Convexification Techniques for Stationary/Dynamic Global Optimization and Set-Based Computing

R. Paulen

Slovak University of Technology in Bratislava

in cooperation with

B. Chachuat
Imperial College London

M. Villanueva
IMT School for Advanced Studies Lucca

B. Houska
ShanghaiTech University

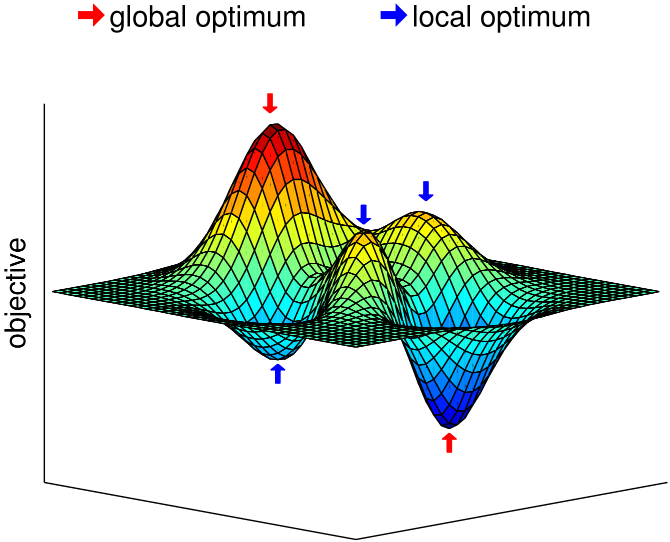
Slides

FrontSeat

Funded by
the European Union

Subscribe

Why Global Optimization?



Outline

Spatial Branch-and-Bound (B&B) Algorithm

Convex Relaxation of Factorable Programs

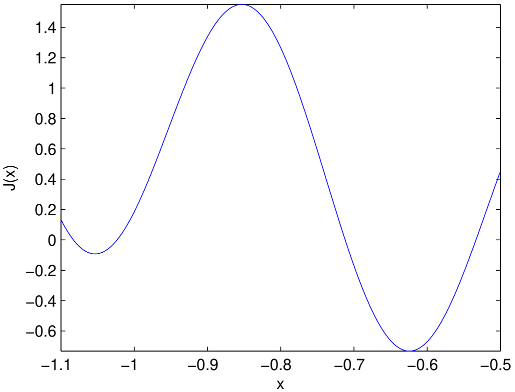
Bounding of Parametric ODEs

Guaranteed Parameter Estimation of Nonlinear Dynamic Systems

Spatial Branch-and-Bound (B&B) Algorithm

B&B algorithm steps

- **Initialization**
- Branching
- Finding upper bound
- Finding lower bound
- Fathoming
- Convergence check



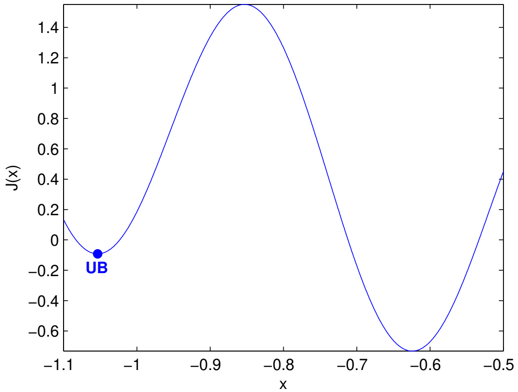
Initialization

Initialize region of interest R.

Spatial Branch-and-Bound (B&B) Algorithm

B&B algorithm steps

- Initialization
- Branching
- **Finding upper bound**
- Finding lower bound
- Fathoming
- Convergence check



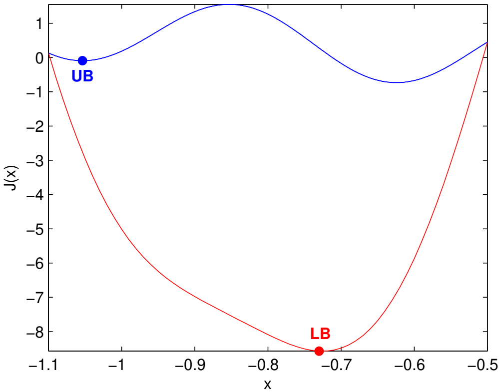
Finding upper bound

Solve original problem – find upper bound (UB).

Spatial Branch-and-Bound (B&B) Algorithm

B&B algorithm steps

- Initialization
- Branching
- Finding upper bound
- **Finding lower bound**
- Fathoming
- Convergence check



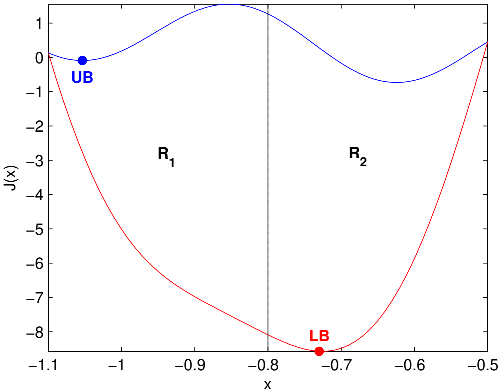
Finding lower bound

Solve relaxed problem – find lower bound (LB).

Spatial Branch-and-Bound (B&B) Algorithm

B&B algorithm steps

- Initialization
- **Branching**
- Finding upper bound
- Finding lower bound
- Fathoming
- Convergence check



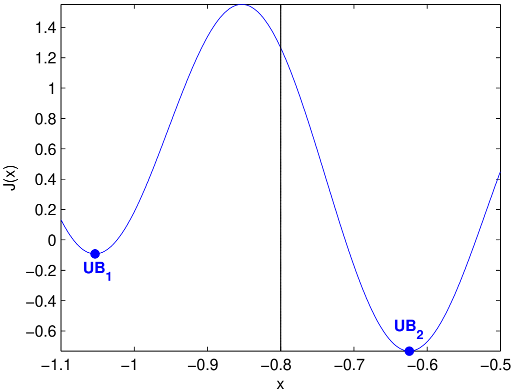
Branching

Partition R to new subregions R_1 and R_2.

Spatial Branch-and-Bound (B&B) Algorithm

B&B algorithm steps

- Initialization
- Branching
- **Finding upper bound**
- Finding lower bound
- Fathoming
- Convergence check



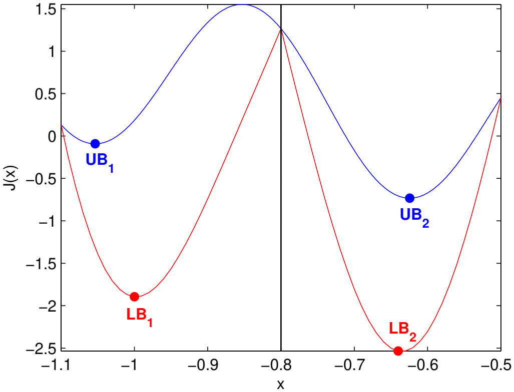
Finding upper bound

Solve original problem – find upper bound (UB).

Spatial Branch-and-Bound (B&B) Algorithm

B&B algorithm steps

- Initialization
- Branching
- Finding upper bound
- **Finding lower bound**
- Fathoming
- Convergence check



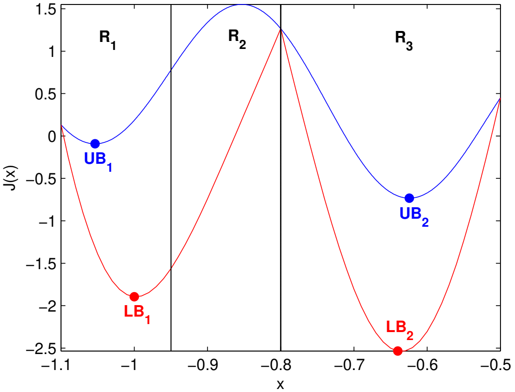
Finding lower bound

Solve relaxed problem – find lower bound (LB).

Spatial Branch-and-Bound (B&B) Algorithm

B&B algorithm steps

- Initialization
- **Branching**
- Finding upper bound
- Finding lower bound
- Fathoming
- Convergence check



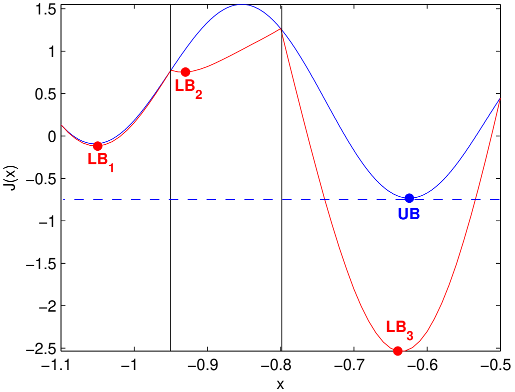
Branching

Partition R to new subregions R_1 and R_2.

Spatial Branch-and-Bound (B&B) Algorithm

B&B algorithm steps

- Initialization
- Branching
- Finding upper bound
- **Finding lower bound**
- Fathoming
- Convergence check



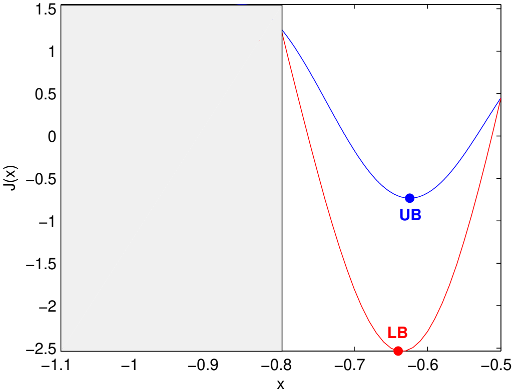
Finding lower bound

Solve relaxed problem – find lower bound (LB).

Spatial Branch-and-Bound (B&B) Algorithm

B&B algorithm steps

- Initialization
- Branching
- Finding upper bound
- Finding lower bound
- **Fathoming**
- Convergence check



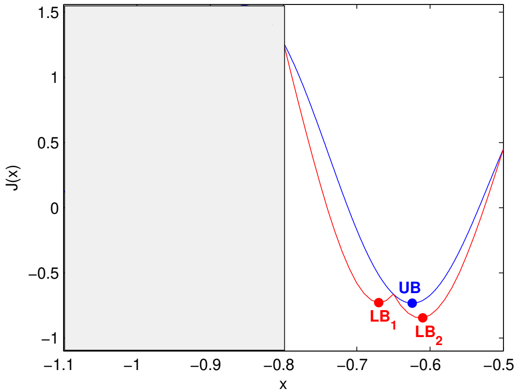
Fathoming

Remove subregions where solution cannot lie.

Spatial Branch-and-Bound (B&B) Algorithm

B&B algorithm steps

- Initialization
- Branching
- Finding upper bound
- **Finding lower bound**
- Fathoming
- Convergence check



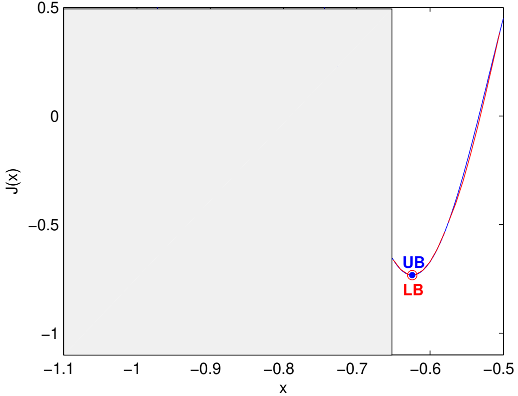
Finding lower bound

Solve relaxed problem – find lower bound (LB).

Spatial Branch-and-Bound (B&B) Algorithm

B&B algorithm steps

- Initialization
- Branching
- Finding upper bound
- Finding lower bound
- Fathoming
- **Convergence check**



Convergence check

Evaluate convergence criterion $\mathbf{UB} - \mathbf{LB} \leq \varepsilon$.

Outline

Spatial Branch-and-Bound (B&B) Algorithm

Convex Relaxation of Factorable Programs

Bounding of Parametric ODEs

Guaranteed Parameter Estimation of Nonlinear Dynamic Systems

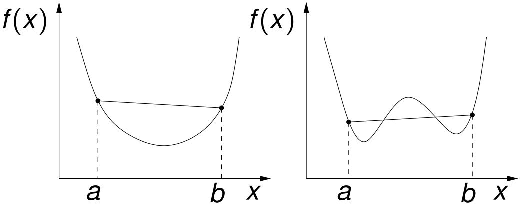
Convex Relaxation of an Optimization Problem

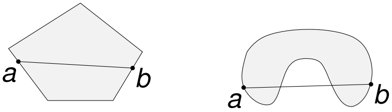
UB: $\min_x f(x)$ s.t. $g(x) \leq 0$

LB: $\min_x f^{\mathrm{cv}}(x)$ s.t. $g^{\mathrm{cv}}(x) \leq 0$

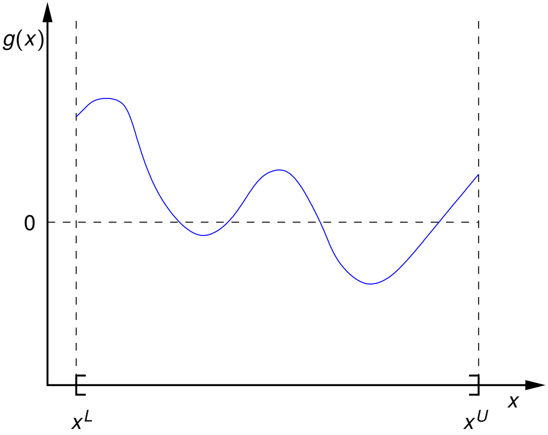
Lower bounding problem of spatial B&B consists of

- convex relaxation of objective function
- convex relaxation of the feasible set (constraint functions)





Convex Relaxation of the Feasible Set



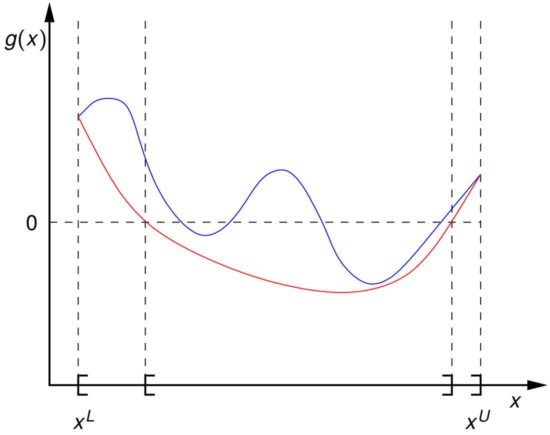
Constraint:

$$g(x) \leq 0$$

Convex relaxation:

$$g^{\mathrm{cv}}(x) \leq 0$$

Convex Relaxation of the Feasible Set



Constraint:

$$g(x) \leq 0$$

Convex relaxation:

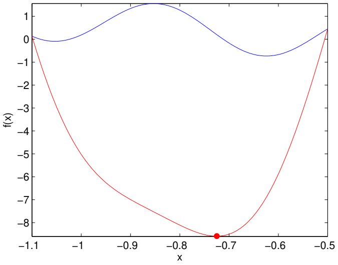
$$g^{\mathrm{cv}}(x) \leq 0$$

Convex Relaxation

Definition

Given a real-valued function $f(x)$ and interval $X = [x^L, x^U]$, identify $f^{\text{cv}}(x)$ such that:

a. $f^{\text{cv}}(x)$ underestimates $f(x)$ ($f^{\text{cv}}(x) \leq f(x),\ \forall x \in X$)
b. $f^{\text{cv}}(x)$ is convex ($\nabla^2 f^{\text{cv}}(x) \geq 0,\ \forall x \in X$)



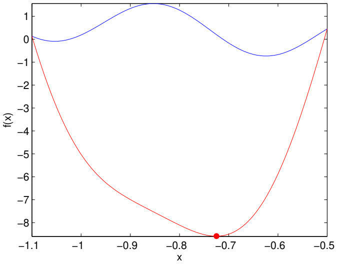
Convex Relaxation

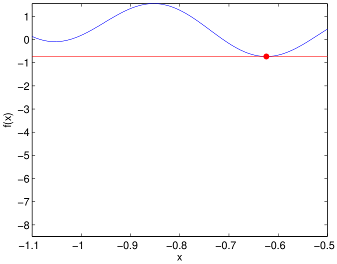
Definition

Given a real-valued function $f(x)$ and interval $X = [x^L, x^U]$, identify $f^{\mathrm{cv}}(x)$ such that:

a. $f^{\mathrm{cv}}(x)$ underestimates $f(x)$ ($f^{\mathrm{cv}}(x) \leq f(x), \ \forall x \in X$)

b. $f^{\mathrm{cv}}(x)$ is convex ($\nabla^2 f^{\mathrm{cv}}(x) \geq 0, \ \forall x \in X$)

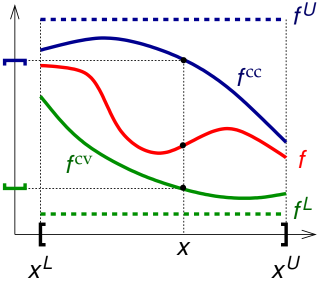




Convex/Concave Relaxations of Factorable Functions

Given a function f, construct functions f^{cv}, f^{cc} such that:

- $f^{\mathrm{cv}}(x) \leq f(x) \leq f^{\mathrm{cc}}(x), \forall x \in X$
- $f^{\mathrm{cv}} : X \to \mathbb{R}$ convex
- $f^{\mathrm{cc}} : X \to \mathbb{R}$ concave

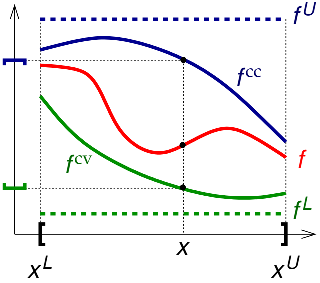


What are the sources of nonconvexity?

Convex/Concave Relaxations of Factorable Functions

Given a function f, construct functions f^{cv}, f^{cc} such that:

- $f^{\text{cv}}(x) \leq f(x) \leq f^{\text{cc}}(x), \forall x \in X$
- $f^{\text{cv}} : X \to \mathbb{R}$ convex
- $f^{\text{cc}} : X \to \mathbb{R}$ concave



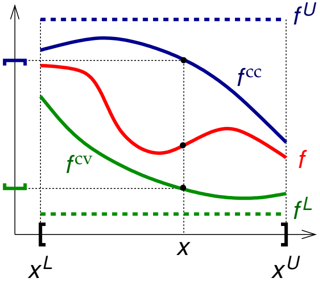
What are the sources of nonconvexity?

$$f(x) = \left(\exp(x_1) - x_2^2\right) x_1 x_2 \xrightarrow[\text{form}]{\text{factored}} \begin{cases} v_1(x) = x_1 \\ v_2(x) = x_2 \\ v_3(x) = \exp(v_1(x)) \\ v_4(x) = -v_2(x)^2 \\ v_5(x) = v_3(x) + v_4(x) \\ v_6(x) = v_5(x) v_1(x) \\ f(x) = v_6(x) v_2(x) \end{cases}$$

Convex/Concave Relaxations of Factorable Functions

Given a function f, construct functions f^{cv}, f^{cc} such that:

- $f^{\text{cv}}(x) \leq f(x) \leq f^{\text{cc}}(x), \forall x \in X$
- $f^{\text{cv}} : X \to \mathbb{R}$ convex
- $f^{\text{cc}} : X \to \mathbb{R}$ concave



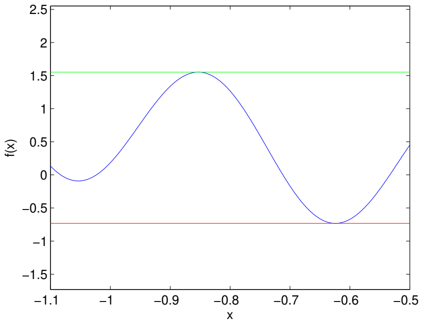
What are the sources of nonconvexity?

$$f(x) = \left(\exp(x_1) - x_2^2\right) x_1 x_2 \xrightarrow[\text{form}]{\text{factored}} \begin{cases} v_1(x) = x_1 \\ v_2(x) = x_2 \\ \boxed{v_3(x) = \exp(v_1(x))} \\ \boxed{v_4(x) = -v_2(x)^2} \\ v_5(x) = v_3(x) + v_4(x) \\ \boxed{v_6(x) = v_4(x) v_1(x)} \\ \boxed{f(x) = v_6(x) v_2(x)} \end{cases}$$

Interval Arithmetics

Set of simple rules:

$$
\begin{aligned}
[a,b]+[c,d] &= [a+c, b+d] \\
-[a,b] &= [-b,-a] \\
[a,b][c,d] &= [\min\{ac,ad,bc,bd\}, \\
&\quad \max\{ac,ad,bc,bd\}] \\
f([a,b]) &= [\min_{[a,b]} f, \max_{[a,b]} f]
\end{aligned}
$$



But...let's consider the following example

$$f(x) = x^2 - x = \left(x - \frac{1}{2}\right)^2 - \frac{1}{4}$$

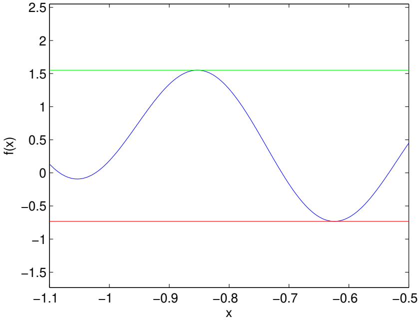
$$x \in X = [0,1]$$

$$f(X) = ?$$

Interval Arithmetics

Set of simple rules:

$$
\begin{aligned}
[a,b]+[c,d] &= [a+c, b+d] \\
-[a,b] &= [-b,-a] \\
[a,b][c,d] &= [\min\{ac,ad,bc,bd\}, \\
&\quad \max\{ac,ad,bc,bd\}] \\
f([a,b]) &= [\min_{[a,b]} f, \max_{[a,b]} f]
\end{aligned}
$$



But...let's consider the following example

$$f(x) = x^2 - x = \left(x - \frac{1}{2}\right)^2 - \frac{1}{4}$$

$$x \in X = [0,1]$$

$$f(X) = ?$$

Convex Relaxation using αBB Method

Find $f^{\text{cv}}(x)$ by subtracting a 'convex-enough' function from $f(x)$:

$$f^{\text{cv}}(x) = f(x) - \alpha(x - x^L)(x - x^U)$$

if $\alpha \geq 0$ then $f^{\text{cv}}(x) \leq f(x),\ \forall x \in X$

if $\alpha = -\min(0, (1/2)\lambda_{\text{min}}([\nabla^2 f(X)]))$

then $\nabla^2 f^{\text{cv}}(x) = \nabla^2 f(x) + 2I\alpha \geq 0,\ \forall x \in X$

Illustrative example:

$$f(x) = x\sin(x),\ \forall x \in [0, 2\pi]$$

$$[\nabla_x^2 f(X)] = 2\cos(X) - X\sin(X)$$

(interval arithmetics)

Adjiman, et al. *Computers & Chemical Engineering* 22 (9), 1998.

Convex Relaxation using αBB Method

Find $f^{\text{cv}}(x)$ by subtracting a 'convex-enough' function from $f(x)$:

$$f^{\text{cv}}(x) = f(x) - \alpha(x - x^L)(x - x^U)$$

$$\text{if } \alpha \geq 0 \text{ then } f^{\text{cv}}(x) \leq f(x), \ \forall x \in X$$

$$\text{if } \alpha = -\min(0, (1/2)\lambda_{\text{min}}([\nabla^2 f(X)]))$$

$$\text{then } \nabla^2 f^{\text{cv}}(x) = \nabla^2 f(x) + 2I\alpha \geq 0, \ \forall x \in X$$

Illustrative example:

$$f(x) = x\sin(x), \ \forall x \in [0, 2\pi]$$

$$[\nabla_x^2 f(X)] = 2\cos(X) - X\sin(X)$$

(interval arithmetics)

Adjiman, et al. *Computers & Chemical Engineering* 22 (9), 1998.

Convex Relaxation using αBB Method

Find $f^{\text{cv}}(x)$ by subtracting a 'convex-enough' function from $f(x)$:

$$f^{\text{cv}}(x) = f(x) - \alpha(x - x^L)(x - x^U)$$
$$\text{if } \alpha \geq 0 \text{ then } f^{\text{cv}}(x) \leq f(x), \ \forall x \in X$$
$$\text{if } \alpha = -\min(0, (1/2)\lambda_{\text{min}}([\nabla^2 f(X)]))$$
$$\text{then } \nabla^2 f^{\text{cv}}(x) = \nabla^2 f(x) + 2I\alpha \geq 0, \ \forall x \in X$$

Illustrative example:

$$f(x) = x\sin(x), \ \forall x \in [0, 2\pi]$$
$$[\nabla_x^2 f(X)] = 2\cos(X) - X\sin(X)$$
(interval arithmetics)

Adjiman, et al. *Computers & Chemical Engineering* 22 (9), 1998.

Convex Relaxation using αBB Method

Find $f^{\mathrm{cv}}(x)$ by subtracting a 'convex-enough' function from $f(x)$:

$$f^{\mathrm{cv}}(x) = f(x) - \alpha(x - x^L)(x - x^U)$$

$$\text{if } \alpha \geq 0 \text{ then } f^{\mathrm{cv}}(x) \leq f(x),\ \forall x \in X$$

$$\text{if } \alpha = -\min(0, (1/2)\lambda_{\mathrm{min}}([\nabla^2 f(X)]))$$

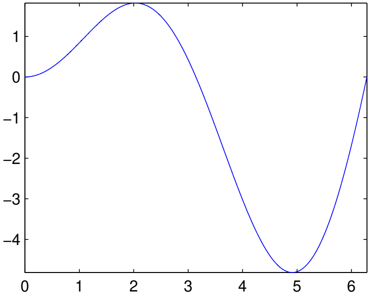
$$\text{then } \nabla^2 f^{\mathrm{cv}}(x) = \nabla^2 f(x) + 2I\alpha \geq 0,\ \forall x \in X$$

Illustrative example:

$$f(x) = x\sin(x),\ \forall x \in [0, 2\pi]$$

$$[\nabla_x^2 f(X)] = 2\cos(X) - X\sin(X)$$

(interval arithmetics)



Adjiman, et al. *Computers & Chemical Engineering* 22 (9), 1998.

Convex Relaxation using αBB Method

Find $f^{\mathrm{cv}}(x)$ by subtracting a 'convex-enough' function from $f(x)$:

$$f^{\mathrm{cv}}(x) = f(x) - \alpha(x - x^L)(x - x^U)$$

$$\text{if } \alpha \geq 0 \text{ then } f^{\mathrm{cv}}(x) \leq f(x),\ \forall x \in X$$

$$\text{if } \alpha = -\min(0, (1/2)\lambda_{\min}([\nabla^2 f(X)]))$$

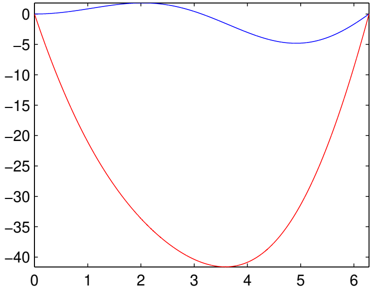
$$\text{then } \nabla^2 f^{\mathrm{cv}}(x) = \nabla^2 f(x) + 2I\alpha \geq 0,\ \forall x \in X$$

Illustrative example:

$$f(x) = x\sin(x),\ \forall x \in [0, 2\pi]$$

$$[\nabla_x^2 f(X)] = 2\cos(X) - X\sin(X)$$

(interval arithmetics)



Adjiman, et al. *Computers & Chemical Engineering* 22 (9), 1998.

Convex Relaxation of Special Functions

McCormick relaxations for bilinear functions

$$(x - x^L)(y - y^L) = xy - x^L y - y^L x + x^L y^L \geq 0$$

$$(x - x^U)(y - y^U) = xy - x^U y - y^U x + x^U y^U \geq 0$$

$$(x - x^U)(y - y^L) = xy - x^U y + y^L x + x^U y^L \leq 0$$

$$(x - x^L)(y - y^U) = xy - x^L y + y^U x + x^L y^U \leq 0$$

Relaxation:

$$w = xy$$

$$\begin{cases} w \geq x^L y + y^L x - x^L y^L \\ w \geq x^U y + y^U x - x^U y^U \\ w \leq x^U y + y^L x - x^U y^L \\ w \leq x^L y + y^U x - x^L y^U \end{cases}$$

Convex Relaxation of Special Functions

McCormick relaxations for bilinear functions

Relaxation:

$$(x - x^L)(y - y^L) = xy - x^L y - y^L x + x^L y^L \geq 0$$
$$(x - x^U)(y - y^U) = xy - x^U y - y^U x + x^U y^U \geq 0$$
$$(x - x^U)(y - y^L) = xy - x^U y + y^L x + x^U y^L \leq 0$$
$$(x - x^L)(y - y^U) = xy - x^L y + y^U x + x^L y^U \leq 0$$

$$w = xy$$

$$\begin{cases} w \geq x^L y + y^L x - x^L y^L \\ w \geq x^U y + y^U x - x^U y^U \\ w \leq x^U y + y^L x - x^U y^L \\ w \leq x^L y + y^U x - x^L y^U \end{cases}$$

Convex Relaxation of Special Functions

McCormick relaxations for bilinear functions

$$(x - x^L)(y - y^L) = xy - x^L y - y^L x + x^L y^L \geq 0$$
$$(x - x^U)(y - y^U) = xy - x^U y - y^U x + x^U y^U \geq 0$$
$$(x - x^U)(y - y^L) = xy - x^U y + y^L x + x^U y^L \leq 0$$
$$(x - x^L)(y - y^U) = xy - x^L y + y^U x + x^L y^U \leq 0$$

Relaxation:

$$w = xy$$

$$\begin{cases} w \geq x^L y + y^L x - x^L y^L \\ w \geq x^U y + y^U x - x^U y^U \\ w \leq x^U y + y^L x - x^U y^L \\ w \leq x^L y + y^U x - x^L y^U \end{cases}$$

Convex Relaxation of Special Functions

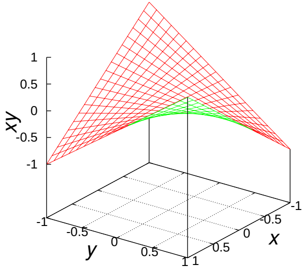
McCormick relaxations for bilinear functions

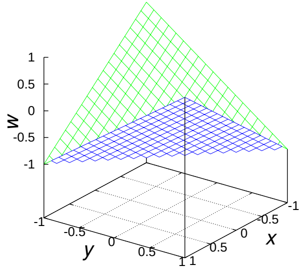
$$(x - x^L)(y - y^L) = xy - x^L y - y^L x + x^L y^L \geq 0$$
$$(x - x^U)(y - y^U) = xy - x^U y - y^U x + x^U y^U \geq 0$$
$$(x - x^U)(y - y^L) = xy - x^U y + y^L x + x^U y^L \leq 0$$
$$(x - x^L)(y - y^U) = xy - x^L y + y^U x + x^L y^U \leq 0$$

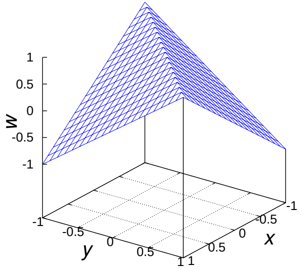
Relaxation:

$$w = xy$$

$$\begin{cases} w \geq x^L y + y^L x - x^L y^L \\ w \geq x^U y + y^U x - x^U y^U \\ w \leq x^U y + y^L x - x^U y^L \\ w \leq x^L y + y^U x - x^L y^U \end{cases}$$







McCormick Relaxations...

...work similarly for bilinear, trilinear, fractional terms. Even...

Theorem [1]

- $f, g : X \to \mathbb{R}$, with $X \subset \mathbb{R}^n$ convex
- $[f^L, f^U]$ and $[g^L, g^U]$ interval bounds for f and g on X
- $[f^{\mathrm{cv}}, f^{\mathrm{cc}}]$ and $[g^{\mathrm{cv}}, g^{\mathrm{cc}}]$ cvx/ccve relaxations for f and g on X

Convex/concave bounds for $h = f \times g$ on X:

$$\begin{aligned}
h^{\mathrm{cv}}(x) = \max\{ & \min\{g^L f^{\mathrm{cv}}(x), g^L f^{\mathrm{cc}}(x)\} + \min\{f^L g^{\mathrm{cv}}(x), f^L g^{\mathrm{cc}}(x)\} - f^L g^L, \\
& \min\{g^U f^{\mathrm{cv}}(x), g^U f^{\mathrm{cc}}(x)\} + \min\{f^U g^{\mathrm{cv}}(x), f^U g^{\mathrm{cc}}(x)\} - f^U g^U\} \\
h^{\mathrm{cc}}(x) = \min\{ & \max\{g^L f^{\mathrm{cv}}(x), g^L f^{\mathrm{cc}}(x)\} + \max\{f^U g^{\mathrm{cv}}(x), f^U g^{\mathrm{cc}}(x)\} - f^U g^L, \\
& \max\{g^U f^{\mathrm{cv}}(x), g^U f^{\mathrm{cc}}(x)\} + \max\{f^L g^{\mathrm{cv}}(x), f^L g^{\mathrm{cc}}(x)\} - f^L g^U\}
\end{aligned}$$

---

[1] McCormick, G.P., *Mathematical Programming*, **10**:147-175, 1976

Relaxation of Special Univariate Terms

Univariate Terms: $w = \varphi(x)$, $x^L \leq x \leq x^U$

- Convex Case: e.g., $\varphi : v \mapsto v^{2n}$, $\varphi : v \mapsto \exp(v)$

$$w = \varphi(x) \quad \xrightarrow{\text{relax}} \quad \begin{cases} w \geq \varphi(x) \\ w \leq \varphi(x^L) + \frac{\varphi(x^U)-\varphi(x^L)}{x^U-x^L}(x - x^L) \end{cases}$$

- Concave Case: e.g., $\varphi : v \mapsto \sqrt{v}$, $\varphi : v \mapsto \log(v)$

$$w = \varphi(x) \quad \xrightarrow{\text{relax}} \quad \begin{cases} w \geq \varphi(x^L) + \frac{\varphi(x^U)-\varphi(x^L)}{x^U-x^L}(x - x^L) \\ w \leq \varphi(x) \end{cases}$$

Relaxation of Special Univariate Terms

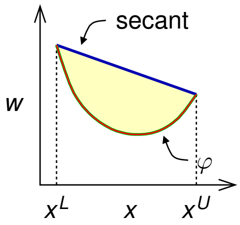
Univariate Terms: $w = \varphi(x)$, $x^L \leq x \leq x^U$

- Convex Case: e.g., $\varphi : v \mapsto v^{2n}$, $\varphi : v \mapsto \exp(v)$

$$w = \varphi(x) \quad \xrightarrow{\text{relax}} \quad \begin{cases} w \geq \varphi(x) \\ w \leq \varphi(x^L) + \frac{\varphi(x^U)-\varphi(x^L)}{x^U-x^L}(x - x^L) \end{cases}$$

- Concave Case: e.g., $\varphi : v \mapsto \sqrt{v}$, $\varphi : v \mapsto \log(v)$

$$w = \varphi(x) \quad \xrightarrow{\text{relax}} \quad \begin{cases} w \geq \varphi(x^L) + \frac{\varphi(x^U)-\varphi(x^L)}{x^U-x^L}(x - x^L) \\ w \leq \varphi(x) \end{cases}$$



Relaxation of Special Univariate Terms

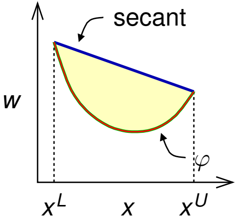
Univariate Terms: $w = \varphi(x)$, $x^L \leq x \leq x^U$

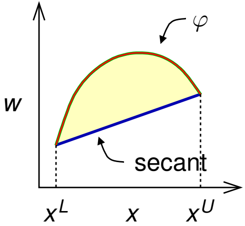
- Convex Case: e.g., $\varphi : v \mapsto v^{2n}$, $\varphi : v \mapsto \exp(v)$

$$w = \varphi(x) \quad \xrightarrow{\text{relax}} \quad \begin{cases} w \geq \varphi(x) \\ w \leq \varphi(x^L) + \frac{\varphi(x^U)-\varphi(x^L)}{x^U - x^L}(x - x^L) \end{cases}$$

- Concave Case: e.g., $\varphi : v \mapsto \sqrt{v}$, $\varphi : v \mapsto \log(v)$

$$w = \varphi(x) \quad \xrightarrow{\text{relax}} \quad \begin{cases} w \geq \varphi(x^L) + \frac{\varphi(x^U)-\varphi(x^L)}{x^U - x^L}(x - x^L) \\ w \leq \varphi(x) \end{cases}$$





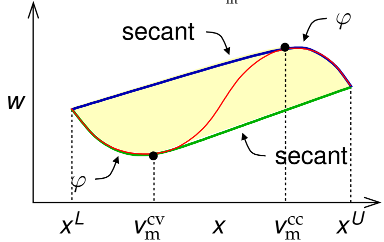
Relaxation of Special Univariate Terms

Univariate Terms: $w = \varphi(x)$, $x^L \leq x \leq x^U$

- Concavoconvex Case: e.g., $\varphi : v \mapsto v^{2n+1}$

$$w = \varphi(x) \quad \xrightarrow{\text{relax}} \quad \begin{cases} w \geq \begin{cases} \varphi(x), & \text{if } x \leq v_{\rm m}^{\rm cv} \\ \varphi(x^U) + \frac{\varphi(x^U)-\varphi(v_{\rm m}^{\rm cv})}{x^U - v_{\rm m}^{\rm cv}}(x - x^U), & \text{otherwise} \end{cases} \\ w \leq \begin{cases} \varphi(x), & \text{if } x \geq v_{\rm m}^{\rm cc} \\ \varphi(x^L) + \frac{\varphi(x^L)-\varphi(v_{\rm m}^{\rm cc})}{x^L - v_{\rm m}^{\rm cc}}(x - x^L), & \text{otherwise} \end{cases} \end{cases}$$

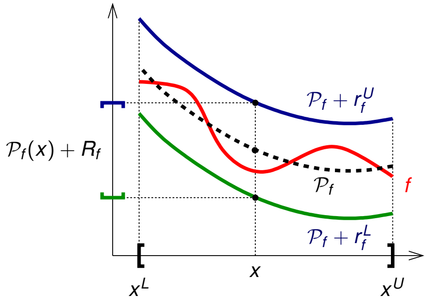
$$v_{\rm m}^{\rm cv}, v_{\rm m}^{\rm cc} \in [x^L, x^U] : \ \varphi'(v_{\rm m}^{\rm cv}) = \frac{\varphi(x^U)-\varphi(v_{\rm m}^{\rm cv})}{x^U - v_{\rm m}^{\rm cv}} \text{ and } \varphi'(v_{\rm m}^{\rm cc}) = \frac{\varphi(x^L)-\varphi(v_{\rm m}^{\rm cc})}{x^L - v_{\rm m}^{\rm cc}}$$



Taylor Model Estimators

Taylor Models:

- $f(x) \in \mathcal{T}_f(x) := \mathcal{P}_f(x) + R_f, \forall x \in X$
- $\mathcal{P}_f$ polynomial of order q
- R_f remainder interval term
- Use of recursive arithmetic rules
- Range: $f(X) \subseteq \mathcal{P}_f(X) + R_f$



TM of algebraic function

$$\mathcal{T}_f^q = f(\text{mid}(X)) + \left.\frac{\partial f}{\partial x}\right|_{\text{mid}(X)} (x - \text{mid}(X)) + \cdots + r_f^q \qquad \forall x \in X$$

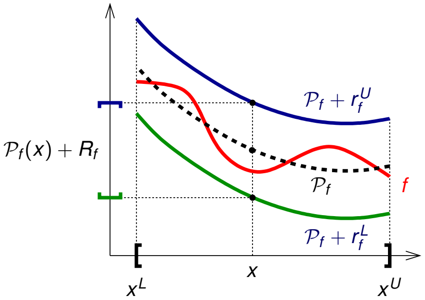
- remainder can be evaluated using interval arithmetics

$$r_f^q = \left.\frac{\partial^{(q+1)} f}{\partial x^{(q+1)}}\right|_X (X - \text{mid}(X))^{(q+1)}$$

Taylor Model Estimators

Taylor Models:

- $f(x) \in \mathcal{T}_f(x) := \mathcal{P}_f(x) + R_f, \forall x \in X$
- $\mathcal{P}_f$ polynomial of order q
- R_f remainder interval term
- Use of recursive arithmetic rules
- Range: $f(X) \subseteq \mathcal{P}_f(X) + R_f$



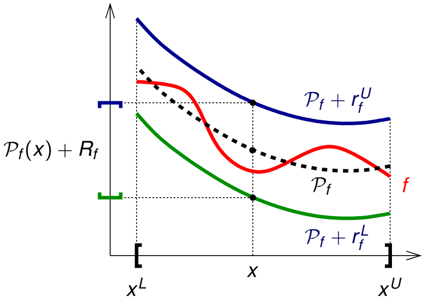
TM of algebraic function

$$\mathcal{T}_f^q = f(\text{mid}(X)) + \left.\frac{\partial f}{\partial x}\right|_{\text{mid}(X)} (x - \text{mid}(X)) + \cdots + r_f^q \qquad \forall x \in X$$

- remainder can be evaluated using interval arithmetics

$$r_f^q = \left.\frac{\partial^{(q+1)} f}{\partial x^{(q+1)}}\right|_X (X - \text{mid}(X))^{(q+1)}$$

Taylor Model Estimators



Taylor Models:

- $f(x) \in \mathcal{T}_f(x) := \mathcal{P}_f(x) + R_f, \forall x \in X$
- $\mathcal{P}_f$ polynomial of order q
- R_f remainder interval term
- Use of recursive arithmetic rules
- Range: $f(X) \subseteq \mathcal{P}_f(X) + R_f$

TM of algebraic function

$$\mathcal{T}_f^q = f(\text{mid}(X)) + \left.\frac{\partial f}{\partial x}\right|_{\text{mid}(X)} (x - \text{mid}(X)) + \cdots + r_f^q \qquad \forall x \in X$$

- remainder can be evaluated using interval arithmetics

$$r_f^q = \left.\frac{\partial^{(q+1)} f}{\partial x^{(q+1)}}\right|_X (X - \text{mid}(X))^{(q+1)}$$

Outline

Spatial Branch-and-Bound (B&B) Algorithm

Convex Relaxation of Factorable Programs

Bounding of Parametric ODEs

Guaranteed Parameter Estimation of Nonlinear Dynamic Systems

Bounding of Parametric ODEs

Problem Definition

Given a parameter box $P = [p^L, p^U]$ and a set of parametric ODEs $\dot{x} = f(x, p)$ and $x(t_0) = g(p)$, identify $X(t) = [x^L(t), x^U(t)]$ such that

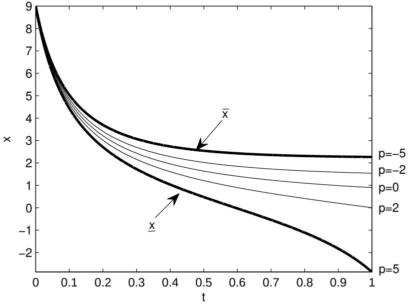
$$x(t, p) \in X(t), \ \forall t \in [t_0, t_N], \ \forall p \in P$$

ODE

$$\dot{x} = -x^2 - p$$
$$p \in [-5, 5]$$

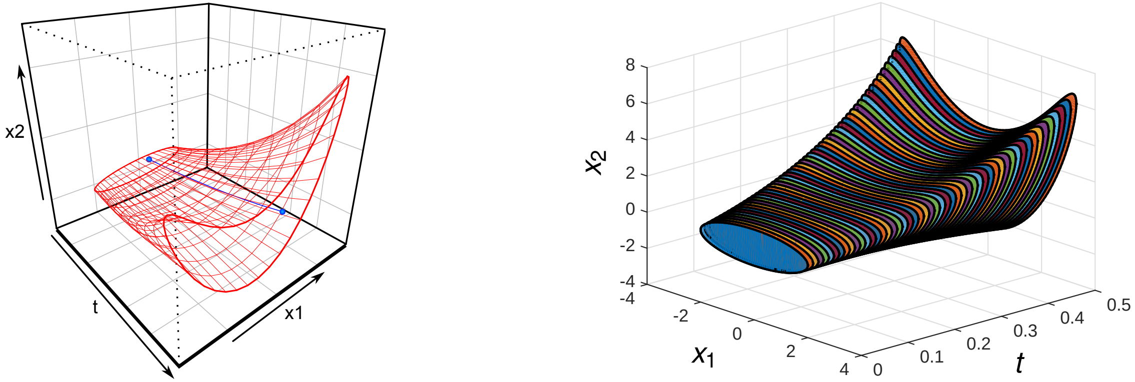
ODE Bounds

$$\dot{x}^U = -(x^U)^2 + 5$$
$$\dot{x}^L = -(x^L)^2 - 5$$



Bounding of Parametric ODEs

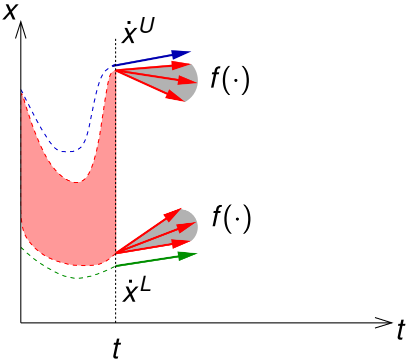
$$\dot{x}_1 = x_1$$
$$\dot{x}_2 = 2x_2 + x_1^2 + (1/8)x_2^2$$
$$x_1(0,p) = p_1, \quad x_2(0,p) = p_2, \quad p = (p_1, p_2)^{\mathrm{T}}$$

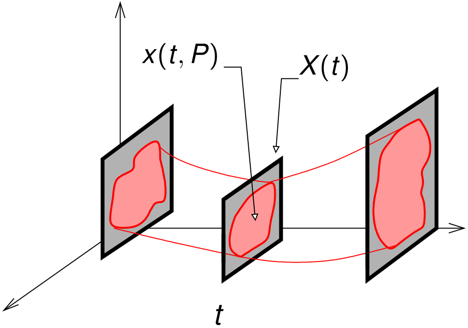


Techniques for Bounding of Parametric ODEs

1. interval methods (differential inequalities) (Walter, 1970)
2. Taylor models methods (Berz, 1997; Chachuat et al., 2012)

$$\dot{x}^U \geq \sup\{f(x,p) : p \in P, x \in [x^L, x^U]\}$$
$$\dot{x}^L \leq \inf\{f(x,p) : p \in P, x \in [x^L, x^U]\}$$



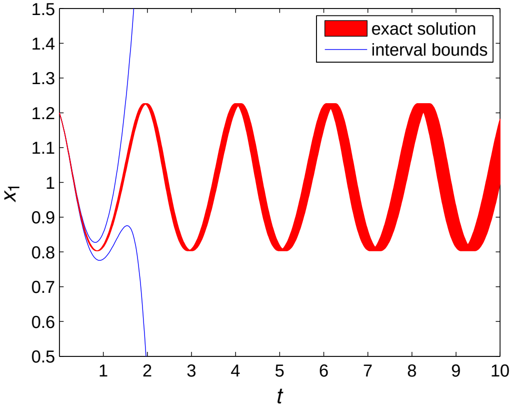


Techniques for Bounding of Parametric ODEs

1. interval methods (differential inequalities) (Walter, 1970)
2. Taylor models methods (Berz, 1997; Chachuat et al., 2012)

Lotka-Volterra system

$$\dot{x}_1 = p x_1 (1 - x_2), \; x_1(0) = 1.2,$$
$$\dot{x}_2 = p x_2 (x_1 - 1), \; x_2(0) = 1.1,$$
$$p \in [2.95, 3.05]$$



Order of convergence is linear: $\Delta X = a \Delta P$

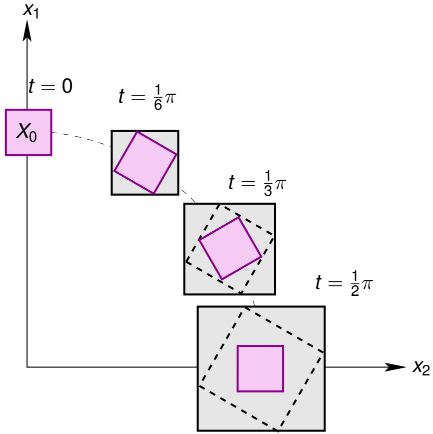
Techniques for Bounding of Parametric ODEs

1. interval methods (differential inequalities) (Walter, 1970)
2. Taylor models methods (Berz, 1997; Chachuat et al., 2012)

Wrapping effect

$$\dot{x}_1 = x_2$$
$$\dot{x}_2 = -x_1$$
$$x(0) \in X_0;$$

$$x(t) = \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix} X_0$$



Techniques for Bounding of Parametric ODEs

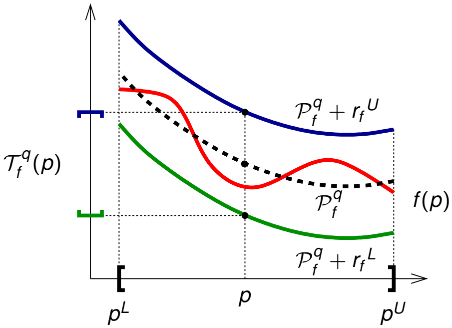
1. interval methods (differential inequalities) (Walter, 1970)
2. Taylor models methods (Berz, 1997; Chachuat et al., 2012)

Idea: Enclose the image of solution to ODE using

$$\mathcal{T}^q_{x(t)}(p) = \mathcal{P}^q_{x(t)} + r^q_{x(t)}$$

where $\mathcal{P}^q_{x(t)}(p)$ is qth order Taylor expansion of $x(t)$ and

$$\mathcal{P}^q_{x(t)} - x(t,p) \in r^q_{x(t)}, \ \forall p \in P$$



TM of ODE solution

$$\mathcal{T}^q_{x(t)} = x(t)(\mathrm{mid}(P)) + \left.\frac{\partial x(t)}{\partial p}\right|_{\mathrm{mid}(P)} (p - \mathrm{mid}(P)) + \cdots + r^q_f \qquad \forall p \in P$$

Techniques for Bounding of Parametric ODEs

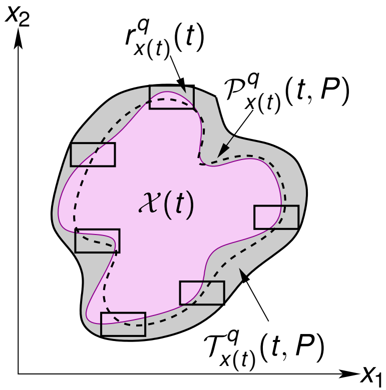
1. interval methods (differential inequalities) (Walter, 1970)
2. Taylor models methods (Berz, 1997; Chachuat et al., 2012)

Idea: Enclose the image of solution to ODE using

$$\mathcal{T}^q_{x(t)}(p) = \mathcal{P}^q_{x(t)} + r^q_{x(t)}$$

where $\mathcal{P}^q_{x(t)}(p)$ is qth order Taylor expansion of $x(t)$ and

$$\mathcal{P}^q_{x(t)} - x(t,p) \in r^q_{x(t)},\ \forall p \in P$$



Enclosure of solution to parametric ODE found by:

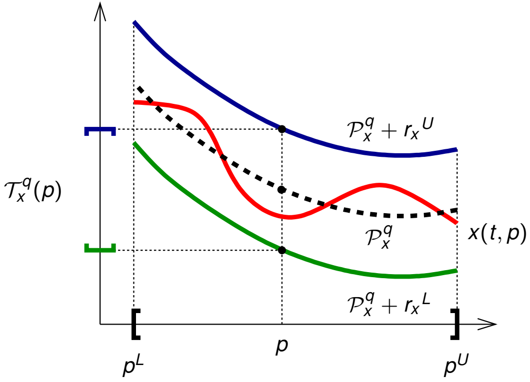
- coefficients of $\mathcal{P}^q_{x(t)}$ – parametric sensitivities $\left(\frac{\partial^q x(t)}{\partial p^q}\Big|_{\mathrm{mid}(P)}\right)$

Techniques for Bounding of Parametric ODEs

1. interval methods (differential inequalities) (Walter, 1970)
2. Taylor models methods (Berz, 1997; Chachuat et al., 2012)

$$\dot{r}_x^U \geq \sup\{f(\mathcal{T}_x^q, p) - \dot{\mathcal{P}}_x^q : p \in P, r_x \in [{r_x}^L, {r_x}^U]\}$$

$$\dot{r}_x^L \leq \inf\{f(\mathcal{T}_x^q, p) - \dot{\mathcal{P}}_x^q : p \in P, r_x \in [{r_x}^L, {r_x}^U]\}$$



$$\dot{x} = f(x, p)$$

$$\mathcal{T}_x^q = \mathcal{P}_x^q + r_x, \quad r_x \in [{r_x}^L, {r_x}^U]$$

$$\Downarrow$$

$$\dot{\mathcal{T}}_x^q = f(\mathcal{T}_x^q, p)$$

$$\Downarrow$$

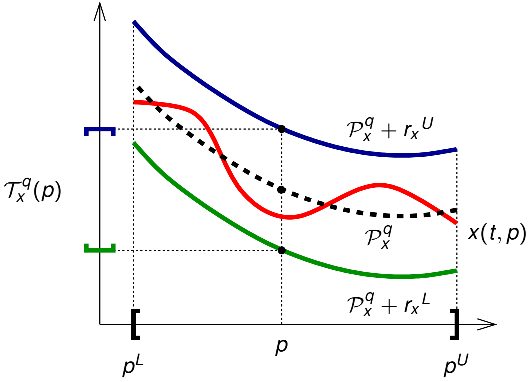
$$\dot{r}_x = f(\mathcal{T}_x^q, p) - \dot{\mathcal{P}}_x^q$$

Techniques for Bounding of Parametric ODEs

1. interval methods (differential inequalities) (Walter, 1970)
2. Taylor models methods (Berz, 1997; Chachuat et al., 2012)

$$\dot{r}_x^U \geq \sup\{f(\mathcal{T}_x^q, p) - \dot{\mathcal{P}}_x^q : p \in P, r_x \in [{r_x}^L, {r_x}^U]\}$$

$$\dot{r}_x^L \leq \inf\{f(\mathcal{T}_x^q, p) - \dot{\mathcal{P}}_x^q : p \in P, r_x \in [{r_x}^L, {r_x}^U]\}$$



$$\dot{x} = f(x, p)$$

$$\mathcal{T}_x^q = \mathcal{P}_x^q + r_x, \quad r_x \in [{r_x}^L, {r_x}^U]$$

$$\Downarrow$$

$$\dot{\mathcal{T}}_x^q = f(\mathcal{T}_x^q, p)$$

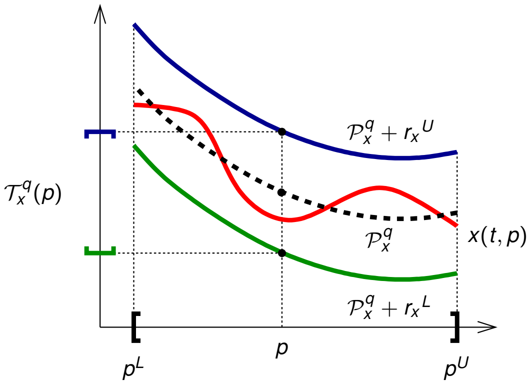
$$\Downarrow$$

$$\dot{r}_x = f(\mathcal{T}_x^q, p) - \dot{\mathcal{P}}_x^q$$

Techniques for Bounding of Parametric ODEs

1. interval methods (differential inequalities) (Walter, 1970)
2. Taylor models methods (Berz, 1997; Chachuat et al., 2012)

$$\dot{r}_x^U \geq \sup\{f(\mathcal{T}_x^q, p) - \dot{\mathcal{P}}_x^q : p \in P, r_x \in [{r_x}^L, {r_x}^U]\}$$
$$\dot{r}_x^L \leq \inf\{f(\mathcal{T}_x^q, p) - \dot{\mathcal{P}}_x^q : p \in P, r_x \in [{r_x}^L, {r_x}^U]\}$$



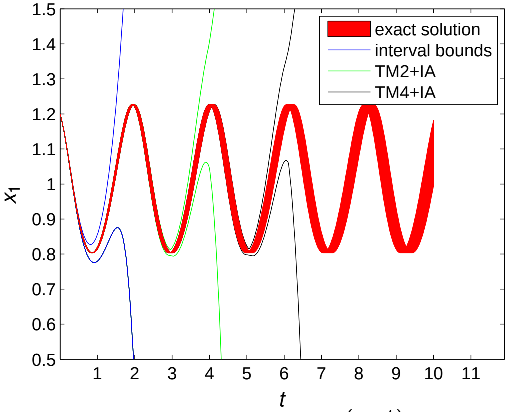
$$\dot{x} = f(x,p)$$
$$\mathcal{T}_x^q = \mathcal{P}_x^q + r_x, \quad r_x \in [{r_x}^L, {r_x}^U]$$
$$\Downarrow$$
$$\dot{\mathcal{T}}_x^q = f(\mathcal{T}_x^q, p)$$
$$\Downarrow$$
$$\dot{r}_x = f(\mathcal{T}_x^q, p) - \dot{\mathcal{P}}_x^q$$

Techniques for Bounding of Parametric ODEs

1. interval methods (differential inequalities) (Walter, 1970)
2. Taylor models methods (Berz, 1997; Chachuat et al., 2012)

Lotka-Volterra system

$$\dot{x}_1 = px_1(1 - x_2), \quad x_1(0) = 1.2,$$
$$\dot{x}_2 = px_2(x_1 - 1), \quad x_2(0) = 1.1,$$
$$p \in [2.95, 3.05]$$



Order of convergence is much higher than linear: $\Delta X = a(\Delta P)^{(q+1)}$

Outline

Spatial Branch-and-Bound (B&B) Algorithm

Convex Relaxation of Factorable Programs

Bounding of Parametric ODEs

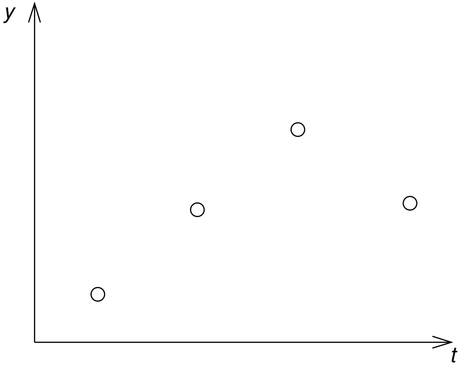
Guaranteed Parameter Estimation of Nonlinear Dynamic Systems

Classical Parameter Estimation

$$\text{model: } \hat{y} = F(p) \begin{cases} \text{dynamic system:} & \dot{x} = f(x, p),\ x(0) = g(p) \\ \text{output equation:} & \hat{y} = h(x, p) \end{cases}$$

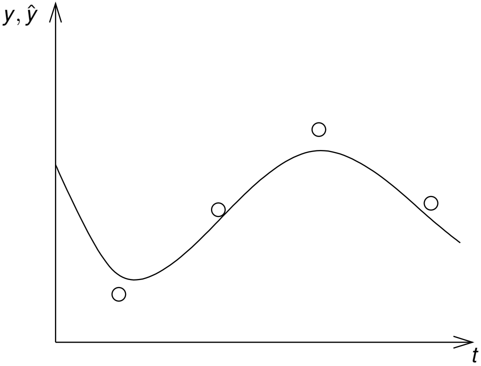
Classical Parameter Estimation

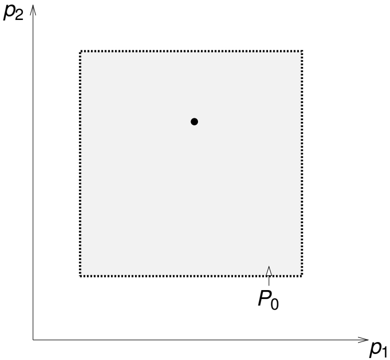
$$\text{model: } \hat{y} = F(p) \begin{cases} \text{dynamic system:} & \dot{x} = f(x,p),\ x(0) = g(p) \\ \text{output equation:} & \hat{y} = h(x,p) \end{cases}$$



Classical Parameter Estimation

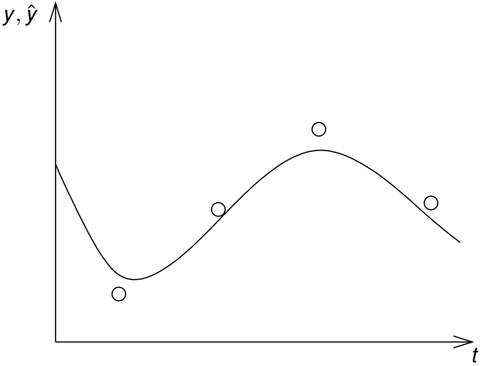
$$\text{model: } \hat{y} = F(p) \begin{cases} \text{dynamic system:} & \dot{x} = f(x,p),\; x(0) = g(p) \\ \text{output equation:} & \hat{y} = h(x,p) \end{cases}$$

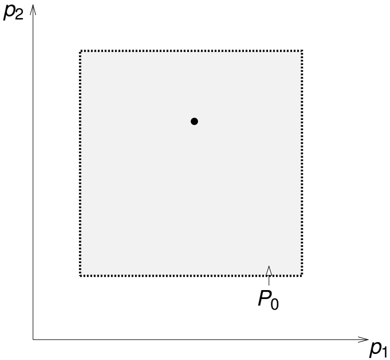




Classical Parameter Estimation

$$\min_{p \in P_0} \|\hat{y} - y\|_2^2$$
$$\text{s.t. } \hat{y} = F(p)$$

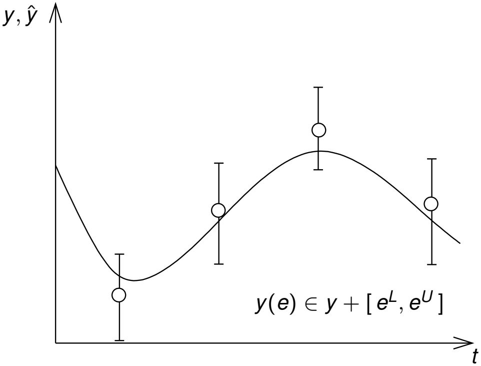


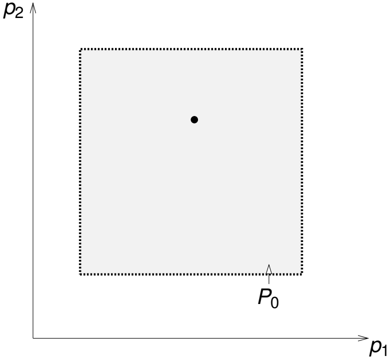


Guaranteed Parameter Estimation

$$\min_{p \in P_0} \|\hat{y} - y\|_2^2$$

$$\text{s.t. } \hat{y} = F(p)$$

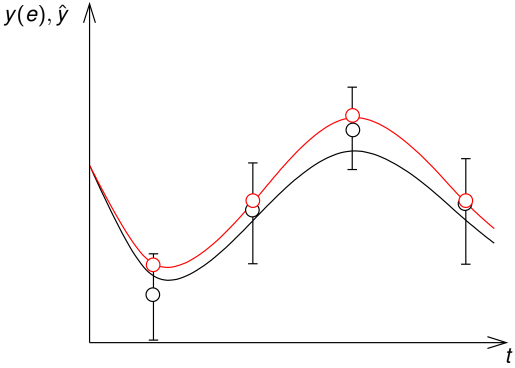


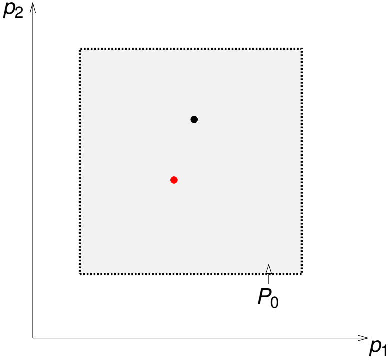


Guaranteed Parameter Estimation

$$e \in [e^L, e^U] : \min_{p \in P_0} \|\hat{y} - y + e\|_2^2$$

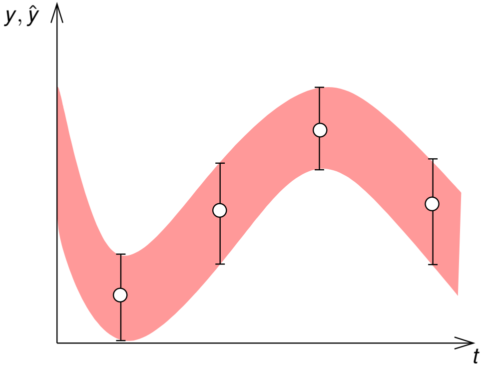
$$\text{s.t. } \hat{y} = F(p)$$

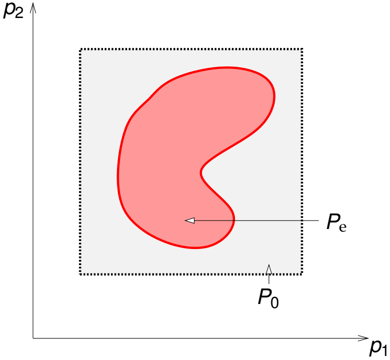




Guaranteed Parameter Estimation

$$\forall e \in [e^L, e^U] : \min_{p \in P_0} \|\hat{y} - y + e\|_2^2$$
$$\text{s.t. } \hat{y} = F(p)$$



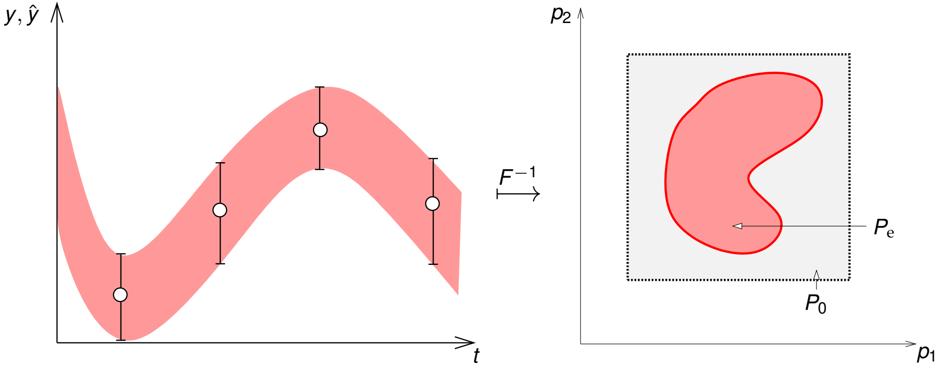


Finding Solution to Guaranteed PE Problem

- Find a set

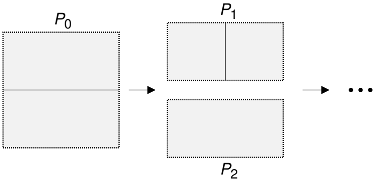
$$P_e := \left\{ p \in P_0 \,\middle|\, \begin{array}{l} \hat{y} = F(p), \\ e^L \le \hat{y} - y \le e^U \end{array} \right\} \equiv \{p \in P_0 | g(p) \le 0\}$$

- Apply set inversion via interval analysis (Jaulin, 1993)

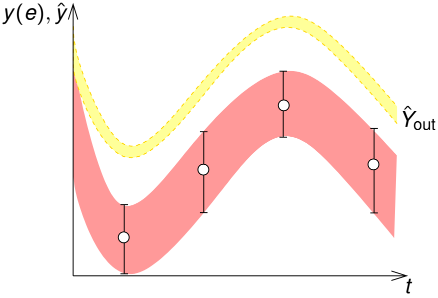


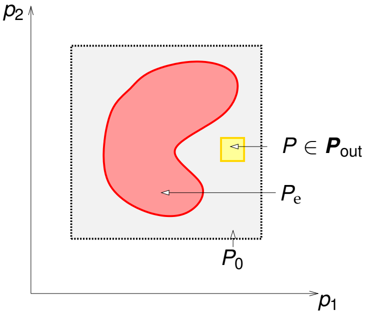
Finding Solution via Set Inversion

1. Apply branching on parameter space



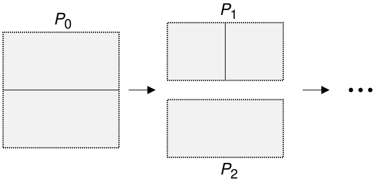
2. Employ exclusion tests on parameter space partitions



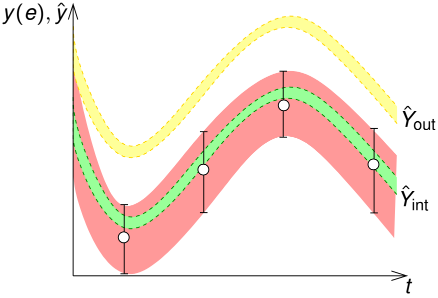


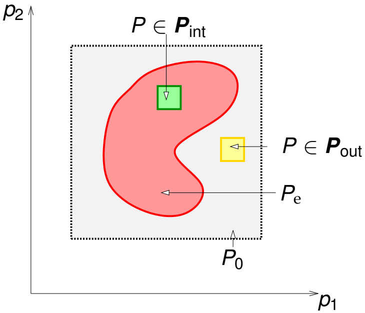
Finding Solution via Set Inversion

1. Apply branching on parameter space



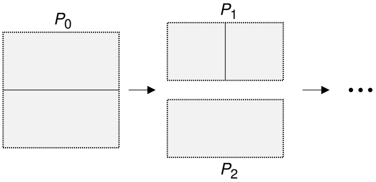
2. Employ exclusion tests on parameter space partitions



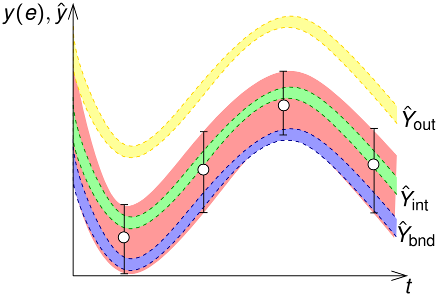


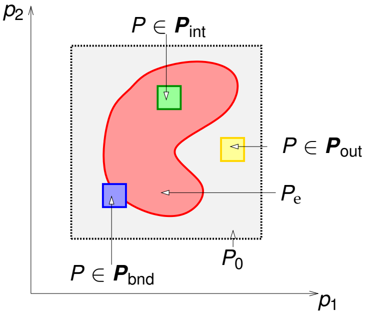
Finding Solution via Set Inversion

1. Apply branching on parameter space



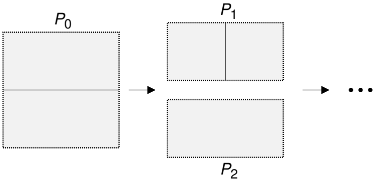
2. Employ exclusion tests on parameter space partitions



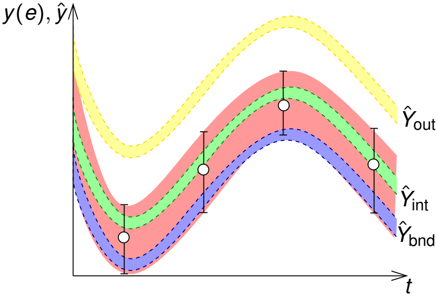


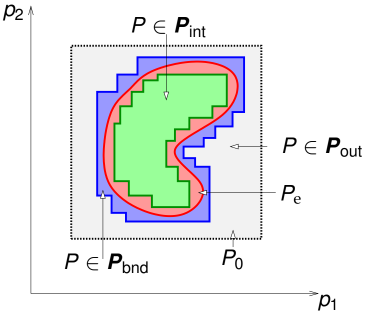
Finding Solution via Set Inversion

1. Apply branching on parameter space



2. Employ exclusion tests on parameter space partitions

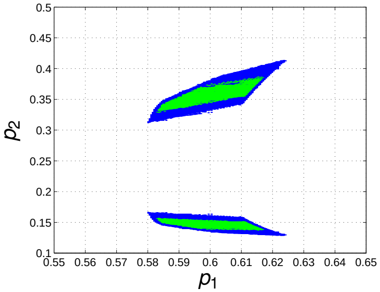


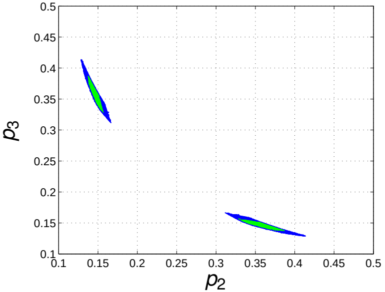


Case Study

$$\dot{x}_1 = -(p_1 + p_3)x_1 + p_2 x_2, \qquad x_1(0) = 1,$$
$$\dot{x}_2 = p_1 x_1 - p_2 x_2, \qquad x_2(0) = 0.$$

$$(p_1, p_2, p_3) = (0.6, 0.15, 0.35) \rightarrow y = x_2 + [-0.05, 0.05]$$



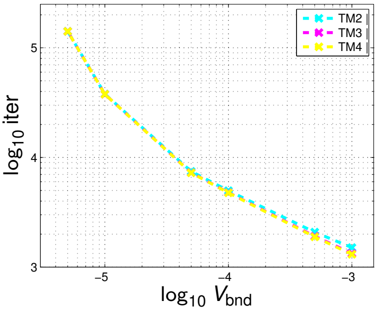


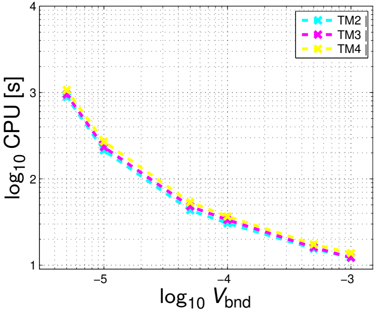
$V_{\text{bnd}} = 5 \times 10^{-6}$ (2nd-order TM, $>$70,000 boxes, 900s CPU time)

Case Study

$$\dot{x}_1 = -(p_1 + p_3)x_1 + p_2 x_2, \qquad x_1(0) = 1,$$
$$\dot{x}_2 = p_1 x_1 - p_2 x_2, \qquad x_2(0) = 0.$$

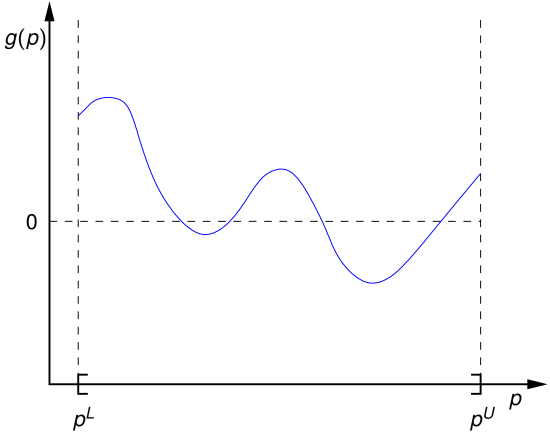
$$(p_1, p_2, p_3) = (0.6, 0.15, 0.35) \rightarrow y = x_2 + [-0.05, 0.05]$$





Implementation in C++ libraries MC++ and CRONOS (Chachuat et al.)

Optimization-based Domain Reduction



Constraint:

$$g(x) \leq 0$$

Convex relaxation:

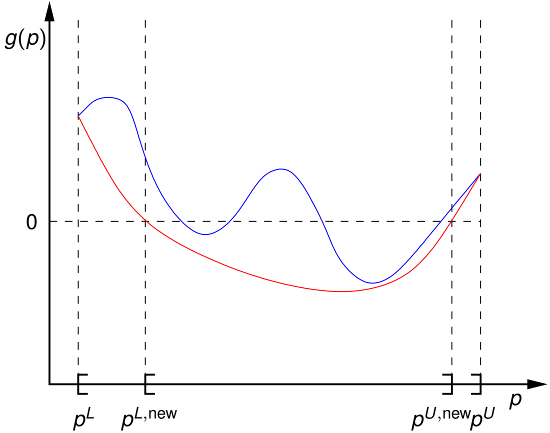
$$g^{\mathrm{cv}}(p) \leq 0$$

$$p^{L,\mathrm{new}} / p^{U,\mathrm{new}} =$$

$$\min_{p \in P} / \max_{p \in P} p$$

$$\text{s.t. } g^{\mathrm{cv}}(p) \leq 0$$

Optimization-based Domain Reduction



Constraint:

$$g(x) \leq 0$$

Convex relaxation:

$$g^{\text{cv}}(p) \leq 0$$

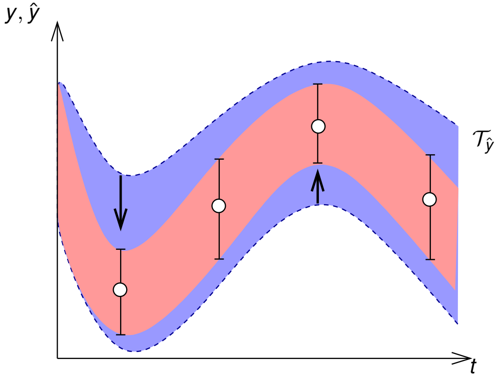
$$p^{L,\text{new}} / p^{U,\text{new}} = \min_{p \in P} / \max_{p \in P} p \quad \text{s.t. } g^{\text{cv}}(p) \leq 0$$

Domain Reduction for Guaranteed PE

$$p^{L,\text{new}} / p^{U,\text{new}} = \min_{p \in P} / \max_{p \in P} p$$

$$\text{s.t.} \ \left(\mathcal{P}^q_{\hat{y}}(p)\right)^{\text{cv}} + r_{\hat{y}}{}^L \leq y + e^L$$

$$\left(-\mathcal{P}^q_{\hat{y}}(p)\right)^{\text{cv}} - r_{\hat{y}}{}^U \leq -y - e^U$$



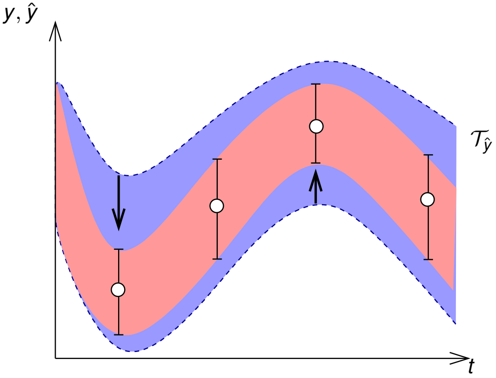
Convex relaxation of $\mathcal{P}^q_{\hat{y}}(p)$ by polyhedral relaxation $\Rightarrow$ LP

Domain Reduction for Guaranteed PE

$$p^{L,\text{new}} / p^{U,\text{new}} = \min_{p \in P} / \max_{p \in P} p$$

$$\text{s.t.} \quad \left(\mathcal{P}^q_{\hat{y}}(p)\right)^{\text{cv}} + r_{\hat{y}}{}^L \leq y + e^L$$

$$\left(-\mathcal{P}^q_{\hat{y}}(p)\right)^{\text{cv}} - r_{\hat{y}}{}^U \leq -y - e^U$$



Convex relaxation of $\mathcal{P}^q_{\hat{y}}(p)$ by polyhedral relaxation $\Rightarrow$ LP

Polyhedral Relaxation of Polynomial Functions

1. Decomposition

$$\mathcal{P}(p) = a_1 p_1 + a_2 p_2 + a_3 p_2^2 + a_4 p_1 p_2^2$$

$$\mathcal{P}(v) = a_1 v_1 + a_2 v_2 + a_3 v_3 + a_4 v_4$$
$$v_1 = p_1,\ v_2 = p_2,\ v_3 = v_2^2,\ v_4 = v_1 v_3$$

2. Affine relaxation of bilinear terms (McCormick, 1976)
3. Outer-approximation

Polyhedral Relaxation of Polynomial Functions

1. Decomposition $$\mathcal{P}(p) = a_1 p_1 + a_2 p_2 + a_3 p_2^2 + a_4 p_1 p_2^2$$

$$\mathcal{P}(v) = a_1 v_1 + a_2 v_2 + a_3 v_3 + a_4 v_4$$
$$v_1 = p_1,\ v_2 = p_2,\ v_3 = v_2^2,\ v_4 = v_1 v_3$$

2. Affine relaxation of bilinear terms (McCormick, 1976)
3. Outer-approximation

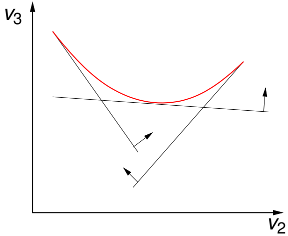
Polyhedral Relaxation of Polynomial Functions

1. Decomposition $$\mathcal{P}(p) = a_1 p_1 + a_2 p_2 + a_3 p_2^2 + a_4 p_1 p_2^2$$

$$\mathcal{P}(v) = a_1 v_1 + a_2 v_2 + a_3 v_3 + a_4 v_4$$

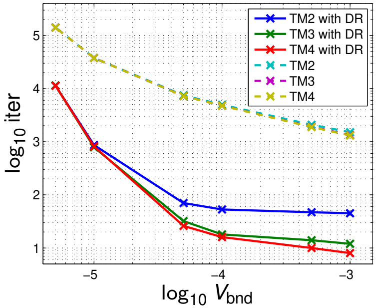
$$v_1 = p_1,\ v_2 = p_2,\ v_3 = v_2^2,\ v_4 = v_1 v_3$$

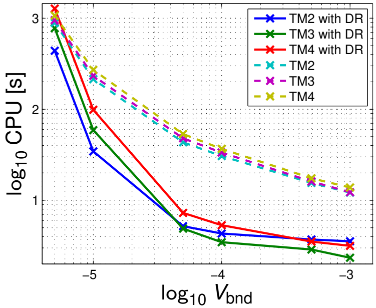
2. Affine relaxation of bilinear terms (McCormick, 1976)
3. Outer-approximation



Case Study (Continued)

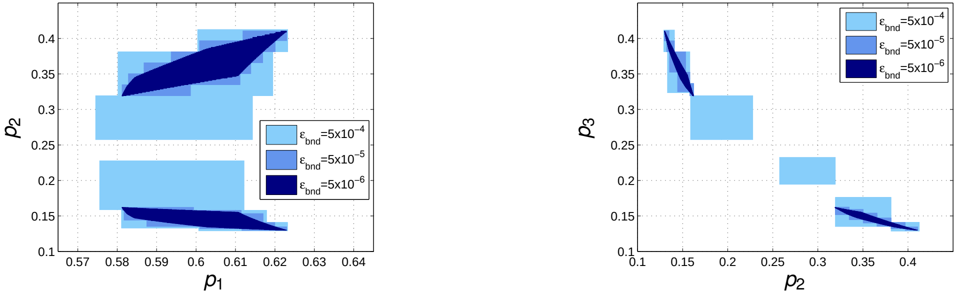
$$\dot{x}_1 = -(p_1 + p_3)x_1 + p_2 x_2, \qquad x_1(0) = 1,$$
$$\dot{x}_2 = p_1 x_1 - p_2 x_2, \qquad x_2(0) = 0.$$
$$(p_1, p_2, p_3) = (0.6, 0.15, 0.35) \rightarrow y = x_2 + [-0.05, 0.05]$$





Case Study (Continued)

$$\dot{x}_1 = -(p_1 + p_3)x_1 + p_2 x_2, \qquad x_1(0) = 1,$$
$$\dot{x}_2 = p_1 x_1 - p_2 x_2, \qquad x_2(0) = 0.$$
$$(p_1, p_2, p_3) = (0.6, 0.15, 0.35) \rightarrow y = x_2 + [-0.05, 0.05]$$



$$V_{bnd} = 5 \times 10^{-6} \text{ (2nd-order TM, 11,250 boxes, 350s CPU time)}$$
$$V_{bnd} = 5 \times 10^{-5} \text{ (2nd-order TM, 34 boxes, 5s CPU time)}$$

Conclusions

- effective convexification techniques are vital for solution of static/dynamic optimization and set-based problems
- efficient bounding techniques for parametric ODEs are essential ingredients
- optimization-based domain-reduction techniques can further enhance the convergence and properties of the obtained solutions

Slides

Funded by the European Union

Subscribe

Software