Approaches to Improve Bilinear Relaxations in Reduced-Space

Matthew Wilhelm, PhD Candidate
Matthew Stuber, Assistant Professor

November 9th, 2021
Global Optimization

- Nonconvex MINLP formulations naturally arise in many applications.
- MINLP solvers generally rely on some variation of spatial branch-and-bound\(^2,3\).
- Relaxed subproblems are used to compute bounds and are often derived from relaxed functions\(^2,3\).


AIChE Annual Meeting 2021
Reduced-Space Relaxations

Factorable Function

\[ y = \sin(w) \]

\[ w = z_1 + z_2 \]

\[ y = \sin(z_1 + z_2) \]

Compute \( y^{cv}(z), y^{cc}(z), y^L, y^U \)

Compute \( w^{cv}(z), w^{cc}(z), w^L, w^U \)

Initialize \( z^{cv}, z^{cc}, z^L, z^U \)


<table>
<thead>
<tr>
<th>Notation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>( q^{cv}(z) )</td>
<td>Convex relaxation</td>
</tr>
<tr>
<td>( q^{cc}(z) )</td>
<td>Concave relaxation</td>
</tr>
<tr>
<td>( q^L )</td>
<td>Lower bound</td>
</tr>
<tr>
<td>( q^U )</td>
<td>Upper bound</td>
</tr>
</tbody>
</table>
Reduced-Space Relaxations

Factorable Function

\[ \sin\]

\[ +\]

\[ z_1\]

\[ z_2\]

**PROS**

- Tight relaxations
- Favorable convergence properties
- Relaxed problem same dimension as decision space
- Ease of implementation

**CONS**

- Multidimensional.
- Nonsmooth
- Overly conservative.

Repeatedly enclosing nonconvex functions with convex/concave functions.

---

Reduced-Space Relaxations

Factorable Function

- Auxiliary ODEs or discretize-and-relax
- Repeated relaxation of right-hand-side function

Parametric ODEs

Reduced-Space Relaxations

Factorable Function

Parametric ODEs

Implicit Functions

- A fixed-point method, involves operation similar linear solve.
- Repeated addition, subtraction, multiplication.

Reduced-Space Relaxations

**Select Contributions to McCormick Relaxation Theory**

a) Special treatment of specific functional forms
b) Improvements to Multivariate Relaxations

---

A Priori Estimators

Underestimators $u_x$ and $u_y$ of $x$ and $y$ are known such that

\begin{align*}
x^L &\leq u_x \leq \min(x, a_x), \quad x \leq x^U \\
y^L &\leq u_y \leq \min(y, a_y), \quad y \leq y^U
\end{align*}

\begin{align*}
w &\geq (y^U - y^L)u_x + y^Lx + axy - axy^U \\
w &\geq (x^U - x^L)u_y + a_yx + x^Ly - a_yx^U \\
w &\geq (a_y - y^L)u_x + (a_x - x^L)u_y + y^Lx + x^Ly - axa_y \\
w &\geq (y^U - a_y)u_x + (x^U - a_x)u_y + a_yx + axy + axa_y - axy^U - x^Ua_y
\end{align*}


4x overestimating inequalities

8x (under/over)-estimators from over-estimators $v_x, v_y$ with upper bounds $b_x, b_y$
Relaxation of Bilinear Term

McCormick Relaxation of Bilinear Term\(^{13}\)

\[ w = x \times y \]

**Underestimators**

1. \[ w \geq x^L y + x y^L - x^L y^L \]
2. \[ w \geq x^U y + x y^U - x^U y^U \]

**Overestimators**

1. \[ w \leq x^U y + x y^L - x^U y^L \]
2. \[ w \leq x y^U + x^L y - x^L y^U \]

---

Relaxation of Composite Bilinear Term

McCormick Relaxation of Composite Bilinear Term

\[ w(z) = x(z) \times y(z) \]

Underestimators Convex Relaxations

\[ w_1^{cv}(z) = \min(x^L y^{cv}(z), x^L y^{cc}(z)) + \min(y^L x^{cv}(z), y^L x^{cc}(z)) - x^L y^L \]

\[ w_2^{cv}(z) = \min(x^U y^{cv}(z), x^U y^{cc}(z)) + \min(y^U x^{cv}(z), y^U x^{cc}(z)) - x^U y^U \]

Overestimators Concave Relaxations

\[ w_1^{cc}(z) = \max(x^U y^{cv}(z), x^U y^{cc}(z)) + \max(y^L x^{cv}(z), y^L x^{cc}(z)) - x^U y^L \]

\[ w_2^{cc}(z) = \max(y^U x^{cv}(z), y^U x^{cc}(z)) + \max(x^L y^{cv}(z), x^L y^{cc}(z)) - x^L y^U \]


Relaxations of **Composite** Bilinear Term A Priori

\[
\begin{align*}
  w_1^{cv}(z) &= (y^U - a_y)u_x^{cv}(z) + (x^U - a_x)u_y^{cv}(z) + \min(a_yx^{cv}(z), a_yx^{cc}(z)) + \min(a_xy^{cv}(z), a_xy^{cc}(z)) + a_xa_y - a_xy^U - x^Ua_y \\
  w_2^{cv}(z) &= (a_y - y^L)u_x^{cv}(z) + (a_x - x^L)u_y^{cv}(z) + \min(y^Lx^{cv}(z), y^Lx^{cc}(z)) + \min(y^Lx^{cv}(z), y^Lx^{cc}(z)) - a_xa_y \\
  w_3^{cv}(z) &= (y^U - y^L)u_x^{cv}(z) + \min(y^Lx^{cv}(z), y^Lx^{cc}(z)) + \min(a_xy^{cv}(z), a_xy^{cc}(z)) - a_xy^U \\
  w_4^{cv}(z) &= (x^U - x^L)u_y^{cv}(z) + \min(a_yx^{cv}(z), a_yx^{cc}(z)) + \min(y^Lx^{cv}(z), y^Lx^{cc}(z)) - a_yx^U
\end{align*}
\]

+ 12 additional inequalities

**Requires** convex/concave relaxations of some a priori under/over-estimators and respective extrema:

\[
  u_x^{cv}(z), u_y^{cv}(z), a_x, a_y, v_x^{cc}(z), v_y^{cc}(z), b_x, b_y
\]
A Priori Estimators, Reduced Space

Where to get convex/concave relaxations of a priori estimators (and extremal values)?
Approach 1: McCormick Relaxation

Where to get convex/concave relaxations of a priori estimators (and extremal values)?

Use convex/concave McCormick relaxations\(^4\) for under/overestimators

\[
\begin{align*}
x(z) & \geq x^{cv}(z) = u_x(z) = u_{x}^{cv}(z) \\
x(z) & \leq x^{cc}(z) = u_x(z) = u_{x}^{cc}(z) \\
y(z) & \geq y^{cv}(z) = u_y(z) = u_{y}^{cv}(z) \\
y(z) & \leq y^{cc}(z) = u_y(z) = u_{y}^{cc}(z)
\end{align*}
\]


\[f(x) = 2x - 3 \exp(x) \sin(2x)\]
Approach 1: McCormick Relaxation

Where to get convex/concave relaxations of a priori estimators (and extremal values)?

Concave optimization problems over box $Z$

$$a_x \geq \max_{z \in Z} x^{cv}(z)$$

$$b_x \leq \min_{z \in Z} x^{cc}(z)$$

$$a_y \geq \max_{z \in Z} y^{cv}(z)$$

$$b_y \leq \min_{z \in Z} y^{cc}(z)$$

$f(x) = 2x - 3 \exp(x) \sin(2x)$
Approach 2: Subgradients

Where to get convex/concave relaxations of a priori estimators (and extremal values)?

Use **affine relaxations** implied by **subgradients** computed using **McCormick framework**

\[
x(z) \geq x_{cv}(\bar{z}) + \nabla s_{x}^{cv}(\bar{z})(z - \bar{z}) \\
x(z) \leq x_{cc}(\bar{z}) + \nabla s_{x}^{cc}(\bar{z})(z - \bar{z}) \\
y(z) \geq y_{cv}(\bar{z}) + \nabla s_{y}^{cv}(\bar{z})(z - \bar{z}) \\
y(z) \leq y_{cc}(\bar{z}) + \nabla s_{y}^{cc}(\bar{z})(z - \bar{z})
\]


\[
f(x) = 2x - 3 \exp(x) \sin(2x)
\]
Approach 2: Subgradients

Where to get convex/concave relaxations of a priori estimators (and extremal values)?

Use affine relaxations implied by subgradients computed using McCormick framework\(^4\):

\[
\begin{align*}
x(z) &\geq x^{CV}(\bar{z}) + \nabla s^{CV}_x(\bar{z})(z - \bar{z}) = u_x(z) = u^{CV}_x(z) \\
x(z) &\leq x^{CC}(\bar{z}) + \nabla s^{CC}_x(\bar{z})(z - \bar{z}) = u_x(z) = u^{CC}_x(z) \\
y(z) &\geq y^{CV}(\bar{z}) + \nabla s^{CV}_y(\bar{z})(z - \bar{z}) = u_y(z) = u^{CV}_y(z) \\
y(z) &\leq y^{CC}(\bar{z}) + \nabla s^{CC}_y(\bar{z})(z - \bar{z}) = u_y(z) = u^{CC}_y(z)
\end{align*}
\]

**Approach 2: Subgradients**

Where to get convex/concave relaxations of a priori estimators (and extremal values)?

Use **affine relaxations** implied by **subgradients** computed using McCormick framework

\[
\begin{align*}
    a_x &= (x^{cv}(\bar{z}) + \nabla S_x^{cv}(\bar{z})(Z - \bar{z}))^U \geq u_x^{cv}(Z) \\
    b_x &= (x^{cc}(\bar{z}) + \nabla S_x^{cc}(\bar{z})(Z - \bar{z}))^L \leq u_x^{cc}(Z) \\
    a_y &= (y^{cv}(\bar{z}) + \nabla S_y^{cv}(\bar{z})(Z - \bar{z}))^U \geq u_y^{cv}(Z) \\
    b_y &= (y^{cc}(\bar{z}) + \nabla S_y^{cc}(\bar{z})(Z - \bar{z}))^L \leq u_y^{cc}(Z)
\end{align*}
\]

1-D Example

**Function**

\[
f(z) = (z - z^2)(z^3 - \exp(z)), \quad z \in [-0.5, 1]
\]

- **Subgradient-Prior:** Use plane defined by subgradient at reference point for under/over-estimator.
- **Max-Concave:** McCormick relaxation are used as under/over-estimators.
- **Interval:** Use interval extensions of subgradients to tighten natural interval bounds for each factor.
  - McCormick Relaxation\(^8\)
  - Multivariate McCormick Relaxation\(^{11}\)
  - A priori Calculation (Subgradient)
  - A priori Calculation (Max-Concave)

1-D Example

**Function**

\[ f(z) = (z - z^2)(z^3 - \exp(z)), \quad z \in [-0.5, 1] \]

- **Subgradient-Prior**: Use plane defined by subgradient at reference point for under/over-estimator.
- **Max-Concave**: McCormick relaxation are used as under/over-estimators.
- **Interval**: Use interval extensions of subgradients to tighten natural interval bounds for each factor.
  - McCormick Relaxation\(^8\)
  - Multivariate McCormick Relaxation\(^{11}\)
  - A priori Calculation (Subgradient)
  - A priori Calculation (Max-Concave)

1-D Example

Function

\[ f(z) = z - z^2 z^3 - \exp(z), \]

Example Function

\[ z \in [-0.5, 1] \]

Subgradient Prior:
Use plane defined by subgradient at reference point for under/over estimator.

Max-Concave:
McCormick relaxation are used as under/over estimators.

Interval:
Use interval extensions of subgradients to tighten natural interval bounds for each factor.

Rationale

\[ w \geq (y^U - y^L)u_x + y^L x + a_x(y - y^U) \]

Improving \( u_x \) may increase or decrease \( a_x \).

Observation

Subgradient-based a priori derived from McCormick relaxations are tighter than directly use of McCormick relaxation.

Approach 3: Affine Arithmetic

AF1 – Affine Form:\textsuperscript{16}:

\[ v = v_0 + \sum_{i=1}^{n} v_i \epsilon_i + v_{n+1} \epsilon_\pm \]

where \( \epsilon_i, \epsilon_\pm \in [-1,1] \)

\[ v(z) \geq v_0 + \sum_{i=1}^{n} v_i \epsilon_i(z) - v_{n+1} \]
\[ v(z) \leq v_0 + \sum_{i=1}^{n} v_i \epsilon_i(z) + v_{n+1} \]

\[ \epsilon_i(z) = \frac{z_i - \text{mid}(Z_i)}{\text{rad}(Z_i)} \]

AF2 – Affine Form:\textsuperscript{16}:

\[ v = v_0 + \sum_{i=1}^{n} v_i \epsilon_i + v_{n+1} \epsilon_\pm + v_{n+2} \epsilon_+ + v_{n+3} \epsilon_- \]

where \( \epsilon_i, \epsilon_\pm \in [-1,1], \ \epsilon_+ \in [0,1], \ \epsilon_- \in [-1,0] \)

\[ v(z) \geq v_0 + \sum_{i=1}^{n} v_i \epsilon_i(z) - v_{n+1} - v_{n+3} \]
\[ v(z) \leq v_0 + \sum_{i=1}^{n} v_i \epsilon_i(z) + v_{n+1} + v_{n+2} \]

Approach 3: Affine Arithmetic

\[ w = (x^2 - x) \times (y^2 - y), \]

\[ Z_1 \]

\[ Z_2 \]

\[ X \times Y = [0.1, 0.9]^2 \]

a) McCormick relaxation (MC)

b) Affine relaxations from affine arithmetic (AF1)

c) Intersection of MC & AF1 on each factor

d) A priori composite relaxations (via AA)


Implementation

- IntervalArithmetic.jl\(^{24}\) used for validated calculations.
- Relaxations computed via McCormick.jl\(^{25}\) (v0.11.1)
- EAGO.jl\(^{26}\) (v0.7.0) global optimizer used to solve problems in Julia\(^{27}\) 1.6.2. Tolerances set to 1E-4.
- In-house code implementation of AF2 affine arithmetic using IntervalArithmetic.jl for validated calculations.

3. Wilhelm, M. E., McCormick.jl, (2021), GitHub repository, [https://github.com/PSORLab/McCormick.jl](https://github.com/PSORLab/McCormick.jl)
BENCHMARKING PROBLEMS

Adapted from Example 5\textsuperscript{16}

\[
\min_{x \in X} \langle c, x \rangle + \sum_{i \geq j} q_{i,j} y_i(x)y_j(x)
\]

\[
y(x) = (x_1^2, x_1^3, x_1^4, \ldots, x_n^2, x_n^3, x_n^4)
\]

\[
x \in X = [-2,2]^n
\]

25x instances for \(n \in \{5,7,9\}\) and \(v\) proportion of nonzero entries of \(Q\) with \(v \in \{0.3, 0.5, 0.7\}\) were generated for randomly such that

\[
q_{i,j} \sim U(-2,2) \quad c_i \sim U(-512,-2)
\]

<table>
<thead>
<tr>
<th>Optimizer</th>
<th>CPU Time (s)</th>
<th>% Solved (&lt;100s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCIP</td>
<td>9.87s</td>
<td>96.4%</td>
</tr>
<tr>
<td>EAGO</td>
<td>7.89s</td>
<td>96.8%</td>
</tr>
<tr>
<td>EAGO + A priori (Subgradient)</td>
<td>5.42s</td>
<td>99.5%</td>
</tr>
<tr>
<td>EAGO + A priori (Affine)</td>
<td>7.42</td>
<td>98.2%</td>
</tr>
</tbody>
</table>

Dynamic Applications

Objective

$$\min_{p \in P} \sum_{1 \leq i \leq 3} x_i(p, t_f)$$

$$0 \leq t \leq t_f = 0.1$$

$$p \in P = [0,30]^3$$

Initial Condition

$$x_0(p) = p$$

ODE System

$$x_1'(p, t) = -x_1(p, t)x_2(p, t) + x_3(p, t)$$

$$x_2'(p, t) = 2x_1(p, t)x_2(p, t)$$

$$x_3'(p, t) = x_1(p, t) + x_2(p, t)$$

Euler Discretization (h = 0.01)

$$x_{i+1,1} = x_{i,1} + h(-x_{i,1}x_{i,2} + x_{i,3})$$

$$x_{i+1,2} = x_{i,2} + h(2x_{i,1}x_{i,2})$$

$$x_{i+1,3} = x_{i,3} + h(x_{i,1} + x_{i,2})$$

Method | CPU Time (s) | Iterations |
---|---|---|
McCormick relaxation | 22.1 | 692k |
A priori relaxation (via subgradient) | 12.7 | 134k |
Future Work

- Investigate a priori relaxations for min, max, mid, and division operators
- Extension to relaxation of multilinear operators such as trilinear envelopes\(^{21,22}\)
- Evaluate effect on existing algorithms:
  - Relaxation of parametric ordinary differential equations\(^{23}\)
  - Convex/concave relaxations of implicit functions\(^{24}\)
- Multivariate treatment of constrained McCormick relaxation

Acknowledgements

Members of the process systems and operations research laboratory at the University of Connecticut (https://psor.uconn.edu/)

Funding:
National Science Foundation, Award No.: 1932723

Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.