A Different Perspective on Perspective Cuts

Jeff Linderoth

Unemployed
linderot@cs.wisc.edu

Oktay Günlük

Mathematical Sciences Department
IBM T.J. Watson Research Center
gunluk@us.ibm.com

MIP 2007
Centre de recherches mathématiques
Université de Montréal
August 1, 2007

Indicator MINLPs

- We focus on (convex) MINLPs that are driven by 0-1 indicator variables $z_i, i \in \mathcal{I}$
- Each indicator variable i controls a collection of variables V_i
- If $z_i = 0$, the components of x controlled by z_i must collapse to a point: $z_i = 0 \Rightarrow x_{V_i} = \hat{x}_{V_i}$
  - WLOG $\hat{x}_{V_i} = 0$ from now on
- If $z_i = 1$, the components of x controlled by z_i belong to a convex set $z_i = 1 \Rightarrow x_{V_i} \in \Gamma_i$
- Γ_i is specified by (convex) nonlinear inequality constraints and bounds on the variables

$$\Gamma_i \stackrel{\text{def}}{=} \{x_{V_i} \mid f_k(x_{V_i}) \leq 0 \ \forall k \in K_i, l \leq x_{V_i} \leq u\}.$$

Indicator MINLPs

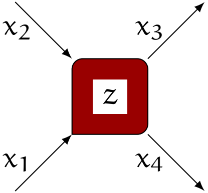
$$
\begin{array}{llcll}
\min & c^T x + d^T z & & & \\
\text{s.t.} & g_m(x,z) & \leq & 0 & \forall m \in M \\
& z_i f_k(x_{V_i}) & \leq & 0 & \forall i \in \mathcal{I} \ \forall k \in K_i \\
& \ell_j z_i \leq x_j & \leq & u_j z_i & \forall i \in \mathcal{I} \ \forall j \in V_i \\
& x \in X & & z \in Z \cap \mathbb{B}^p, &
\end{array}
$$

- X, Z polyhedral sets
- Typically, $g_m(x,z) = \bar{g}_m(x) + a_m^T z$ is linear in z, or even $a_m = 0$.
- If $z \in Z \cap \mathbb{B}^p$ is fixed, then the problem is convex.

Indicators Everywhere

Process Flow Applications

- $z = 0 \Rightarrow x_1 = x_2 = x_3 = x_4 = 0$
- $z = 1 \Rightarrow f(x_1, x_2, x_3, x_4) \leq 0$



Separable Function Epigraphs

$$y_i \geq f_i(x_i) \ \forall i \in \mathcal{I}$$

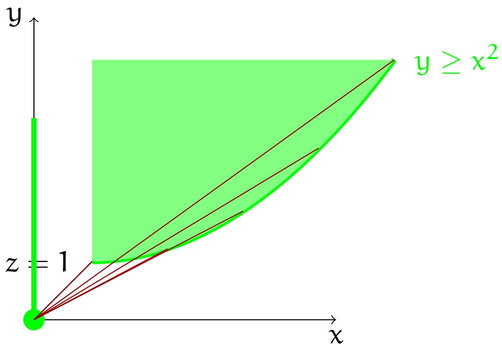
$$\ell z_i \leq x_i \leq u z_i \ \forall i \in \mathcal{I}$$

- Note that here I am already lying
- $z = 0$ does not imply $y = 0$
- Nevertheless, results apply to epigraph-type indicator MINLPs.

A Very Simple Example

$$R \stackrel{\text{def}}{=} \left\{(x, y, z) \in \mathbb{R}^2 \times \mathbb{B} \mid y \geq x^2, 0 \leq x \leq uz\right\}$$

- $z = 0 \Rightarrow x = 0, y \geq 0$
- $z = 1 \Rightarrow x \leq u, y \geq x^2$



Deep Insights

- $\text{conv}(R) \equiv$ line connecting $(0, 0, 0)$ to $y = x^2$ in the $z = 1$ plane

Characterization of Convex Hull

- Work out the algebra to get:

Deep Theorem #1

$$\mathrm{conv}(R) = \left\{ (x, y, z) \in \mathbb{R}^3 \mid yz \geq x^2, 0 \leq x \leq uz, 0 \leq z \leq 1, y \geq 0 \right\}$$

$$x^2 \leq yz, y, z \geq 0 \equiv$$

Second Order Cone Programming

- There are effective and robust algorithms for optimizing linear objectives over $\mathrm{conv}(R)$

Higher Dimensions

- Using an extended formulation, we can describe the convex hull of a higher-dimensional analogue of R:

$$Q \stackrel{\text{def}}{=} \left\{ (w, x, z) \in \mathbb{R}^{1+n} \times \mathbb{B}^n \mid w \geq \sum_{i=1}^{n} q_i x_i^2, \; u_i z_i \geq x_i \geq 0, \forall i \right\}$$

- First we write an extended formulation of Q, introducing variables y_i:

$$\bar{Q} \stackrel{\text{def}}{=} \left\{ (w, x, y, z) \in \mathbb{R}^{1+3n} \mid w \geq \sum_i q_i y_i, (x_i, y_i, z_i) \in R_i, \;\; \forall i \right\}$$

$$R_i \stackrel{\text{def}}{=} \left\{ (x_i, y_i, z_i) \in \mathbb{R}^2 \times \mathbb{B} \mid y_i \geq x_i^2, 0 \leq x_i \leq u_i z_i \right\}$$

Extended Formulations

- $\bar{Q}$ is indeed an extended formulation in the sense that projecting out the y variables from $\bar{Q}$ gives Q: $\mathrm{Proj}_{(w,x,z)}\,\bar{Q} = Q$.
- The convex hull of $\bar{Q}$ is obtained by replacing R_i with its convex hull description $\mathrm{conv}(R_i)$:

$$\mathrm{conv}(\bar{Q}) = \Big\{ w \in \mathbb{R},\ x \in \mathbb{R}^n, y \in \mathbb{R}^n, z \in \mathbb{R}^n \ :\ w \ge \sum_i q_i y_i,$$
$$(x_i, y_i, z_i) \in \mathrm{conv}(R_i),\quad i = 1, 2, \ldots, n \Big\}.$$

- Again, the description of $\mathrm{conv}(\bar{Q})$ is SOC-representable.
- You get one rotated cone for each i

Descriptions in the Original Space

- We can also write also write a convex hull description in the original space of variables, by projecting out y:

$$Q^c = \Big\{ (w,x,z) \in \mathbb{R}^{1+n+n} : \\ w \prod_{i \in S} z_i \geq \sum_{i \in S} \left(q_i x_i^2 \prod_{l \in S \setminus \{i\}} z_l \right) \quad S \subseteq \{1,2,\ldots,n\} \\ u_i z_i \geq x_i \geq 0, \quad x_i \geq 0, \quad i = 1,2,\ldots,n \Big\} \quad (\Pi)$$

Theorem

$$\mathrm{Proj}_{(w,x,z)}(\bar{Q}^c) = Q^c = \mathrm{conv}(Q).$$

- Q^c consists of an exponential number of nonlinear inequalities.

Extending the Intuition

- To deal with general convex sets, let $W = W^1 \cup W^0$:

$$\begin{aligned} W^0 &= \{(x,z) \in \mathbb{R}^{n+1} \mid x = 0, z = 0\} \\ W^1 &= \{(x,z) \in \mathbb{R}^{n+1} \mid f_k(x) \leq 0 \text{ for } k \in K, u \geq x \geq 0, z = 1\} \end{aligned}$$

- Write an extended formulation (XF) for $\mathrm{conv}(W)$

$$\Big\{(x, x_0, x_1, z, z_0, z_1, \alpha) \in \mathbb{R}^{3n+4} \mid 1 \geq \alpha \geq 0, x^0 = 0, z^0 = 0$$
$$x = \alpha x^1 + (1-\alpha)x^0, z = \alpha z^1 + (1-\alpha)z^0,$$
$$f_i(x^1) \leq 0 \text{ for } i \in I, u \geq x^1 \geq 0, z^1 = 1\Big\}.$$

Simplify, Simplify, Simplify

- Substitute out x^0, z^0 and z^1: They are fixed in (XF)
- $z = \alpha$ after these substitutions, so substitute it out as well.
- $x = \alpha x^1 = z x^1$, so we can eliminate x^1 by replacing it with x/z provided that $z > 0$.

Lemma

If W^1 is convex, then $\mathrm{conv}(W) = W^- \cup W^0$, where

$$W^- = \left\{(x, z) \in \mathbb{R}^{n+1} \mid f_k(x/z) \leq 0 \;\forall k \in K, uz \geq x \geq 0, 1 \geq z > 0\right\}$$

Lemma Extension

$$\mathrm{conv}(W) = \mathrm{closure}(W^-)$$

Convexify, Convexify, Convexify

- Note: $f_k(x/z)$ is not necessarily convex, even if $f_k(x)$ is.
- However, $zf_k(x/z)$ is convex if $f_k(x)$ is.
- Multiplying both sides of the inequality by $z > 0$ doesn't change the set W^-:

$$W^- = \left\{(x, z) \in \mathbb{R}^{n+1} \mid zf_k(x/z) \leq 0 \ \forall k \in K, uz \geq x \geq 0, 1 \geq z > 0\right\}$$

- You can, if you wish, multiply by z^p

Giving You Some Perspective

- For a convex function $f(x) : \mathbb{R}^n \to \mathbb{R}$, the function

$$\mathcal{P}(f(z,x)) = zf(x/z)$$

  is known as the perspective function of f
- The epigraph of $\mathcal{P}(f(z,x))$ is a cone pointed at the origin whose lower shape is $f(x)$

Exploiting Your Perspective

- If z_i is an indicator that the (nonlinear, convex) inequality $f(x) \leq 0$ must hold, (otherwise $x = 0$), replace the inequality with its perspective version:

$$z_i f(x/z_i) \leq 0$$

- The resulting (convex) inequality is a much tighter relaxation of the feasible region.

An Axioma Connection

Stubbs (1996)

- In his Ph.D. thesis, Stubbs gives (without proof) $\mathrm{conv}(\bar{Q})$, our original (high-dimensional) set

Ceria and Soares (1999)

- Describe $K = \cup_{i \in M} K_i$, with $K_i = \{x \mid f_i(x) \leq 0\}$ in a higher-dimensional space.
- $x \in \mathrm{conv}(K) \Leftrightarrow$

$$x = \sum_{i \in M} \lambda_i x_i, \mathcal{P}(f_i(\lambda_i, x_i)) \leq 0, \lambda \in \Delta_{|M|}$$

Other Smart People

Frangioni and Gentile (2006)

- Study: $y \geq f(x), x \leq uz$, give perspective cut:

$$y \geq f(x) + \nabla f(x)^T (x - \hat{x}) - (\hat{x}^T \nabla f(\hat{x}) + f(\hat{x}))(z - 1)$$

- This is first-order Taylor expansion of perspective $zf(x/z) + y \leq 0$ about $(\hat{x}, f(\hat{x}), 1)$
- Feasible inequality by convexity of $f(x)$

Aktürk, Atamtürk, and Gürel (2007)

- Apply perspective reformulation (of epigraph indicator MINLP) to nonlinear machine scheduling problem
- Explain that formulations are representable as SOCP.

Facility Location

- M: Facilities
- N: Customers
- x_{ij}: percentage of customer i's demand served from facility j
- $z_i = 1 \Leftrightarrow$ facility i is opened
- Fixed cost for opening facility i
- Quadratic cost for serving j from i
- Problem studied by Günlük, Lee, and Weismantel ('07), and classes of strong cutting planes derived

Separable Quadratic UFL—Formulation

$$z^* \stackrel{\text{def}}{=} \min \sum_{i \in M} c_i z_i + \sum_{i \in M} \sum_{j \in N} q_{ij} x_{ij}^2 y_{ij}$$

subject to

$$\begin{array}{rcll} x_{ij} & \leq & z_i & \forall i \in M, \forall j \in N \\ \sum_{i \in M} x_{ij} & = & 1 & \forall j \in N \\ x_{ij} & \geq & 0 & \forall i \in M, \forall j \in N \\ z_i & \in & \{0,1\} & \forall i \in M \\ x_{ij}^2 - z_i y_{ij} & \leq & 0 & \forall i \in M, \forall j \in N \end{array}$$

Strength of Relaxations

- z_{R}: Value of NLP relaxation
- z_{GLW}: Value of NLP relaxation after GLW cuts
- z_{P}: Value of perspective relaxation
- z^*: Optimal solution value

$\lvert M \rvert$	$\lvert N \rvert$	z_{R}	z_{GLW}	z_{P}	z^*
10	30	140.6	326.4	346.5	348.7
15	50	141.3	312.2	380.0	384.1
20	65	122.5	248.7	288.9	289.3
25	80	121.3	260.1	314.8	315.8
30	100	128.0	327.0	391.7	393.2

Woo Hoo!

Design of Uncongested Network

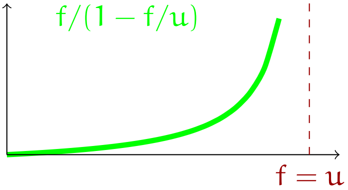
- Capacitated directed network: $G = (N, \mathcal{A})$
- Set of commodities: K
- Node demands: b_i^k $\forall i \in N, \forall k \in K$
- Each arc $(i, j) \in \mathcal{A}$ has
  - Fixed cost: c_{ij}
  - Capacity: u_{ij}
  - Queueing weight: r_{ij}

- $z_{ij} \in \{0, 1\}$: Indicates whether arc $(i, j) \in \mathcal{A}$ is opened.
- x_{ij}^k: The quantity of commodity k routed on arc (i, j)

Network Design

- Let $f_{ij} \stackrel{\text{def}}{=} \sum_{k \in K} x_{ij}^k$ be the flow on arc (i,j).
- A measure of queueing delay is:

$$\rho(f) \stackrel{\text{def}}{=} \sum_{(i,j)\in A} r_{ij} \frac{f_{ij}}{1 - f_{ij}/u_{ij}}$$



Our Network Design Problem

Design network to keep total queueing delay less than a given value β, and this is to be accomplished at minimum cost.

Network Design Formulation

$$
\begin{array}{lrcll}
 & \min \sum_{(i,j)\in A} c_{ij} z_{ij} & & & \\
\text{s.t.} & \sum_{(j,i)\in A} x_{ij}^k - \sum_{(i,j)\in A} x_{ij}^k & = & b_i^k & \forall i \in N, \forall k \in K \\
 & \sum_{k\in K} x_{ij}^k - f_{ij} & = & 0 & \forall (i,j) \in A \\
 & f_{ij} & \leq & u_{ij} z_{ij} & \forall (i,j) \in A \\
 & y_{ij} & \geq & \dfrac{r_{ij} f_{ij}}{1 - f_{ij}/u_{ij}} & \forall (i,j) \in A \\
 & \sum_{(i,j)\in A} y_{ij} & \leq & \beta &
\end{array}
$$

Perspective Formulations and Cones

- Consider the nonlinear inequality:

$$y \geq \frac{rf}{1 - f/u} \Leftrightarrow ruf \leq y(u - f)$$

- Since $z_{ij} = 0 \Rightarrow f_{ij} = 0$, we can write the perspective reformulation:

$$y/z \geq \frac{rf/z}{1 - f/zu} \Leftrightarrow ruzf \leq y(uz - f)$$

Cones Are Everywhere!

- The inequalities $ruf \leq y(u - f)$ and $urfz \leq y(uz - f)$ are SOC-representable:

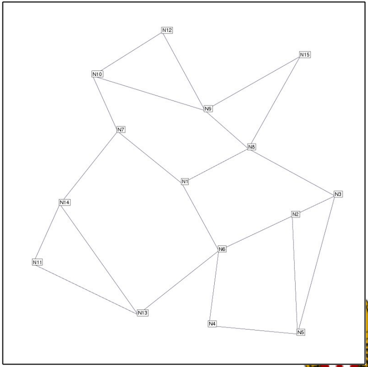
$$\begin{aligned} ruf \leq y(u - f) \quad &\Leftrightarrow \quad rf^2 \leq (y - rf)(u - f) \\ rufz \leq y(uz - f) \quad &\Leftrightarrow \quad rf^2 \leq (y - rf)(uz - f) \end{aligned}$$

since $y \geq rf$, $u \geq f$, $uz \geq f$

Results (Under Construction)

- ZIB SNDLIB instance: ATL.
- $|N| = |K| = 15$, $|A| = 22$
- Instance solved using (beta) version of Mosek (v5) conic MIP solver
- No fancy cutting planes (cut-set inequalities) added

ATL Network



Results

	Nodes	Time
No Perspective	3686	517.1
W/Perspective	414	52.5

Conclusions

Other Conclusions

- Strong reformulations for MINLPs are likely to be just as important as they are for MILPs
- Strong formulations for MINLPs may require nonlinear inequalities. (Duh!)
- Much of the work we present here has (recently) found its way into the literature.

Our "contributions"

- Give convex hull for the union of a (general) bounded convex set and a point
- Give description in original space of variables
- Exploit SOC-representability of strong reformulations to solve instances much more effectively