]\in\mathbb{IR}$ are two intervals.

Hausdorff metric $q(X,Y):=\max\left\{\left|x^{\mathrm{L}}-y^{\mathrm{L}}\right|,\left|x^{\mathrm{U}}-y^{\mathrm{U}}\right|\right\}$.

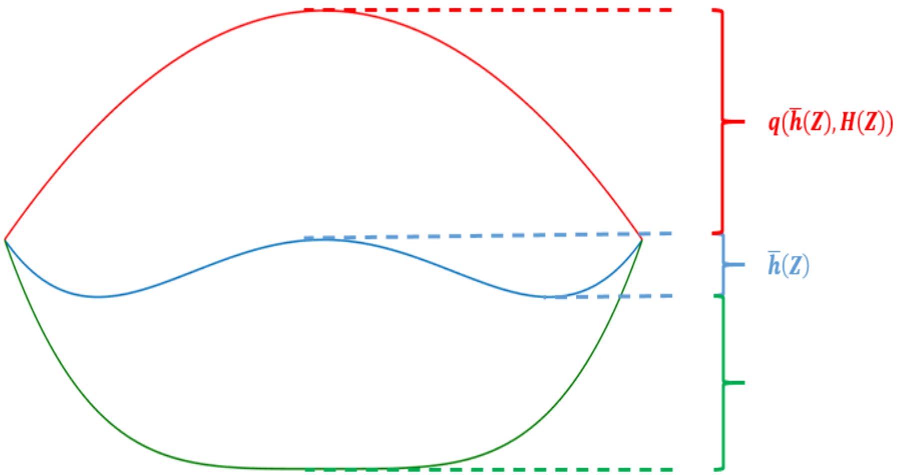
- ## Inclusion function

$h:\mathbb{R}^n\supset X\to\mathbb{R}$ continuous.

Image of $Z\subset X$ under h: $\overline{h}(Z):=\left[h^{\mathrm{L}}(Z),h^{\mathrm{U}}(Z)\right]$.

$H:\mathbb{I}X\supset\mathcal{X}\to\mathbb{IR}$ is an inclusion function for h on $\mathcal{X}$ if

$$\overline{h}(Z)\subset H(Z),\forall Z\in\mathcal{X}.$$



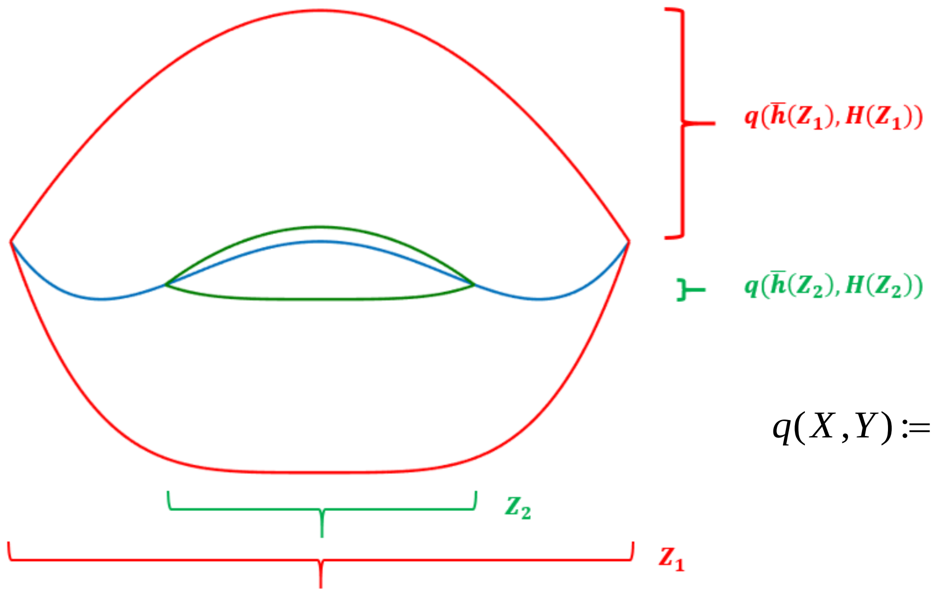
Hausdorff Convergence

- Hausdorff Convergence Order

$h:\mathbb{R}^n \supset X \to \mathbb{R}$ continuous, H inclusion function of h on $\mathbb{I}X$.

H has Hausdorff convergence of order $\beta > 0$ on X if $\exists \tau > 0$ s.t. $\forall Z \in \mathbb{I}X$,

$$q(\bar{h}(Z), H(Z)) \le \tau w(Z)^{\beta}.$$



$$q(X,Y) := \max\left\{\left|x^{\mathrm{L}} - y^{\mathrm{L}}\right|, \left|x^{\mathrm{U}} - y^{\mathrm{U}}\right|\right\}.$$

Bompadre, A. et al., J. Global Optim., 2012.

Pointwise Convergence

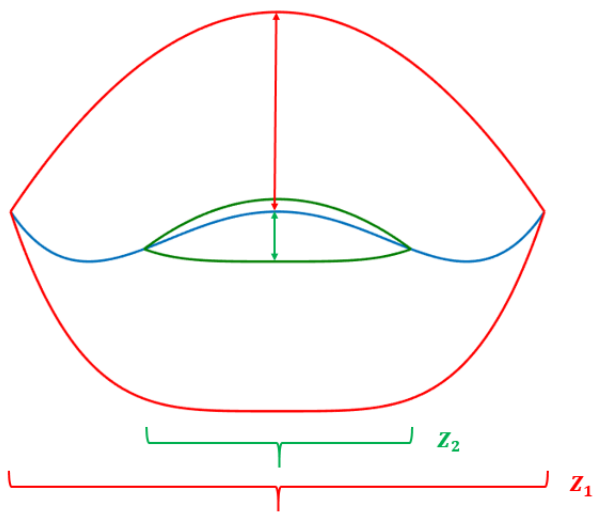
- Pointwise Convergence Order

$h:\mathbb{R}^n \supset X \to \mathbb{R}$ continuous, $(h_Z^{cv}, h_Z^{cc})\big|_{Z\in\mathbb{I}X}$ scheme of relaxations of h in X.

$(h_Z^{cv}, h_Z^{cc})\big|_{Z\in\mathbb{I}X}$ has pointwise convergence of order $\gamma > 0$ on X if $\exists\, \tau > 0$ s.t. $\forall Z \in \mathbb{I}X$,

$$\sup_{x\in Z}\left|h(x) - h_Z^{cv}(x)\right| \le \tau w(Z)^\gamma,$$

$$\sup_{x\in Z}\left|h(x) - h_Z^{cc}(x)\right| \le \tau w(Z)^\gamma.$$



Bompadre, A. et al., J. Global Optim., 2012.

Propagation of convergence orders

- γ-order pointwise convergence of a scheme of relaxations implies $(\gamma \leq)\beta$-order Hausdorff convergence of the scheme

Bompadre, A. et al., J. Global Optim., 2012.

Scholz, D., J. Global Optim., 2012.

Propagation of convergence orders

- γ-order pointwise convergence of a scheme of relaxations implies $(\gamma \leq)\beta$-order Hausdorff convergence of the scheme

- Envelopes and αBB relaxations have second-order pointwise convergence for C^2 functions

Bompadre, A. et al., J. Global Optim., 2012.

Scholz, D., J. Global Optim., 2012.

Propagation of convergence orders

- γ-order pointwise convergence of a scheme of relaxations implies $(\gamma \leq)\beta$-order Hausdorff convergence of the scheme
- Envelopes and αBB relaxations have second-order pointwise convergence for C^2 functions
- Natural interval extensions have first-order pointwise convergence for Lipschitz continuous functions
- Centered forms have second-order Hausdorff convergence for C^1 functions

Bompadre, A. et al., J. Global Optim., 2012.

Scholz, D., J. Global Optim., 2012.

Propagation of convergence orders

Convergence order of factors	Convergence order of operation result
Sum: $g(\mathbf{z}) = g_1(\mathbf{z}) + g_2(\mathbf{z})$	
Schemes for g_i have β_i	$\beta \geq 1$ (no order propagation)
Schemes for g_i have γ_i	$\gamma \geq \min\{\gamma_1, \gamma_2\}$
Product: $g(\mathbf{z}) = g_1(\mathbf{z}) \cdot g_2(\mathbf{z})$	
Schemes for g_i have β_i	$\beta \geq 1$ (no order propagation)
Schemes for g_i have γ_i	$\gamma \geq \min\{\gamma_1, \gamma_2, 2\}$
Composition: $g(\mathbf{z}) = F \circ f(\mathbf{z})$	
Scheme for F has β_F	$\beta \geq \min\{\beta_F, \beta_{f,T}\}$
Inclusion for f has $\beta_{f,T}$	
Scheme for F has γ_F	$\gamma \geq \min\{\gamma_F, \gamma_f\}$
Scheme for f has γ_f	

Bound on convergence order of McCormick estimators assuming Lipschitz continuity of the factors

Bompadre, A. et al., J. Global Optim., 2012.

Bompadre, A. et al., J. Global Optim., 2013.

More Definitions

- Distance between sets

Let $Y, Z \subset \mathbb{R}^n$.

The distance between Y and Z is defined as

$$d(Y, \mathrm{Z}) := \inf_{\substack{y \in Y, \\ z \in Z}} \left\| y - z \right\|.$$

More Definitions

- Distance between sets

Let $Y, Z \subset \mathbb{R}^n$.

The distance between Y and Z is defined as

$$d(Y, \mathrm{Z}) := \inf_{\substack{y \in Y, \\ z \in Z}} \|y - z\|.$$

- Convergence and Pointwise Convergence

$h : \mathbb{R}^n \supset X \to \mathbb{R}$ continuous, $(h_Z^{\text{cv}})\big|_{Z \in \mathbb{I}X}$ scheme of convex relaxations of h on X.

$(h_Z^{\text{cv}})\big|_{Z \in \mathbb{I}X}$ has convergence of order $\beta > 0$ on X if $\exists \tau > 0$ s.t. $\forall Z \in \mathbb{I}X$,

$$\inf_{x \in Z} h(x) - \inf_{x \in Z} h_Z^{\text{cv}}(x) \le \tau w(Z)^{\beta}.$$

$(h_Z^{\text{cv}})\big|_{Z \in \mathbb{I}X}$ has pointwise convergence of order $\gamma > 0$ on X if $\exists \tau > 0$ s.t. $\forall Z \in \mathbb{I}X$,

$$\sup_{x \in Z} \left| h(x) - h_Z^{\text{cv}}(x) \right| \le \tau w(Z)^{\gamma}.$$

Formulation

$$\begin{aligned} \min_{x \in X}\ & f(x) \\ \text{s.t. } & g(x) \le 0, \\ & h(x) = 0, \end{aligned}$$

where $X \subset \mathbb{R}^n$ is a nonempty compact convex set,

$f: X \to \mathbb{R}$, $g: X \to \mathbb{R}^{m_I}$, $h: X \to \mathbb{R}^{m_E}$ are continuous.

Definition of Convergence Order

- Convergence order of a lower bounding scheme

For any $Z \in \mathbb{I}X$, let $\mathcal{F}(Z) := \{x \in Z : g(x) \leq 0, h(x) = 0\}$ denote the feasible set of the problem with x restricted to Z.

Definition of Convergence Order

- Convergence order of a lower bounding scheme

For any $Z \in \mathbb{I}X$, let $\mathcal{F}(Z) := \{x \in Z : g(x) \le 0, h(x) = 0\}$ denote the feasible set of the problem with x restricted to Z.

Let $(\mathcal{L}(Z))\big|_{Z \in \mathbb{I}X}$ denote a scheme of lower bounding problems.

Associate with $(\mathcal{L}(Z))\big|_{Z \in \mathbb{I}X}$ a scheme of triples $(\mathcal{O}(Z), \mathcal{I}_I(Z), \mathcal{I}_E(Z))\big|_{Z \in \mathbb{I}X}$, where

$(\mathcal{O}(Z))\big|_{Z \in \mathbb{I}X}$ is a scheme of lower bounds,

$(\mathcal{I}_I(Z))\big|_{Z \in \mathbb{I}X}$ and $(\mathcal{I}_E(Z))\big|_{Z \in \mathbb{I}X}$ are schemes of subsets of $\mathbb{R}^{m_I}$ and $\mathbb{R}^{m_E}$, respectively, satisfying

$$d(\mathcal{I}_I(Z), \mathbb{R}_-^{m_I}) \le d(\overline{g}(Z), \mathbb{R}_-^{m_I}),$$

$$d(\mathcal{I}_E(Z), \{0\}) \le d(\overline{h}(Z), \{0\}), \text{ and}$$

$$\mathcal{O}(Z) = +\infty \Leftrightarrow d(\mathcal{I}_I(Z), \mathbb{R}_-^{m_I}) > 0 \text{ or } d(\mathcal{I}_E(Z), \{0\}) > 0, \ \forall Z \in \mathbb{I}X.$$

Definition of Convergence Order

- Convergence order of a lower bounding scheme

$(\mathcal{L}(Z))\big|_{Z \in \mathbb{I}X}$: lower bounding scheme.

$(\mathcal{O}(Z))\big|_{Z \in \mathbb{I}X}$: scheme of lower bounds.

$(\mathcal{I}_I(Z))\big|_{Z \in \mathbb{I}X}$: scheme estimating feasibility of inequality constraints.

$(\mathcal{I}_E(Z))\big|_{Z \in \mathbb{I}X}$: scheme estimating feasibility of equality constraints.

Definition of Convergence Order

- Convergence order of a lower bounding scheme

$(\mathcal{L}(Z))\big|_{Z\in\mathbb{I}X}$: lower bounding scheme.

$(\mathcal{O}(Z))\big|_{Z\in\mathbb{I}X}$: scheme of lower bounds.

$(\mathcal{I}_I(Z))\big|_{Z\in\mathbb{I}X}$: scheme estimating feasibility of inequality constraints.

$(\mathcal{I}_E(Z))\big|_{Z\in\mathbb{I}X}$: scheme estimating feasibility of equality constraints.

The lower bounding scheme $(\mathcal{L}(Z))\big|_{Z\in\mathbb{I}X}$ is said to have convergence of order $\beta > 0$ at

1. a feasible point $x \in X$ if $\exists \tau > 0$ s.t. $\forall Z \in \mathbb{I}X$ with $x \in Z$,

$$\min_{z\in\mathcal{F}(Z)} f(z) - \mathcal{O}(Z) \leq \tau w(Z)^{\beta}.$$

2. an infeasible point $x \in X$ if $\exists \overline{\tau} > 0$ s.t. $\forall Z \in \mathbb{I}X$ with $x \in Z$,

$$d(\overline{g}(Z), \mathbb{R}_{-}^{m_I}) - d(\mathcal{I}_I(Z), \mathbb{R}_{-}^{m_I}) \leq \overline{\tau} w(Z)^{\beta}, \text{ and}$$

$$d(\overline{h}(Z), \{0\}) - d(\mathcal{I}_E(Z), \{0\}) \leq \overline{\tau} w(Z)^{\beta}.$$

Definition of Convergence Order

- Convergence order of a lower bounding scheme

$(\mathcal{L}(Z))\big|_{Z\in\mathbb{I}X}$: lower bounding scheme.

$(\mathcal{O}(Z))\big|_{Z\in\mathbb{I}X}$: scheme of lower bounds.

$(\mathcal{I}_I(Z))\big|_{Z\in\mathbb{I}X}$: scheme estimating feasibility of inequality constraints.

$(\mathcal{I}_E(Z))\big|_{Z\in\mathbb{I}X}$: scheme estimating feasibility of equality constraints.

The lower bounding scheme $(\mathcal{L}(Z))\big|_{Z\in\mathbb{I}X}$ is said to have convergence of order $\beta > 0$ at

1. a feasible point $x \in X$ if $\exists \tau > 0$ s.t. $\forall Z \in \mathbb{I}X$ with $x \in Z$,

$$\min_{z\in\mathcal{F}(Z)} f(z) - \mathcal{O}(Z) \le \tau w(Z)^{\beta}.$$

2. an infeasible point $x \in X$ if $\exists \overline{\tau} > 0$ s.t. $\forall Z \in \mathbb{I}X$ with $x \in Z$,

$$d(\overline{g}(Z), \mathbb{R}_-^{m_I}) - d(\mathcal{I}_I(Z), \mathbb{R}_-^{m_I}) \le \overline{\tau} w(Z)^{\beta}, \text{ and}$$

$$d(\overline{h}(Z), \{0\}) - d(\mathcal{I}_E(Z), \{0\}) \le \overline{\tau} w(Z)^{\beta}.$$

The lower bounding scheme has convergence of order β on X if it has convergence of order (at least) β at each $x \in X$, with the constants $\tau, \overline{\tau}$ independent of x.

Convergence Order
Convex relaxation-based scheme

- Convergence order of a lower bounding scheme

Let $(f_Z^{\mathrm{cv}})\big|_{Z\in\mathbb{I}X}$ and $(g_Z^{\mathrm{cv}})\big|_{Z\in\mathbb{I}X}$ denote continuous schemes of convex relaxations of f and g in X, and let $(h_Z^{\mathrm{cv}}, h_Z^{\mathrm{cc}})\big|_{Z\in\mathbb{I}X}$ denote a continuous scheme of relaxations of h in X.

Convergence Order
Convex relaxation-based scheme

- Convergence order of a lower bounding scheme

Let $(f_Z^{\mathrm{cv}})\big|_{Z\in\mathbb{I}X}$ and $(g_Z^{\mathrm{cv}})\big|_{Z\in\mathbb{I}X}$ denote continuous schemes of convex relaxations of f and g in X, and let $(h_Z^{\mathrm{cv}}, h_Z^{\mathrm{cc}})\big|_{Z\in\mathbb{I}X}$ denote a continuous scheme of relaxations of h in X.

The convex relaxation-based lower bounding scheme is defined by

$$\begin{aligned}
\mathcal{O}(Z) &:= \min_{x\in Z} f_Z^{\mathrm{cv}}(x) \\
&\quad\ \text{s.t. } g_Z^{\mathrm{cv}}(x)\le 0, \\
&\quad\quad\ h_Z^{\mathrm{cv}}(x)\le 0, \\
&\quad\quad\ h_Z^{\mathrm{cc}}(x)\ge 0, \\
\mathcal{I}_I(Z) &:= \overline{g}_Z^{\mathrm{cv}}(Z), \\
\mathcal{I}_E(Z) &:= \left\{ w\in\mathbb{R}^{m_E} : h_Z^{cv}(z)\le w\le h_Z^{cc}(z) \text{ for some } z\in Z \right\}.
\end{aligned}$$

Convergence Order
Convex relaxation-based scheme

- Convergence order of a lower bounding scheme

Let $(f_Z^{\text{cv}})\big|_{Z\in\mathbb{I}X}$ and $(g_Z^{\text{cv}})\big|_{Z\in\mathbb{I}X}$ denote continuous schemes of convex relaxations of f and g in X, and let $(h_Z^{\text{cv}}, h_Z^{\text{cc}})\big|_{Z\in\mathbb{I}X}$ denote a continuous scheme of relaxations of h in X.

The convex relaxation-based lower bounding scheme is defined by

$$\begin{aligned}
\mathcal{O}(Z) := \ & \min_{x\in Z} f_Z^{\text{cv}}(x) \\
& \text{s.t. } g_Z^{\text{cv}}(x) \le 0, \\
& \qquad h_Z^{\text{cv}}(x) \le 0, \\
& \qquad h_Z^{\text{cc}}(x) \ge 0, \\
\mathcal{I}_I(Z) := \ & \overline{g}_Z^{\text{cv}}(Z), \\
\mathcal{I}_E(Z) := \ & \left\{ w \in \mathbb{R}^{m_E} : h_Z^{cv}(z) \le w \le h_Z^{cc}(z) \text{ for some } z \in Z \right\}.
\end{aligned}$$

For any $Z \in \mathbb{I}X$, let $\mathcal{F}^{\text{cv}}(Z) := \left\{ x \in Z : g_Z^{\text{cv}}(x) \le 0, h_Z^{\text{cv}}(x) \le 0, h_Z^{\text{cc}}(x) \ge 0 \right\}$ denote the feasible set of the convex relaxation-based lower bounding scheme with x restricted to Z.

Convergence Order
Convex relaxation-based scheme

- Convergence order of a lower bounding scheme

Let $\mathcal{F}(Z) := \{x \in Z : g(x) \le 0, h(x) = 0\}$,

$\mathcal{F}^{\text{cv}}(Z) := \{x \in Z : g_Z^{\text{cv}}(x) \le 0, h_Z^{\text{cv}}(x) \le 0, h_Z^{\text{cc}}(x) \ge 0\}$.

The convex relaxation-based lower bounding scheme is said to have convergence of order $\beta > 0$ at

1. a feasible point $x \in X$ if $\exists \tau \ge 0$ s.t. $\forall Z \in \mathbb{I}X$ with $x \in Z$,

$$\min_{z \in \mathcal{F}(Z)} f(z) - \min_{z \in \mathcal{F}^{\text{cv}}(Z)} f_Z^{\text{cv}}(z) \le \tau w(Z)^{\beta}.$$

2. an infeasible point $x \in X$ if $\exists \bar{\tau} \ge 0$ s.t. $\forall Z \in \mathbb{I}X$ with $x \in Z$,

$$d(\bar{g}(Z), \mathbb{R}_-^{m_I}) - d(\bar{g}_Z^{\text{cv}}(Z), \mathbb{R}_-^{m_I}) \le \bar{\tau} w(Z)^{\beta}, \text{ and}$$

$$d(\bar{h}(Z), \{0\}) - d(I_E(Z), \{0\}) \le \bar{\tau} w(Z)^{\beta},$$

where $(I_E(Z))\big|_{Z \in \mathbb{I}X}$ is defined as

$$(I_E(Z))\big|_{Z \in \mathbb{I}X} := \left(\{w \in \mathbb{R}^{m_E} : h_Z^{\text{cv}}(x) \le w \le h_Z^{\text{cc}}(x) \text{ for some } x \in Z\}\right)_{Z \in \mathbb{I}X}.$$

Convergence Order
Convex relaxation-based scheme

- Convergence order of a lower bounding scheme

Let $\mathcal{F}(Z) := \{x \in Z : g(x) \le 0, h(x) = 0\}$,

$\mathcal{F}^{\mathrm{cv}}(Z) := \{x \in Z : g_Z^{\mathrm{cv}}(x) \le 0, h_Z^{\mathrm{cv}}(x) \le 0, h_Z^{\mathrm{cc}}(x) \ge 0\}$.

The convex relaxation-based lower bounding scheme is said to have convergence of order $\beta > 0$ at

1. a feasible point $x \in X$ if $\exists \tau \ge 0$ s.t. $\forall Z \in \mathbb{I}X$ with $x \in Z$,

$$\min_{z \in \mathcal{F}(Z)} f(z) - \min_{z \in \mathcal{F}^{\mathrm{cv}}(Z)} f_Z^{\mathrm{cv}}(z) \le \tau w(Z)^{\beta}.$$

"The lower bound has to converge to the minimum objective value with order at least β"

2. an infeasible point $x \in X$ if $\exists \overline{\tau} \ge 0$ s.t. $\forall Z \in \mathbb{I}X$ with $x \in Z$,

$$d(\overline{g}(Z), \mathbb{R}_-^{m_I}) - d(\overline{g}_Z^{\mathrm{cv}}(Z), \mathbb{R}_-^{m_I}) \le \overline{\tau} w(Z)^{\beta}, \text{ and}$$

$$d(\overline{h}(Z), \{0\}) - d(I_E(Z), \{0\}) \le \overline{\tau} w(Z)^{\beta},$$

where $(I_E(Z))|_{Z \in \mathbb{I}X}$ is defined as

$$(I_E(Z))|_{Z \in \mathbb{I}X} := \left(\left\{w \in \mathbb{R}^{m_E} : h_Z^{\mathrm{cv}}(x) \le w \le h_Z^{\mathrm{cc}}(x) \text{ for some } x \in Z\right\}\right)_{Z \in \mathbb{I}X}.$$

Convergence Order
Convex relaxation-based scheme

- Convergence order of a lower bounding scheme

Let $\mathcal{F}(Z) := \{x \in Z : g(x) \le 0, h(x) = 0\}$,

$\mathcal{F}^{\text{cv}}(Z) := \{x \in Z : g_Z^{\text{cv}}(x) \le 0, h_Z^{\text{cv}}(x) \le 0, h_Z^{\text{cc}}(x) \ge 0\}$.

The convex relaxation-based lower bounding scheme is said to have convergence of order $\beta > 0$ at

1. a feasible point $x \in X$ if $\exists \tau \ge 0$ s.t. $\forall Z \in \mathbb{I}X$ with $x \in Z$,

$$\min_{z \in \mathcal{F}(Z)} f(z) - \min_{z \in \mathcal{F}^{\text{cv}}(Z)} f_Z^{\text{cv}}(z) \le \tau w(Z)^{\beta}.$$

"The lower bound has to converge to the minimum objective value with order at least β"

2. an infeasible point $x \in X$ if $\exists \overline{\tau} \ge 0$ s.t. $\forall Z \in \mathbb{I}X$ with $x \in Z$,

$$d(\overline{g}(Z), \mathbb{R}_{-}^{m_I}) - d(\overline{g}_Z^{\text{cv}}(Z), \mathbb{R}_{-}^{m_I}) \le \overline{\tau} w(Z)^{\beta}, \text{ and}$$

$$d(\overline{h}(Z), \{0\}) - d(I_E(Z), \{0\}) \le \overline{\tau} w(Z)^{\beta},$$

where $(I_E(Z))|_{Z \in \mathbb{I}X}$ is defined as

$$(I_E(Z))|_{Z \in \mathbb{I}X} := \left(\{w \in \mathbb{R}^{m_E} : h_Z^{\text{cv}}(x) \le w \le h_Z^{\text{cc}}(x) \text{ for some } x \in Z\}\right)_{Z \in \mathbb{I}X}.$$

"The image of constraint relaxations has to converge (in distance) to the image of the true constraints with order at least β"

Convergence Order at Infeasible Points

- Convergence order of a lower bounding scheme

$$
\begin{aligned}
g_1(x) &= -x^2 + 4x - 2, \\
g_2(x) &= -x^2 + 2x + 1, \\
g_1^{\mathrm{cv}}(x) &= -(x^{\mathrm{L}} + x^{\mathrm{U}})x + x^{\mathrm{L}}x^{\mathrm{U}} + 4x - 2, \\
g_2^{\mathrm{cv}}(x) &= -(x^{\mathrm{L}} + x^{\mathrm{U}})x + x^{\mathrm{L}}x^{\mathrm{U}} + 2x + 1.
\end{aligned}
$$

Convergence Order at Infeasible Points

- Convergence order of a lower bounding scheme

$$g_1(x) = -x^2 + 4x - 2,$$
$$g_2(x) = -x^2 + 2x + 1,$$
$$g_1^{\mathrm{cv}}(x) = -(x^{\mathrm{L}} + x^{\mathrm{U}})x + x^{\mathrm{L}}x^{\mathrm{U}} + 4x - 2,$$
$$g_2^{\mathrm{cv}}(x) = -(x^{\mathrm{L}} + x^{\mathrm{U}})x + x^{\mathrm{L}}x^{\mathrm{U}} + 2x + 1.$$

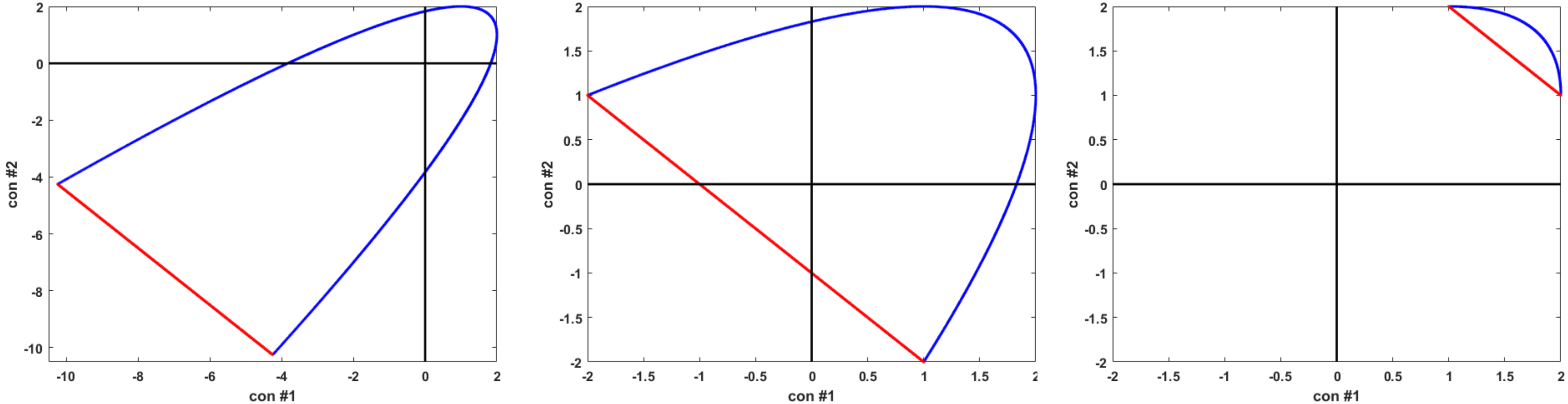
$$g_1(1.5) = \mathrm{g}_2(1.5) = 1.75.$$

Convergence Order at Infeasible Points

- Convergence order of a lower bounding scheme

$$g_1(x) = -x^2 + 4x - 2,$$
$$g_2(x) = -x^2 + 2x + 1,$$
$$g_1^{\mathrm{cv}}(x) = -(x^{\mathrm{L}} + x^{\mathrm{U}})x + x^{\mathrm{L}}x^{\mathrm{U}} + 4x - 2,$$
$$g_2^{\mathrm{cv}}(x) = -(x^{\mathrm{L}} + x^{\mathrm{U}})x + x^{\mathrm{L}}x^{\mathrm{U}} + 2x + 1.$$

$$g_1(1.5) = g_2(1.5) = 1.75.$$



Images on [-1.5,4.5] Images on [0,3] Images on [1,2]

Convergence Order at Infeasible Points

- Convergence order of a lower bounding scheme

$$g_1(x) = -x^2 + 4x - 2,$$

$$g_2(x) = -x^2 + 2x + 1,$$

$$g_1^{cv}(x) = -\max\left\{\left(x^{L}\right)^2, \left(x^{U}\right)^2\right\} + 4x - 2,$$

$$g_2^{cv}(x) = -\max\left\{\left(x^{L}\right)^2, \left(x^{U}\right)^2\right\} + 2x + 1.$$

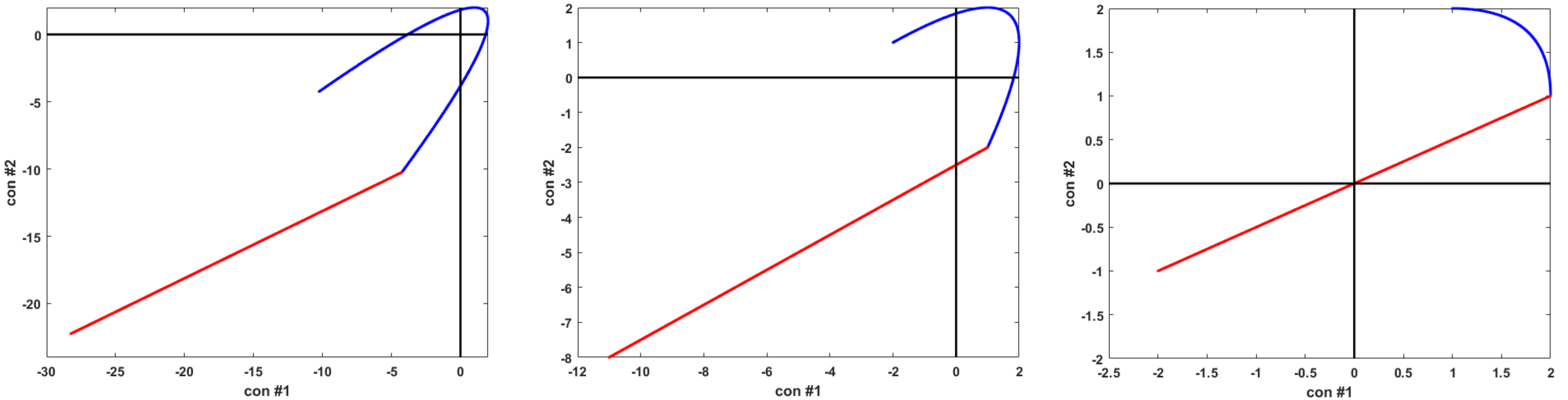
$$g_1(1.5) = g_2(1.5) = 1.75.$$

Convergence Order at Infeasible Points

- Convergence order of a lower bounding scheme

$$g_1(x) = -x^2 + 4x - 2,$$
$$g_2(x) = -x^2 + 2x + 1,$$
$$g_1^{cv}(x) = -\max\left\{\left(x^L\right)^2, \left(x^U\right)^2\right\} + 4x - 2,$$
$$g_2^{cv}(x) = -\max\left\{\left(x^L\right)^2, \left(x^U\right)^2\right\} + 2x + 1.$$

$$g_1(1.5) = g_2(1.5) = 1.75.$$



Images on [-1.5,4.5] Images on [0,3] Images on [1,2]

Conditions for first-order convergence

- Sufficient conditions for first-order convergence

Theorem: Suppose

1. $f, g_j, j = 1, \cdots, m_I,$ and $h_k, k = 1, \cdots, m_E,$ are Lipschitz continuous on X.
2. The schemes $(f_Z^{\mathrm{cv}})\big|_{Z \in \mathbb{I}X}, (g_{j,Z}^{\mathrm{cv}})\big|_{Z \in \mathbb{I}X}, j = 1, \cdots, m_I,$ and $(h_{k,Z}^{\mathrm{cv}}, h_{k,Z}^{\mathrm{cc}})\big|_{Z \in \mathbb{I}X}, k = 1, \cdots, m_E,$ are at least first-order pointwise convergent on X.

Then, the convex relaxation-based lower bounding scheme is at least first-order convergent on X.

Conditions for first-order convergence

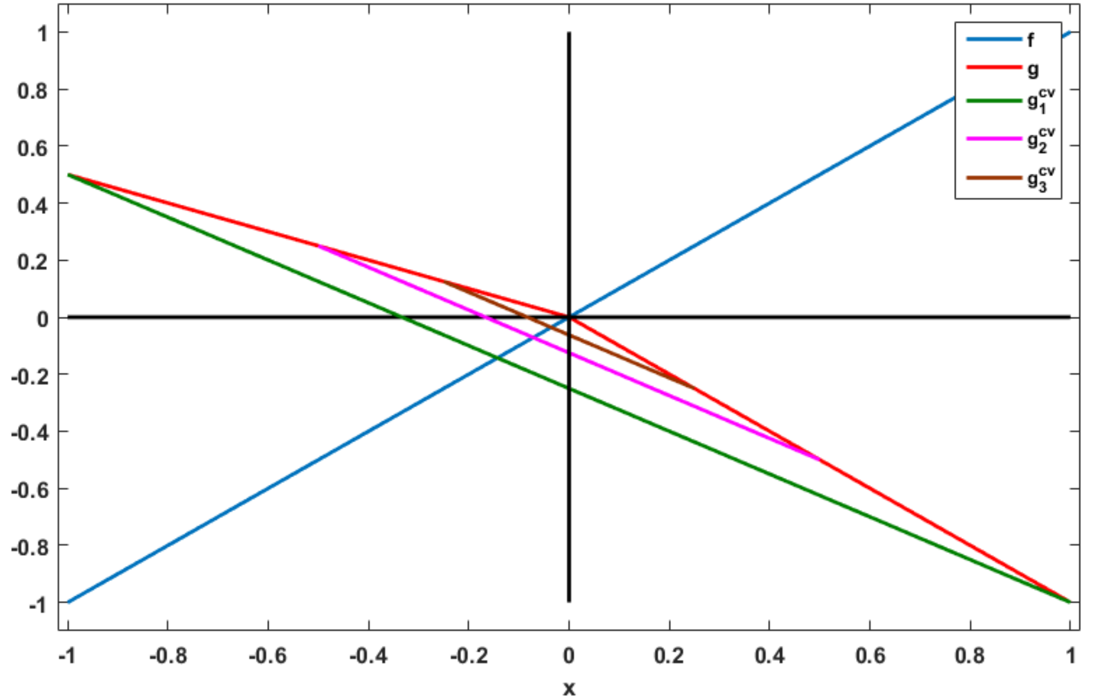
- Sufficient conditions for first-order convergence

Theorem: Suppose

1. $f, g_j, j = 1, \cdots, m_I$, and $h_k, k = 1, \cdots, m_E$, are Lipschitz continuous on X.

2. The schemes $(f_Z^{\text{cv}})\big|_{Z \in \mathbb{I}X}, (g_{j,Z}^{\text{cv}})\big|_{Z \in \mathbb{I}X}, j = 1, \cdots, m_I$, and $(h_{k,Z}^{\text{cv}}, h_{k,Z}^{\text{cc}})\big|_{Z \in \mathbb{I}X}, k = 1, \cdots, m_E$, are at least first-order pointwise convergent on X.

Then, the convex relaxation-based lower bounding scheme is at least first-order convergent on X.



Conditions for first-order convergence

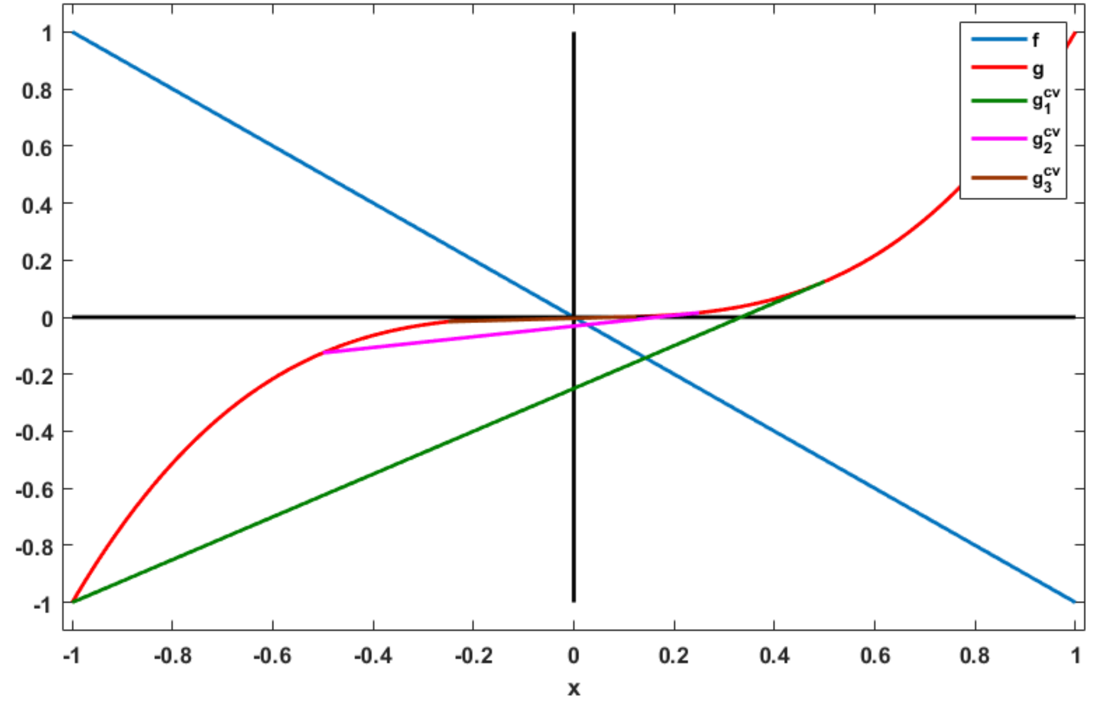
- Sufficient conditions for first-order convergence

Theorem: Suppose

1. $f, g_j, j=1,\cdots,m_I$, and $h_k, k=1,\cdots,m_E$, are Lipschitz continuous on X.

2. The schemes $(f_Z^{\mathrm{cv}})\big|_{Z\in\mathbb{I}X}, (g_{j,Z}^{\mathrm{cv}})\big|_{Z\in\mathbb{I}X}, j=1,\cdots,m_I$, and $(h_{k,Z}^{\mathrm{cv}}, h_{k,Z}^{\mathrm{cc}})\big|_{Z\in\mathbb{I}X}, k=1,\cdots,m_E$, are at least first-order pointwise convergent on X.

Then, the convex relaxation-based lower bounding scheme is at least first-order convergent on X.



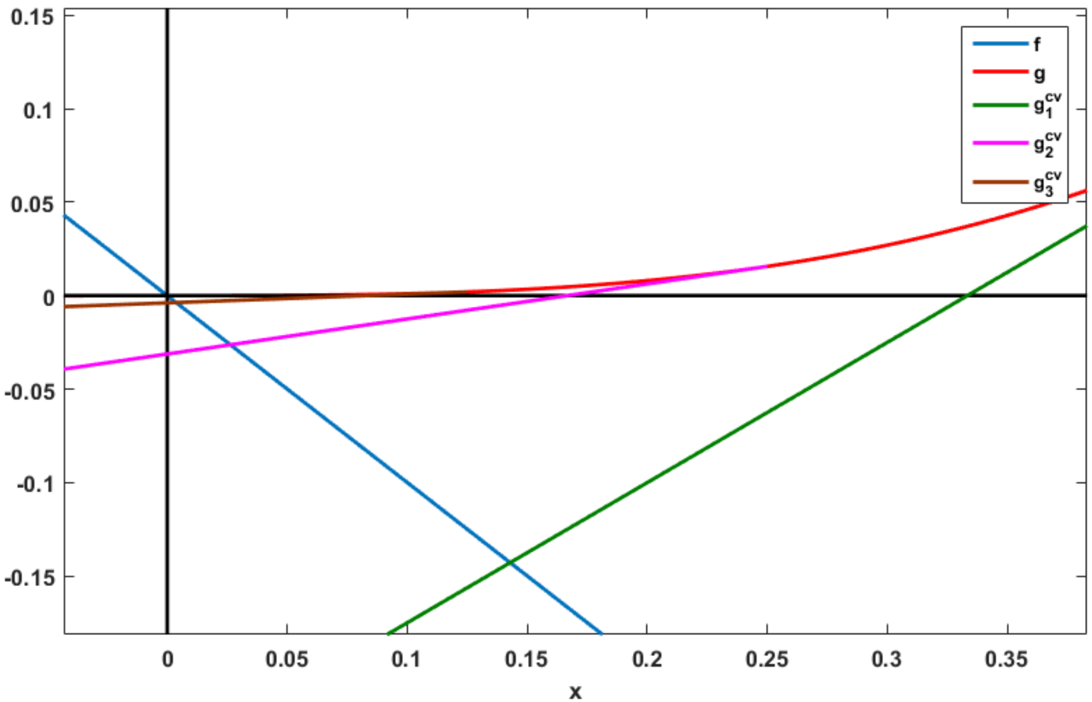
Conditions for first-order convergence

- Sufficient conditions for first-order convergence

Theorem: Suppose

1. $f, g_j, j=1,\cdots,m_I$, and $h_k, k=1,\cdots,m_E$, are Lipschitz continuous on X.
2. The schemes $(f_Z^{cv})\big|_{Z\in\mathbb{I}X}, (g_{j,Z}^{cv})\big|_{Z\in\mathbb{I}X}, j=1,\cdots,m_I$, and $(h_{k,Z}^{cv}, h_{k,Z}^{cc})\big|_{Z\in\mathbb{I}X}, k=1,\cdots,m_E$, are at least first-order pointwise convergent on X.

Then, the convex relaxation-based lower bounding scheme is at least first-order convergent on X.



Conditions for second-order convergence

- Sufficient conditions for second-order convergence

Theorem: Suppose

1. $f, g_j, j=1,\cdots,m_I$, and $h_k, k=1,\cdots,m_E$, are $\mathcal{C}^2$ on X.
2. The schemes $(f_Z^{\mathrm{cv}})\big|_{Z\in\mathbb{I}X}, (g_{j,Z}^{\mathrm{cv}})\big|_{Z\in\mathbb{I}X}, j=1,\cdots,m_I$, and $(h_{k,Z}^{\mathrm{cv}}, h_{k,Z}^{\mathrm{cc}})\big|_{Z\in\mathbb{I}X}, k=1,\cdots,m_E$, are at least second-order pointwise convergent on X.

Then, the convex relaxation-based lower bounding scheme is at least second-order convergent at

1. $x \in X$ for which $\exists(\mu,\lambda)\in\mathbb{R}_+^{m_I}\times\mathbb{R}^{m_E}$ such that (x,μ,λ) is a KKT point
2. $x \in X$ with $g(x) < 0$ (when $m_E = 0$)
3. infeasible $x \in X$

Conditions for second-order convergence

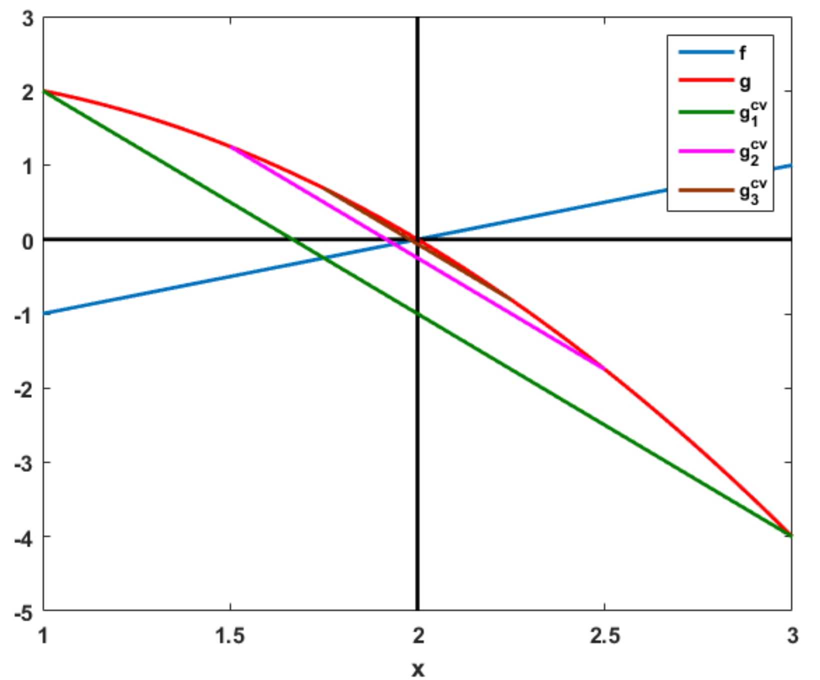
- Sufficient conditions for second-order convergence

Theorem: Suppose

1. $f, g_j, j=1,\cdots,m_I$, and $h_k, k=1,\cdots,m_E$, are $\mathcal{C}^2$ on X.
2. The schemes $(f_Z^{\mathrm{cv}})\big|_{Z\in\mathbb{I}X}, (g_{j,Z}^{\mathrm{cv}})\big|_{Z\in\mathbb{I}X}, j=1,\cdots,m_I$, and $(h_{k,Z}^{\mathrm{cv}}, h_{k,Z}^{\mathrm{cc}})\big|_{Z\in\mathbb{I}X}, k=1,\cdots,m_E$, are at least second-order pointwise convergent on X.

Then, the convex relaxation-based lower bounding scheme is at least second-order convergent at

1. $x\in X$ for which $\exists(\mu,\lambda)\in\mathbb{R}_+^{m_I}\times\mathbb{R}^{m_E}$ such that (x,μ,λ) is a KKT point
2. $x\in X$ with $g(x)<0$ (when $m_E=0$)
3. infeasible $x\in X$



Reduced-Space B&B Formulation

$$\min_{(x,y)\in X\times Y} f(x,y)$$
$$\text{s.t.} \quad g(x,y) \le 0,$$
$$h(x,y) = 0,$$

where $X \subset \mathbb{R}^{n_x}, Y \subset \mathbb{R}^{n_y}$ are nonempty compact convex sets,
$f : X \times Y \to \mathbb{R}$, $g : X \times Y \to \mathbb{R}^{m_I}$, $h : X \times Y \to \mathbb{R}^{m_E}$ are continuous, and
$f(\cdot, y)$ and $g(\cdot, y)$ are convex on X and $h(\cdot, y)$ is affine on X for each $y \in Y$.

- Some widely-applicable reduced-space B&B algorithms are
  - Dür's Lagrangian duality-based B&B algorithm (2001)
  - Epperly and Pistikopoulos' convex relaxation-based B&B algorithm for problems with special structures (1997)

Dür, M., Math. Program., 2001.

Epperly, T. G. W. and Pistikopoulos, E. N., J. Global Optim., 1997.

Reduced-Space Convergence Order

- Convergence order of a lower bounding scheme

For any $Z \in \mathbb{I}X$, let $\mathcal{F}(Z) := \{(x, y) \in X \times Z : g(x, y) \le 0, h(x, y) = 0\}$ denote the feasible set of the problem with y restricted to Z.

Let $(\mathcal{L}(Z))\big|_{Z \in \mathbb{I}Y}$ denote a scheme of lower bounding problems.

Associate with $(\mathcal{L}(Z))\big|_{Z \in \mathbb{I}Y}$ a scheme of triples $(\mathcal{O}(Z), \mathcal{I}_I(Z), \mathcal{I}_E(Z))\big|_{Z \in \mathbb{I}Y}$, where

$(\mathcal{O}(Z))\big|_{Z \in \mathbb{I}Y}$ is a scheme of lower bounds,

$(\mathcal{I}_I(Z))\big|_{Z \in \mathbb{I}Y}$ and $(\mathcal{I}_E(Z))\big|_{Z \in \mathbb{I}Y}$ are schemes of subsets of $\mathbb{R}^{m_I}$ and $\mathbb{R}^{m_E}$, respectively, satisfying

$$d(\mathcal{I}_I(Z), \mathbb{R}_-^{m_I}) \le d(\overline{g}(X \times Z), \mathbb{R}_-^{m_I}),$$

$$d(\mathcal{I}_E(Z), \{0\}) \le d(\overline{h}(X \times Z), \{0\}), \text{ and}$$

$$\mathcal{O}(Z) = +\infty \Leftrightarrow d(\mathcal{I}_I(Z), \mathbb{R}_-^{m_I}) > 0 \text{ or } d(\mathcal{I}_E(Z), \{0\}) > 0, \ \forall Z \in \mathbb{I}Y.$$

Reduced-Space Convergence Order

- Convergence order of a lower bounding scheme

$(\mathcal{L}(Z))\big|_{Z\in\mathbb{I}Y}$: lower bounding scheme.

$(\mathcal{O}(Z))\big|_{Z\in\mathbb{I}Y}$: scheme of lower bounds.

$(\mathcal{I}_I(Z))\big|_{Z\in\mathbb{I}Y}$: scheme estimating feasibility of inequality constraints.

$(\mathcal{I}_E(Z))\big|_{Z\in\mathbb{I}Y}$: scheme estimating feasibility of equality constraints.

The lower bounding scheme $(\mathcal{L}(Z))\big|_{Z\in\mathbb{I}Y}$ is said to have convergence of order $\beta > 0$ at

1. a feasible point $y \in Y$ if $\exists \tau > 0$ s.t. $\forall Z \in \mathbb{I}Y$ with $y \in Z$,

$$\min_{(x,z)\in\mathcal{F}(Z)} f(\mathrm{x}, z) - \mathcal{O}(Z) \le \tau w(Z)^{\beta}.$$

2. an infeasible point $y \in Y$ if $\exists \overline{\tau} > 0$ s.t. $\forall Z \in \mathbb{I}Y$ with $y \in Z$,

$$d(\overline{g}(X\times Z), \mathbb{R}_-^{m_I}) - d(\mathcal{I}_I(Z), \mathbb{R}_-^{m_I}) \le \overline{\tau} w(Z)^{\beta}, \text{ and}$$

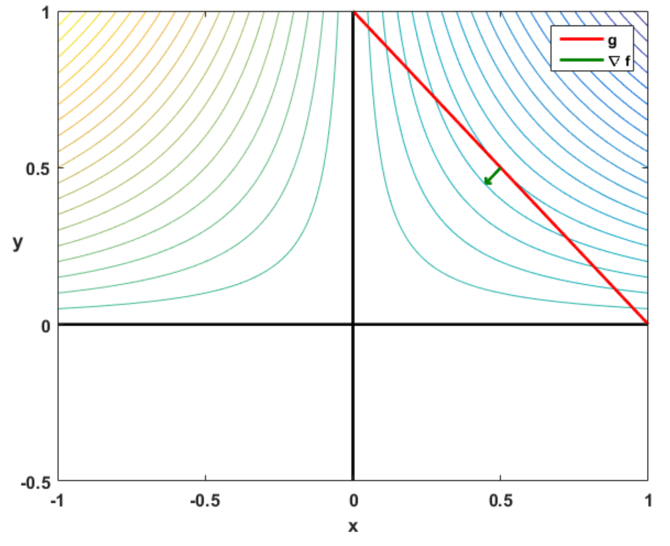
$$d(\overline{h}(X\times Z), \{0\}) - d(\mathcal{I}_E(Z), \{0\}) \le \overline{\tau} w(Z)^{\beta}.$$

The lower bounding scheme has convergence of order β on Y if it has convergence of order (at least) β at each $y \in Y$, with the constants $\tau, \overline{\tau}$ independent of y.

An argument for constraint propagation in reduced-space B&B algorithms

- Consider solving

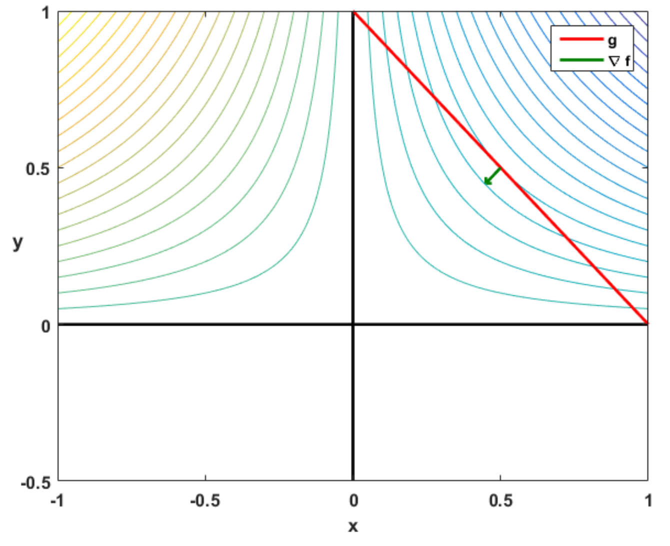
$$\begin{aligned} &\min_{x,y} \; -xy \\ &\text{s.t. } x + y \le 1, \\ &\quad x \in [-1,1],\; y \in [0,1]. \end{aligned}$$



An argument for constraint propagation in reduced-space B&B algorithms

- Consider solving

$$\begin{aligned}\min_{x,y}\ & -xy\\ \text{s.t.}\ & x+y\le 1,\\ & x\in[-1,1],\ y\in[0,1].\end{aligned}$$



via

$$\begin{aligned}\min_{y}\ & v(y)\\ \text{s.t.}\ & y\in[0,1].\end{aligned}$$

where

$$\begin{aligned}v(y)=\min_{x}\ & -xy\\ \text{s.t.}\ & x\le 1-y,\\ & x\in[-1,1].\end{aligned}$$

An argument for constraint propagation in reduced-space B&B algorithms

- Consider solving

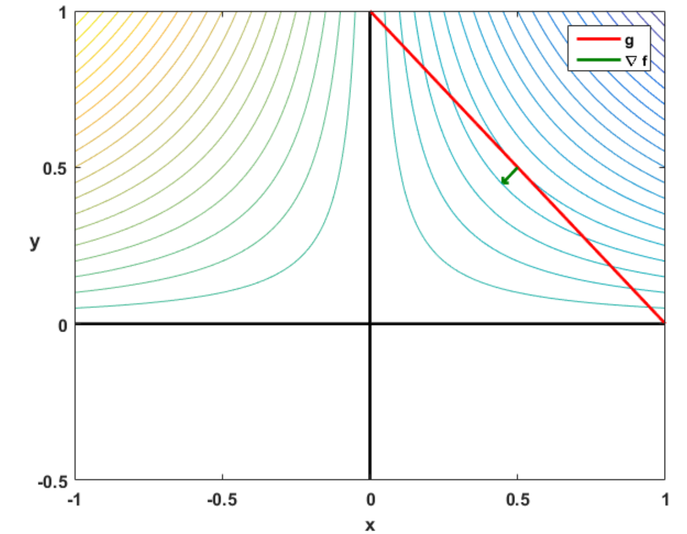
$$\min_{x,y} \ -xy$$
$$\text{s.t. } x + y \le 1,$$
$$x \in [-1,1],\ y \in [0,1].$$

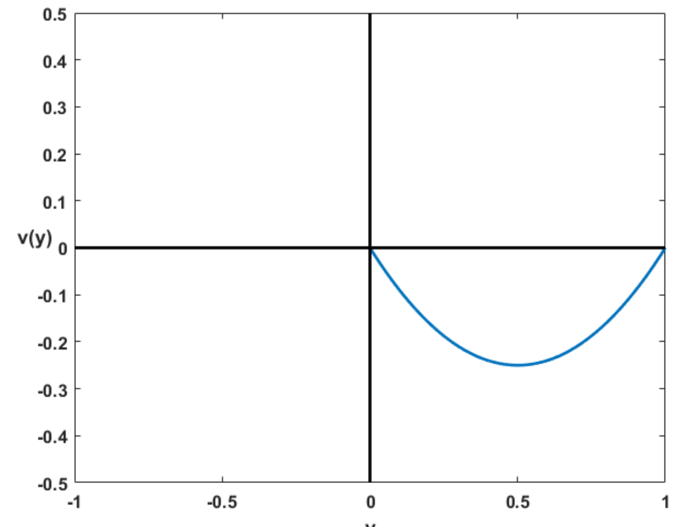
via

$$\min_{y} \ v(y)$$
$$\text{s.t. } y \in [0,1].$$

where

$$v(y) = \min_{x} \ -xy$$
$$\text{s.t. } x \le 1 - y,$$
$$x \in [-1,1].$$





An argument for constraint propagation in reduced-space B&B algorithms

◆ Consider solving

$$\begin{aligned}\min_{x,y} \quad & -xy \\ \text{s.t.} \quad & x+y \le 1, \\ & x \in [-1,1],\ y \in [0,1].\end{aligned}$$

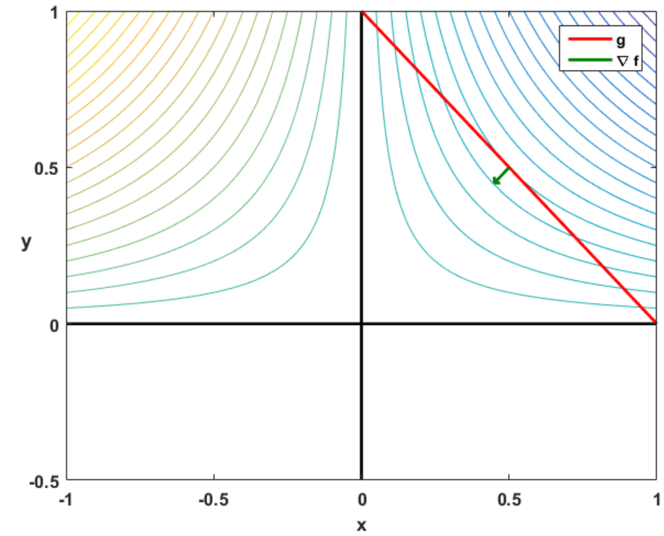
via

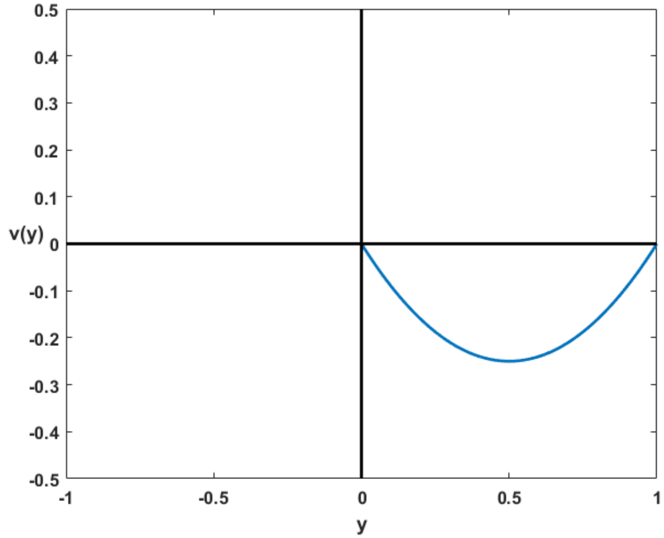
$$\begin{aligned}\min_{y} \quad & v(y) \\ \text{s.t.} \quad & y \in [0,1].\end{aligned}$$

where

$$\begin{aligned}v(y) = \min_{x} \quad & -xy \\ \text{s.t.} \quad & x \le 1-y, \\ & x \in [-1,1].\end{aligned}$$

The algorithms of Epperly and Pistikopoulos and Dür are first-order convergent on Y in the absence of constraint propagation, and will result in clustering.





Issues in reduced-space B&B algorithms

Consider the unconstrained problem

$$\begin{aligned}\min_{x,y}\ & 2x^2 + x^2 y - xy^2 + (y-0.5)^2 \\ \text{s.t.}\ & x \in [-1,1],\ y \in [0,1]\end{aligned}$$

and the corresponding reduced-space lower bounding scheme (Epperly and Pistikopoulos)

$$\begin{aligned}\min_{x,y}\ & 2x^2 + w_1 + w_2 + (y-0.5)^2 \\ \text{s.t.}\ & w_1 \ge x^2 y^{\mathrm{L}}, \qquad w_1 \ge y + x^2 y^{\mathrm{U}} - y^{\mathrm{U}}, \\ & w_2 \ge y^2 - x\left(y^{\mathrm{U}}\right)^2 - \left(y^{\mathrm{U}}\right)^2, \qquad w_2 \ge -\left(y^{\mathrm{L}}\right)^2 - \left(y^{\mathrm{U}} + y^{\mathrm{L}}\right)\left(y - y^{\mathrm{L}}\right) - x\left(y^{\mathrm{L}}\right)^2 + \left(y^{\mathrm{L}}\right)^2, \\ & x \in [-1,1], y \in [y^{\mathrm{L}}, y^{\mathrm{U}}].\end{aligned}$$

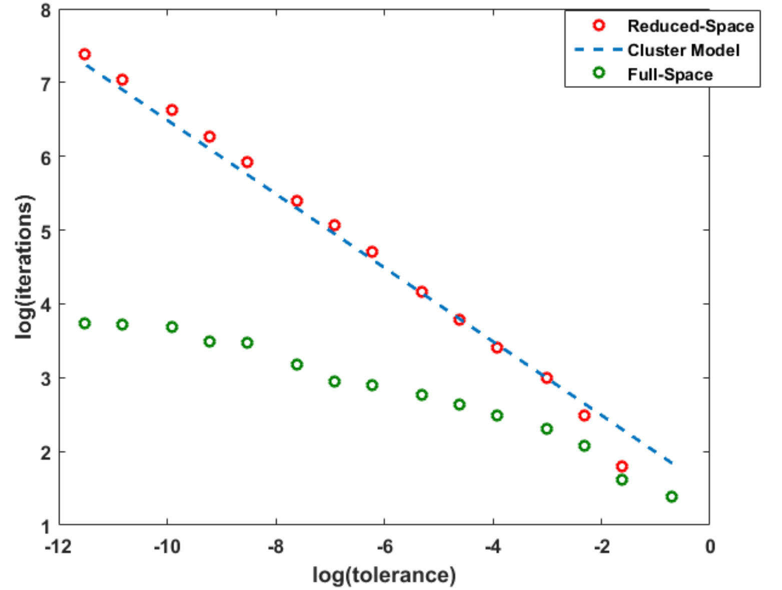
Issues in reduced-space B&B algorithms

Consider the unconstrained problem

$$\min_{x,y}\ 2x^2 + x^2 y - xy^2 + (y-0.5)^2$$
$$\text{s.t. } x \in [-1,1],\ y \in [0,1]$$

and the corresponding reduced-space lower bounding scheme (Epperly and Pistikopoulos)

$$\min_{x,y}\ 2x^2 + w_1 + w_2 + (y-0.5)^2$$
$$\text{s.t. } w_1 \ge x^2 y^{\mathrm{L}}, \qquad w_1 \ge y + x^2 y^{\mathrm{U}} - y^{\mathrm{U}},$$
$$w_2 \ge y^2 - x\left(y^{\mathrm{U}}\right)^2 - \left(y^{\mathrm{U}}\right)^2, \qquad w_2 \ge -\left(y^{\mathrm{L}}\right)^2 - \left(y^{\mathrm{U}} + y^{\mathrm{L}}\right)\left(y - y^{\mathrm{L}}\right) - x\left(y^{\mathrm{L}}\right)^2 + \left(y^{\mathrm{L}}\right)^2,$$
$$x \in [-1,1], y \in [y^{\mathrm{L}}, y^{\mathrm{U}}].$$



Summary

- Illustrated the cluster problem (or lack thereof) in constrained optimization as motivation for convergence order analysis
- Proposed a notion of convergence order for lower bounding schemes for constrained problems
- Established sufficient conditions for first-order and second-order convergence of convex relaxation-based lower bounding schemes
- Highlighted limitations in widely applicable reduced-space branch-and-bound algorithms
  - Demonstrated the importance of constraint propagation towards mitigating the cluster problem

Acknowledgements

- The Barton lab