

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 branchand-bound^{2,3}.
- Relaxed subproblems are used to compute bounds and are often derived from relaxed functions^{2,3}.
- 1. Wilhelm, M.E., and Stuber, M.D.. **EAGO.jl: easy advanced global optimization in Julia.** *Optimization Methods and Software*, 1-26.
- 2. Horst, Reiner, and Hoang Tuy. *Global optimization: Deterministic approaches*. Springer Science & Business Media, 2013.
- Puranik, Yash, and Nikolaos V. Sahinidis. Domain reduction techniques for global NLP and MINLP optimization. *Constraints* 22.3 (2017): 338-376.







 $y = \sin(z_1 + z_2)$

Initialize $\mathbf{z}^{cv}, \mathbf{z}^{cc}, \mathbf{z}^{L}, \mathbf{z}^{U}$

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

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



Convex relaxation Concave relaxation Lower bound Upper bound

4. Mitsos, A, et al. McCormick-based relaxations of algorithms. SIAM Journal on Optimization, 20, (2009): 573 - 601.

3

Factorable Function⁴



PROS

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

Repeatedly enclosing nonconvex functions with convex/concave functions.

4. Mitsos, A, et al. McCormick-based relaxations of algorithms. SIAM Journal on Optimization, 20, (2009): 573 - 601.

CONS

- Multidimensional.
- Nonsmooth
- Overly conservative.







- 4. Mitsos, A, et al. McCormick-based relaxations of algorithms. SIAM Journal on Optimization, 20, (2009): 573 601.
- 5. Wilhelm, ME; Le, AV; and Stuber. MD. Global Optimization of Stiff Dynamical Systems. AIChE Journal: Futures Issue, 65 (12), 2019.
- 6. Sahlodin, A.M., and Chachuat, B. Discretize-then-relax approach for convex/concave relaxations of the solutions of parametric ODEs. Applied Numerical Mathematics 61.7 (2011): 803-820
- 7. Scott, Joseph K., and Paul I. Barton. Improved relaxations for the parametric solutions of ODEs using differential inequalities. Journal of Global Optimization 57.1 (2013): 143-176.



- 4. Mitsos, A, et al. McCormick-based relaxations of algorithms. SIAM Journal on Optimization, 20, (2009): 573 601.
- 5. Wilhelm, ME; Le, AV; and Stuber. MD. Global Optimization of Stiff Dynamical Systems. AIChE Journal: Futures Issue, 65 (12), 2019.
- 6. Sahlodin, A.M., and Chachuat, B. Discretize-then-relax approach for convex/concave relaxations of the solutions of parametric ODEs. Applied Numerical Mathematics 61.7 (2011): 803-820
- 7. Scott, Joseph K., and Paul I. Barton. Improved relaxations for the parametric solutions of ODEs using differential inequalities. Journal of Global Optimization 57.1 (2013): 143-176.
- 8. Stuber, MD et al. Convex and concave relaxations of implicit functions. Optimization Methods and Software 30, (2015), 424-460



- 11. Tsoukalas, A., Mitsos, A. Multivariate McCormick Relaxations. Journal of Global Optimization 59, (2014): 633 662.
- 12. Khan, K. et al. Differentiable McCormick relaxations. *Journal Global Optimization* 67, (2017): 687-729.

^{10.} D Bongartz, J Najman, A Mitsos. Deterministic global optimization of steam cycles using the IAPWS-IF97 model. Optimization and Engineering 21, (2020): 1095-1131.

A Priori Estimators



Relaxation of Bilinear Term

McCormick Relaxation of Bilinear Term¹³

$$w = x \times y$$

Underestimators

(1)
$$w \ge x^L y + xy^L - x^L y^L$$

(2) $w \ge x^U y + xy^U - x^U y^U$

Overestimators

(1)
$$w \le x^U y + xy^L - x^U y^L$$

(2) $w \le xy^U + x^L y - x^L y^U$

13. McCormick, Garth P. Computability of global solutions to factorable nonconvex programs: Part I - Convex underestimating problems. *Mathematical programming* 10.1, (1976): 147-175.

Relaxation of Composite Bilinear Term

McCormick Relaxation of <u>Composite</u> Bilinear Term⁴

$$w(\mathbf{z}) = \mathbf{x}(\mathbf{z}) \times \mathbf{y}(\mathbf{z})$$

Underestimators Convex Relaxations

$$w_1^{cv}(\mathbf{z}) = \min(x^L y^{cv}(\mathbf{z}), x^L y^{cc}(\mathbf{z})) + \min(y^L x^{cv}(\mathbf{z}), y^L x^{cc}(\mathbf{z})) - x^L y^L$$
$$w_2^{cv}(\mathbf{z}) = \min(x^U y^{cv}(\mathbf{z}), x^U y^{cc}(\mathbf{z})) + \min(y^U x^{cv}(\mathbf{z}), y^U x^{cc}(\mathbf{z})) - x^U y^U$$

Overestimators Concave Relaxations

$$w_1^{cc}(\mathbf{z}) = \max(x^U y^{cv}(\mathbf{z}), x^U y^{cc}(\mathbf{z})) + \max(y^L x^{cv}(\mathbf{z}), y^L x^{cc}(\mathbf{z})) - x^U y^L$$
$$w_2^{cc}(\mathbf{z}) = \max(y^U x^{cv}(\mathbf{z}), y^U x^{cc}(\mathbf{z})) + \max(x^L y^{cv}(\mathbf{z}), x^L y^{cc}(\mathbf{z})) - x^L y^U$$

13. McCormick, Garth P. Computability of global solutions to factorable nonconvex programs: Part I - Convex underestimating problems. Mathematical programming 10.1, (1976): 147-175.

^{4.} Mitsos, A, et al. McCormick-based relaxations of algorithms. SIAM Journal on Optimization, 20, (2009): 573 - 601.

A Priori Estimators, Reduced-Space

Relaxations of <u>Composite</u> Bilinear Term A Priori $w(z) = x(z) \times y(z)$

$$w_{1}^{cv}(\mathbf{z}) = (y^{U} - a_{y})u_{x}^{cv}(\mathbf{z}) + (x^{U} - a_{x})u_{y}^{cv}(\mathbf{z}) + \min(a_{y}x^{cv}(\mathbf{z}), a_{y}x^{cc}(\mathbf{z})) + \min(a_{x}y^{cv}(\mathbf{z}), a_{x}y^{cc}(\mathbf{z})) + a_{x}a_{y} - a_{x}y^{U} - x^{U}a_{y}$$

$$w_{2}^{cv}(\mathbf{z}) = (a_{y} - y^{L})u_{x}^{cv}(\mathbf{z}) + (a_{x} - x^{L})u_{y}^{cv}(\mathbf{z}) + \min(y^{L}x^{cv}(\mathbf{z}), y^{L}x^{cc}(\mathbf{z})) + \min(y^{L}x^{cv}(\mathbf{z}), y^{L}x^{cc}(\mathbf{z})) - a_{x}a_{y}$$

$$w_{3}^{cv}(\mathbf{z}) = (y^{U} - y^{L})u_{x}^{cv}(\mathbf{z}) + \min(y^{L}x^{cv}(\mathbf{z}), y^{L}x^{cc}(\mathbf{z})) + \min(a_{x}y^{cv}(\mathbf{z}), a_{x}y^{cc}(\mathbf{z})) - a_{x}y^{U}$$

$$w_{4}^{cv}(\mathbf{z}) = (x^{U} - x^{L})u_{y}^{cv}(\mathbf{z}) + \min(a_{y}x^{cv}(\mathbf{z}), a_{y}x^{cc}(\mathbf{z})) + \min(y^{L}x^{cv}(\mathbf{z}), y^{L}x^{cc}(\mathbf{z})) - a_{y}x^{U}$$

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

 $u_x^{cv}(\mathbf{z}), u_y^{cv}(\mathbf{z}), a_x, a_y, v_x^{cc}(\mathbf{z}), v_y^{cc}(\mathbf{z}), b_x, b_y$

+ 12 additional inequalities



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** relaxations⁴ for under/overestimators

- $\begin{aligned} x(\mathbf{z}) &\geq x^{cv}(\mathbf{z}) = u_{\chi}(\mathbf{z}) = u_{\chi}^{cv}(\mathbf{z}) \\ x(\mathbf{z}) &\leq x^{cc}(\mathbf{z}) = u_{\chi}(\mathbf{z}) = u_{\chi}^{cc}(\mathbf{z}) \\ y(\mathbf{z}) &\geq y^{cv}(\mathbf{z}) = u_{y}(\mathbf{z}) = u_{y}^{cv}(\mathbf{z}) \\ y(\mathbf{z}) &\leq y^{cc}(\mathbf{z}) = u_{y}(\mathbf{z}) = u_{y}^{cc}(\mathbf{z}) \end{aligned}$
- 4. Mitsos, A, et al. **McCormick-based relaxations of algorithms.** *SIAM Journal on Optimization*, 20, (2009): 573 601.



Approach 1: McCormick Relaxation





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 framework⁴

 $\begin{aligned} x(\mathbf{z}) &\geq x^{cv}(\bar{\mathbf{z}}) + \nabla s_{x}^{cv}(\bar{\mathbf{z}})(\mathbf{z} - \bar{\mathbf{z}}) \\ x(\mathbf{z}) &\leq x^{cc}(\bar{\mathbf{z}}) + \nabla s_{x}^{cc}(\bar{\mathbf{z}})(\mathbf{z} - \bar{\mathbf{z}}) \\ y(\mathbf{z}) &\geq y^{cv}(\bar{\mathbf{z}}) + \nabla s_{y}^{cv}(\bar{\mathbf{z}})(\mathbf{z} - \bar{\mathbf{z}}) \\ y(\mathbf{z}) &\leq y^{cc}(\bar{\mathbf{z}}) + \nabla s_{y}^{cc}(\bar{\mathbf{z}})(\mathbf{z} - \bar{\mathbf{z}}) \end{aligned}$

4. Mitsos, A, et al. **McCormick-based relaxations of algorithms.** *SIAM Journal on Optimization*, 20, (2009): 573 - 601.



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 framework⁴ $x(\mathbf{z}) \ge x^{cv}(\overline{\mathbf{z}}) + \nabla s_x^{cv}(\overline{\mathbf{z}})(\mathbf{z} - \overline{\mathbf{z}}) = u_x(\mathbf{z}) = u_x^{cv}(\mathbf{z})$ $x(\mathbf{z}) \le x^{cc}(\overline{\mathbf{z}}) + \nabla s_x^{cc}(\overline{\mathbf{z}})(\mathbf{z} - \overline{\mathbf{z}}) = u_x(\mathbf{z}) = u_x^{cc}(\mathbf{z})$ $y(\mathbf{z}) \ge y^{cv}(\overline{\mathbf{z}}) + \nabla s_y^{cv}(\overline{\mathbf{z}})(\mathbf{z} - \overline{\mathbf{z}}) = u_y(\mathbf{z}) = u_y^{cv}(\mathbf{z})$ $y(\mathbf{z}) \le y^{cc}(\overline{\mathbf{z}}) + \nabla s_y^{cc}(\overline{\mathbf{z}})(\mathbf{z} - \overline{\mathbf{z}}) = u_y(\mathbf{z}) = u_y^{cc}(\mathbf{z})$

4. Mitsos, A, et al. McCormick-based relaxations of algorithms. SIAM Journal on Optimization, 20, (2009): 573 - 601.



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^{8,18-20} $a_{x} = (x^{cv}(\overline{z}) + \nabla s_{x}^{cv}(\overline{z})(Z - \overline{z}))^{U} \ge u_{x}^{cv}(z)$ $b_{x} = (x^{cc}(\overline{z}) + \nabla s_{x}^{cc}(\overline{z})(Z - \overline{z}))^{L} \le u_{x}^{cc}(z)$ $a_{y} = (y^{cv}(\overline{z}) + \nabla s_{y}^{cv}(\overline{z})(Z - \overline{z}))^{U} \ge u_{y}^{cv}(z)$

- $b_{y} = \left(y^{cc}(\overline{\mathbf{z}}) + \nabla s_{y}^{cc}(\overline{\mathbf{z}})(\mathbf{Z} \overline{\mathbf{z}})\right)^{L} \le u_{y}^{cc}(\mathbf{z})$
- 4. Mitsos, A, et al. McCormick-based relaxations of algorithms. *SIAM Journal on Optimization*, 20, (2009): 573 601.
- 6. Sahlodin, A.M., and Chachuat, B. **Discretize-then-relax approach for convex/concave relaxations** of the solutions of parametric ODEs. *Applied Numerical Mathematics* 61.7 (2011): 803-820.
- 8. Stuber, Matthew D., Joseph K. Scott, and Paul I. Barton. **Convex and concave relaxations of implicit functions.** *Optimization Methods and Software* 30.3 (2015): 424-460.
- 9. Najman, J., Mitsos, A. **Tighter McCormick relaxations through subgradient propagation.** *Journal of Global Optimization* **75**, 565–593 (2019).



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 Relaxation⁸
 - Multivariate McCormick Relaxation¹¹
 - A priori Calculation (Subgradient)
 - A priori Calculation (Max-Concave)
 - 8. Mitsos, A, et al. McCormick-based relaxations of algorithms. *SIAM Journal on Optimization*, 20, (2009): 573 601.
 - 11. Tsoukalas, A., and Mitsos, A. **Multivariate McCormick relaxations.** Journal of Global Optimization 59.2-3 (2014): 633-662.

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 Relaxation⁸
 - Multivariate McCormick Relaxation¹¹
 - A priori Calculation (Subgradient)
 - A priori Calculation (Max-Concave)
 - 8. Mitsos, A, et al. McCormick-based relaxations of algorithms. *SIAM Journal on Optimization*, 20, (2009): 573 601.
 - 11. Tsoukalas, A., and Mitsos, A. **Multivariate McCormick relaxations.** *Journal of Global Optimization* 59.2-3 (2014): 633-662.



1-D Example



Approach 3: Affine Arithmetic



16. Messine, F.: Extensions of affine arithmetic: Application to unconstrained global optimization. *JUCS* (11), 992–1015 (2002).

17. Ninin, J., Messine, F., Hansen, P.: A reliable affine relaxation method for global optimization. 4OR13 (3), 247–277 (2015).

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)⁸
- b) Affine relaxations from affine arithmetic (AF1)^{22,23}
- c) Intersection of MC & AF1 on each factor
- d) A priori composite relaxations (via AA)



- 4. Mitsos, A, et al. McCormick-based relaxations of algorithms. SIAM Journal on Optimization, 20, (2009): 573 601.
- 16. Messine, F.: Extensions of affine arithmetic: Application to unconstrained global optimization. JUCS (11), 992–1015 (2002).
- 17. Ninin, J., Messine, F., Hansen, P.: A reliable affine relaxation method for global optimization. 4OR13 (3), 247–277 (2015).



Implementation

- ✤ IntervalArithmetic.jl²⁴ used for validated calculations.
- Relaxations computed via McCormick.jl²⁵ (v0.11.1)
- EAGO.jl²⁶ (v0.7.0) global optimizer used to solve problems in Julia²⁷ 1.6.2. Tolerances set to 1E-4.
- In-house code implementation of AF2 affine arithmetic using IntervalArithmetic.jl for validated calculations.



https://www.github.com/PSORLab/EAGO.jl



1. Wilhelm, M. E., and M. D. Stuber. EAGO. jl: easy advanced global optimization in Julia. Optimization Methods and Software (2020): 1-26

- 18. Sanders, D.P., Benet, L. (2021), GitHub repository, https://github.com/JuliaIntervals/IntervalArithmetic.jl
- 19. Wilhelm, M. E., **McCormick.jl**, (2021), GitHub repository, https://github.com/PSORLab/McCormick.jl.

20. Bezanson, Jeff, et al. Julia: A fresh approach to numerical computing. SIAM review 59.1 (2017): 65-98.

Benchmarking Problems

Adapted from Example 5¹⁶

$$\min_{\mathbf{x} \in X} \langle \mathbf{c}, \mathbf{x} \rangle + \sum_{i \ge j} q_{i,j} y_i(\mathbf{x}) y_j(\mathbf{x})$$
$$\mathbf{y}(\mathbf{x}) = (x_1^2, x_1^3, x_1^4, \cdots, x_n^2, x_n^3, x_n^4)$$
$$\mathbf{x} \in X = [-2, 2]^n$$

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

$$q_{i,j} \sim U(-2,2)$$
 $c_i \sim U(-512,-2)$

Optimizer	CPU Time (s)	% Solved (<100s)
SCIP	9.87s	96.4%
EAGO	7.89s	96.8%
EAGO + A priori (Subgradient)	5.42s	99.5%
EAGO + A priori (Affine)	7.42	98.2%

21. He, et al. A new framework to relax composite functions in nonlinear programs. Mathematical Programming, (2020): 1 - 40.

Dynamic Applications

Objective
$$\min_{\mathbf{p} \in P} \sum_{1 \le i \le 3} \mathbf{x}_i(\mathbf{p}, t_f)$$
$$0 \le t \le t_f = 0.1$$
$$\mathbf{p} \in P = [0, 30]^3$$

Initial Condition
$$\mathbf{x}_0(\mathbf{p}) = \mathbf{p}$$

ODE System

$$x'_{1}(\mathbf{p}, t) = -x_{1}(\mathbf{p}, t)x_{2}(\mathbf{p}, t) + x_{3}(\mathbf{p}, t)$$
$$x'_{2}(\mathbf{p}, t) = 2x_{1}(\mathbf{p}, t)x_{2}(\mathbf{p}, t)$$
$$x'_{3}(\mathbf{p}, t) = x_{1}(\mathbf{p}, t) + x_{2}(\mathbf{p}, t)$$

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

Euler Discretization (h = 0.01) $x_{i+1,1} = x_{i,1} + h(-x_{i,1}x_{i,2} + x_{i,3})$ $x_{i+1,2} = x_{i,2} + h(2x_{i,1}x_{i,2})$ $x_{i+1,3} = x_{i,3} + h(x_{i,1} + x_{i,2})$

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 equations²³
- Convex/concave relaxations of implicit functions²⁴

□ Multivariate treatment of constrained McCormick relaxation

24. Stuber, MD et al. Convex and concave relaxations of implicit functions. Optimization Methods and Software 30, (2015), 424-460.

^{4.} Tsoukalas, A., and Mitsos, A. Multivariate McCormick relaxations. Journal of Global Optimization 59.2-3 (2014): 633-662.

^{21.} Meyer, C.A., and Floudas, C.A. Trilinear monomials with positive or negative domains: Facets of the convex and concave envelopes. Frontiers in global optimization, (2004). 327-352.

^{22.} Meyer, C.A., and Floudas, C.A. Trilinear monomials with mixed sign domains: Facets of the convex and concave envelopes. Journal of Global Optimization 29.2 (2004): 125-155.

^{23.} Harwood, S.M., and Barton, P.I. Affine relaxations for the solutions of constrained parametric ordinary differential equations. Optimal Control Applications and Methods 39.2 (2018): 427-448.

Acknowledgements

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



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.