Strong Linear Relaxations for Global Optimization Problems with Multilinear Terms

Jim Luedtke

Collaborators: Mahdi Namazifar, Jeff Linderoth

University of Wisconsin-Madison

ASCR Project: Next Generations Solvers for Mixed-Integer Nonlinear Programming: Structure, Search, and Implementation

2011 DOE Applied Mathematics Program Meeting
October 19, 2011

Motivation

A generic global optimization problem

$$\begin{aligned} &\min f(x) \\ \text{subjec to } & h_i(x) + \sum_{t \in T_i} a_{it} \prod_{j \in J_t} x_j \leq b_i, \quad i = 1, 2, \ldots, m \\ & \ell \leq x \leq u \end{aligned}$$

- We seek a global optimal solution $\Rightarrow$ Need tight convex relaxations
- We study techniques for relaxing the multilinear terms: $\prod_{j \in J_t} x_j$

Multilinear terms appear in many applications

- Electricity transmission, nuclear core reload, chemical blending processes, any problem with *bilinear* terms (nonconvex quadratic)

Agenda

1. Comparison of existing linear relaxations
2. Term-cover relaxation approach
3. Constructing term-covers
4. Computational experience

What do we relax?

Reformulated global optimization problem

$$\begin{aligned} \min\ & f(x) \\ \text{subject to } & h_i(x) + \sum_{t \in T_i} a_{it} z_t \le b_i, \quad i = 1, 2, \ldots, m \\ & (x, z) \in X \end{aligned}$$

Where $T = \bigcup_i T_i$ and

$$X = \left\{ x \in [\ell, u], z \in \mathbb{R}^{|T|} \mid z_t = \prod_{j \in J_t} x_j, \ \ t \in T \right\}$$

We seek strong (linear) relaxations of X

- With linear relaxations of f and h_i yields LP relaxation
- Ideal case: conv(X)

McCormick [1976] relaxation of bilinear terms

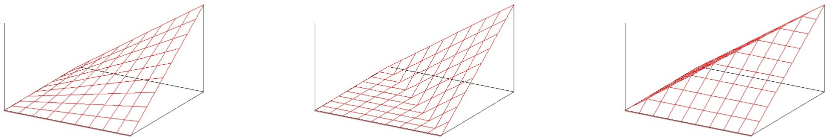
Bilinear special case

$$B = \left\{ (x,z) \in [0,1]^n \times \mathbb{R}^{|T|} \;\middle|\; z_{ij} = x_i x_j, \;\; \forall (i,j) \in T \right\}$$

Relaxation:

$$z_{ij} \geq \max\{0, x_i + x_j - 1\}, \quad z_{ij} \leq \min\{x_i, x_j\}, \quad \forall (i,j) \in T$$

- Similar inequalities apply when $x \in [\ell, u]$
- Inequalities define convex hull for individual terms



Convex Hull

Let $\chi^1, \ldots, \chi^{2^n}$ be vertices of $[\ell, u]$

Convex hull extended formulation (Rikun 1997, Sherali 1997)

$(x, z) \in \text{conv}(X)$ if and only if there exists $\lambda \in \mathbb{R}_+^{2^n}$ such that

$$\sum_{k=1}^{2^n} \lambda_k = 1, \quad x = \sum_{k=1}^{2^n} \lambda_k \chi^k, \quad z_t = \sum_{k=1}^{2^n} \lambda_k \phi_t(\chi^k), \quad \forall t \in T$$

where $\phi_t(\chi^k) = \prod_{j \in J_t} \chi_j^k$

Good news: Polyhedral

Bad news: 2^n variables

Can we compromise?

Our goal: Find a relaxation that is

- Stronger than McCormick
- More compact than full convex hull

Related work

- Bao, Sahinidis, and Tawarmalani (2010)
- Semidefinite programming relaxations

How much better can convex hull formulation be?

Results from L., Namazifar, and Linderoth (2010)

- Consider a single bilinear function: $z = \sum_{(i,j)\in T} a_{ij} x_i x_j$
  - Let G be a graph with edges defined by T, and χ be the coloring number of G
- If G is a bipartite graph, McCormick relaxtion $\equiv$ convex hull (Coppersmith, et. al.)
- If $a_{ij} \geq 0$, maximum McCormick error is within factor $2 - 2/\chi$ of maximum convex hull error
- For general bilinear, maximum McCormick error is within factor $2(\chi - 1)$ of maximum convex hull error

Conclusion: Convex hull not always better, but may be significantly better for “dense” problems

Agenda

1. Comparison of existing linear relaxations
2. Term-cover relaxation approach
3. Constructing term-covers
4. Computational experience

What's a term-cover?

Definition

A finite collection of subsets of $\{1,\ldots,n\}$, $\mathcal{C} = \{C_g\}_{g\in G}$, is called a term-cover of T if for all $t \in T$ there exists $g \in G$ such that

$$J_t \subseteq C_g.$$

$$T = \big\{\{1,2\},\{1,3\},\{1,4\},\{1,5\},\{1,6\},\{2,3\},\\ \{2,4\},\{2,5\},\{3,4\},\{3,5\},\{4,6\},\{5,7\},\{6,7\}\big\}$$

$$\mathcal{C} = \big\{\{1,2,3,4\},\{2,3,4,5\},\{1,5,6,7\},\{4,6\}\big\}$$

Generalizes both McCormick and convex hull

- McCormick relaxation: $V_t = J_t$, $t \in T$
- Convex hull: $V_1 = \{1,\ldots,n\}$

Constructing a relaxation based on a term-cover

Idea: Introduce the convex hull formulation for each "group" in the term-cover

- Let $\{\chi^{g,k}\}_{k=1}^{2^{|C_g|}} = \text{Vert}(\{\ell_j \leq x_j \leq u_j, j \in C_g\})$ for all $g \in G$

Relaxation, $TC_{\mathcal{C}}(X)$, is given by:

$$x_{C_g} = \sum_{k=1}^{2^{|C_g|}} \lambda_k^g \chi^{g,k}, \quad \sum_{k=1}^{2^{|C_g|}} \lambda_k^g = 1, \quad \forall g \in G$$

$$z_t = \sum_{k=1}^{2^{|C_g|}} \lambda_k^g \prod_{j \in J_t} \chi_j^{g,k}, \quad \forall g \in G, t \in T : J_t \subseteq C_g$$

Agenda

Building Good Term Covers

Overarching Goal

Get strong bounds fast

Questions

1. What should be the maximum size (σ) of elements of the term cover?
   - (A): Not too big! You need 2^σ variables for that element.
2. Which (original) variables should appear in the same element of the term cover?
3. How many elements should the term cover contain?

Thoughts on Building Good Terms Covers

1. Variables from a dense part of the graph (hypergraph) associated with the function should be in the same element
2. Covering terms multiple times can help
3. Terms with large coefficients may be more important

- We (read Mahdi) tried lots of different strategies, but none did as well as a simple greedy two-phase heuristic

Two-phase heuristic

Parameters:	Max size $\sigma \geq 2$, Augmentation factor $\nu \geq 1$
Data:	Terms T_i, weights w_i
Graph:	A node for each variable, edge for each term

1. Construct an initial term-cover
   - Heuristically choose a max-weight subgraph of size $\leq \sigma$
   - Remove covered terms and repeat
2. Augment the term-cover
   - Adjust weights according to how many times each term has been covered
   - Heuristically choose a max-weight subgraph of size $\leq \sigma$
   - Repeat until have ν times more elements than initial term-cove

Empirical tests: $\sigma = 6$ and $\nu = 2$ are reasonable

Is It Good?

- MANY different methods were tried for building good term-covers
- This greedy two-phase heuristic was by far the best

A Small Experiment

- Compared bounds to 40 random term covers of the same size.
- Greedy heuristic always beats the best of the 40 random covers.

Agenda

Computational Experiments

Questions to Answer

1. How close is term-cover relaxation to convex hull relaxation?
2. Compare “best” term-cover relaxation to McCormick and SDP relaxations

SDP Relaxation for QCQPs

- Introduce all “product variables”, $y_{ij} = x_i x_j$
- Relax the constraint $Y - xx^T = 0$ to $Y - xx^T \succeq 0$
- This is the state of the art relaxation for nonconvex quadratic problems

Test Instances

- Box-constrained QP: (`spar-n-density-#`)

$$\min_{x\in[0,1]^n} \{x^T Qx + d^T x\}$$

- Quadratically Constrained QP: (`qcp-n-#cons-density-#`)

$$\min_{x\in[0,1]^n} \{x^T Qx + d^T x \mid x^T Q_i x + d_i^t x \leq b_i \ \forall i = 1, \ldots, m\}$$

- Also tested: QP with linear constraints, quadrilinear programming.

Computational Details

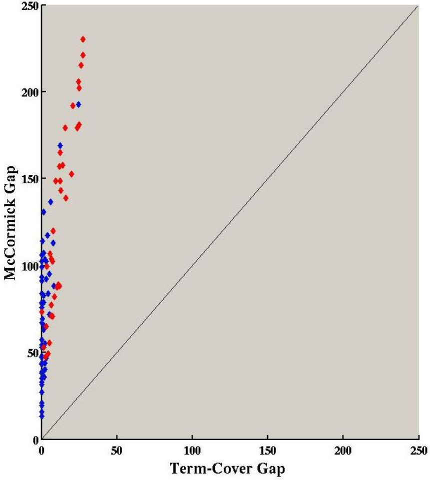
- Relaxations implemented in MINOTAUR: Mixed Integer Nonlinear Optimization Toolkit: Algorithms, Underestimators, Relaxations
  - Lead developer: Ashutosh Mahajan
- LP's solved with Coin's CLP
- SDP's built/solved with YALMIP/SeDuMi
- Quadratic terms relaxed in a manner similar to BARON:
  - Approximate $z = x^2, x \in [\ell, u]$ with one "secant" on top, and $\kappa (= 5)$ linearizations underneath.

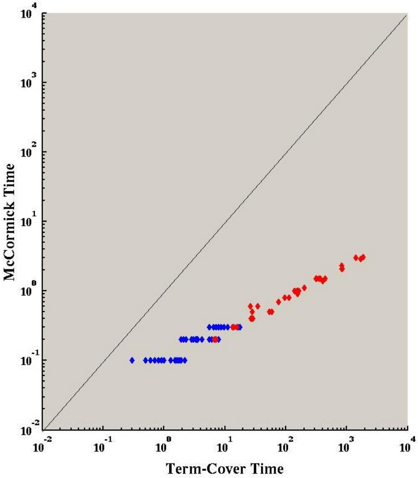
How Close is Term Cover to Convex Hull?

$\nu = 1$	McCormick		Term-Cover 4		Term-Cover 8		Convex Hull	
Problem	**G%**	**T**	**G%**	**T**	**G%**	**T**	**G%**	**T**
qcp-15-05-050-1	23.8	0.1	23.8	0.1	23.8	0.2	23.8	23.9
qcp-15-05-050-2	41.4	0.1	41.3	0.1	38.1	0.1	38.1	29.3
qcp-15-05-075-1	24.8	0.1	16.7	0.1	16.7	0.2	16.7	30.2
qcp-15-05-075-2	19.8	0.1	19.8	0.1	19.8	0.1	19.8	24.9
qcp-15-05-100-1	99.2	0.1	47.7	0.1	47.7	0.1	47.7	21.7
qcp-15-05-100-2	101.4	0.1	58.7	0.1	49.8	0.1	49.8	27.8
qcp-15-10-050-1	115.3	0.1	115.3	0.1	115.3	0.2	115.3	26.6
qcp-15-10-050-2	31.2	0.1	31.2	0.1	31.2	0.2	31.2	30.5
qcp-15-10-075-1	137.1	0.1	95.8	0.1	80.6	0.2	80.6	26.7
qcp-15-10-075-2	21.1	0.1	11.7	0.1	11.7	0.2	11.7	24.3
qcp-15-10-100-1	1.5	0.1	0.0	0.1	0.0	0.2	0.0	27.2
qcp-15-10-100-2	1078.7	0.1	697.9	0.1	492.6	0.2	478.5	28.2

- G%: Gap to optimal solution value. (Instances solved by Couenne)

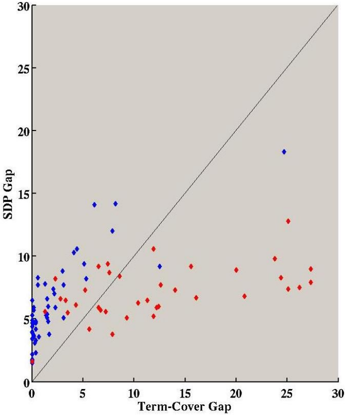
Box QP – Term-Cover vs. McCormick

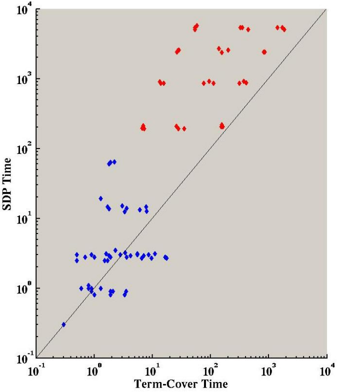




Red points correspond to large instances: 70 – 100 variables
Blue points correspond to small and medium instances: 20 – 60 variables

Box QP – Term-Cover vs. SDP





Surprising!

Linear term-cover relaxation is competitive with SDP relaxations.

QCQP

	SDP		Term-Cover $\sigma = 6, \nu = 2$		McCormick	
Problems	**G%**	**T**	**G%**	**T**	**G%**	**T**
q020-025-1	12.1	1.5	17.0	0.2	17.2	0.0
q020-025-2	13.8	0.4	28.9	0.3	28.9	0.1
q020-050-1	9.3	0.4	15.4	0.5	17.9	0.1
q020-050-2	11.1	0.4	16.8	0.3	43.9	0.1
q020-075-1	5.0	0.4	8.1	0.6	30.1	0.1
q020-075-2	15.5	0.3	18.4	0.4	58.6	0.1
q040-025-1	7.5	3.6	5.3	2.4	16.7	0.2
q040-025-2	18.8	4.0	7.7	15.5	27.3	0.6
q040-050-1	10.9	3.9	5.5	20.4	53.3	0.8
q040-050-2	7.9	4.3	4.0	13.2	23.3	0.9
q040-075-1	27.1	4.0	45.8	19.0	201.1	0.8
q040-075-2	19.4	3.9	24.7	14.3	133.9	0.8
q060-025-1	13.7	92.9	10.8	134.1	45.4	2.1
q060-025-2	22.5	84.7	30.7	168.4	63.3	2.7
q060-050-1	18.8	101.3	34.6	303.3	129.8	3.4
q060-050-2	14.5	94.8	23.0	222.4	91.1	3.2
q060-075-1	24.9	79.5	55.1	150.3	201.4	3.5
q060-075-2	21.8	95.7	51.7	103.1	191.3	3.7
q080-025-1	11.3	1467.1	12.2	659.4	39.0	6.1
q080-025-2	15.8	1515.6	25.5	1822.2	54.3	8.8
q080-050-1	29.1	1445.8	65.8	2112.7	186.1	10.6
q080-050-2	14.6	1599.4	36.1	1285.8	110.0	11.6
q080-075-1	24.2	1438.8	67.3	1166.4	206.3	11.8
q080-075-2	29.8	1378.5	83.2	1225.5	242.4	12.5
Average	**16.6**	**392.6**	**28.9**	**393.4**	**92.2**	**3.5**

Parting thoughts

Key findings

- Strong linear relaxations of multiliear terms are possible
- Surprisingly competitive with SDP
- Advantages of linear relaxations: fast re-solve, easy combination with relaxation of other components

Continuing work

- Embed relaxation into branch and bound scheme
- Transform quadratic instances to better separate convex and concave parts. (“Eigen”-reformulation)
- Develop a dynamic term covering strategy
    - Potential to be faster *and* stronger