MIXED-INTEGER CONVEX OPTIMIZATION

Miles Lubin
with Emre Yamangil, Russell Bent, Juan Pablo Vielma, Chris Coey

April 1, 2016

MIT & Los Alamos National Laboratory

First, **Mixed-integer linear programming (MILP)**,

$$\begin{aligned} \min_{x} \quad & c^T x \\ \text{subject to} \quad & Ax = b, \\ & x \geq 0, \\ & x_i \in \mathbb{Z}, \quad \forall i \in I \end{aligned}$$

- Despite NP-Hardness, many problems of practical interest can be solved to optimality or near optimality
- Algorithms are based on LP relaxations, branch & bound, cuts, preprocessing, heuristics, ...
- 50+ years of commercial investment in developing these techniques

WHO USES MILP?

http://www.gurobi.com/company/example-customers

Minimize unit-production costs + fixed operating costs

subject to:

- Total generation = total demand, hourly over 24h
- Generation and transmission limits
  - If a generator is on, it produces within some interval $[l, u]$, otherwise its production is zero
- Linear approximation of nonlinear powerflow laws

1% reduction in generation gives billions of dollars per year cost reduction. Worth spending time to improve optimality.

$$\begin{aligned} \min_x \quad & ||Ax - b||_2 \\ \text{subject to} \quad & ||x||_0 \leq k, \end{aligned}$$

Suppose we know optimal solution satisfies $||x||_\infty \leq M$. We have an equivalent **mixed-integer quadratic** formulation

$$\begin{aligned} \min_{x,y} \quad & ||Ax - b||_2 \\ \text{subject to} \quad & \sum y_i \leq k, \\ & x_i \leq My_i \\ & x_i \geq -My_i \\ & y \in \{0,1\}^n \end{aligned}$$

EXAMPLE: SPARSE LINEAR REGRESSION

$$\begin{aligned} \min_{x} \quad & ||Ax - b||_2 \\ \text{subject to} \quad & ||x||_0 \leq k, \end{aligned}$$

These are solvable up to 1000s of dimensions using modern commercial solvers (Bertsimas and King, 2016).

We take m measurements y_i by performing a set of experiments a_i,

$$y_i = a_i^T x + \epsilon_i$$

with noise ϵ_i independent standard Gaussian.

Standard confidence set for $x \in \mathbb{R}^n$ has form

$$\{z | (z - \hat{x})^T E^{-1} (z - \hat{x}) \leq \beta\}$$

where $\hat{x}$ is the maximum likelihood estimate of x,

$$E = \left(\sum_{i=1}^{m} a_i a_i^T \right)^{-1},$$

and β depends on n confidence level α.

Say we want to choose experiments a_i in an optimal way from a set of potential experiments $v_1, v_2, \dots, v_p$. Introduce an integer variable m_j counting how many times we perform experiment v_j. Then

$$E = \left(\sum_{j=1}^{p} m_j v_j v_j^T \right)^{-1}.$$

We could choose to minimize the volume of the confidence ellipsoid,

$$\begin{aligned} \max_{m} \quad & \det(E^{-1}) \\ \text{subject to} \quad & \sum m_j = k, \\ & m \geq 0, \\ & m \in \mathbb{Z}^n. \end{aligned}$$

EXAMPLE: EXPERIMENTAL DESIGN

$$\begin{aligned} \max_{m} \quad & \det(E^{-1}) \\ \text{subject to} \quad & \sum m_j = k, \\ & m \geq 0, \\ & m \in \mathbb{Z}^n. \end{aligned}$$

becomes

$$\begin{aligned} \max_{m} \quad & \log(\det(E^{-1})) \\ \text{subject to} \quad & \sum m_j = k, \\ & m \geq 0, \\ & m \in \mathbb{Z}^n. \end{aligned}$$

- Convex optimization problem when we relax integrality
- Boyd (Ch. 7.5) proposes to **solve convex relaxation and round**, but maybe we can solve the original problem to optimality!

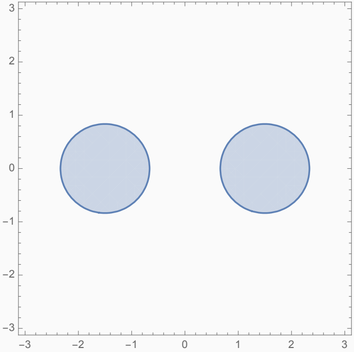
Mixed-integer convex programming (MICP),

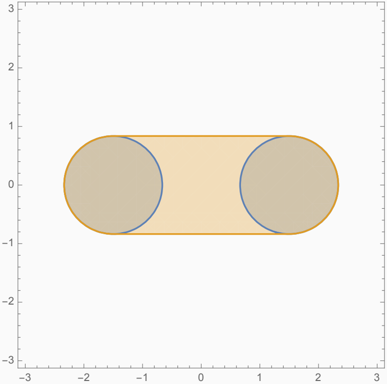
$$\begin{aligned} \min_{x} \quad & f(x) \\ \text{subject to} \quad & g_j(x) \leq 0, \quad \forall j \\ & x_i \in \mathbb{Z}, \quad \forall i \in I \end{aligned}$$

Objective function f and constraints g_j **convex**

Disjunctions/unions of compact convex sets (Stubbs and Mehrotra, 1999)

$$\{x : f_1(x) \leq 0 \text{ or } f_2(x) \leq 0\}$$





x belongs to the convex hull iff $\exists\ u_1, u_2, \lambda_1, \lambda_2$ such that

$$x = \lambda_1 u_1 + \lambda_2 u_2$$

$$\lambda_1 + \lambda_2 = 1$$

$$\lambda_1 \geq 0, \lambda_2 \geq 0$$

$$f_1(u_1) \leq 0$$

$$f_2(u_2) \leq 0$$

$$x = \lambda_1 u_1 + \lambda_2 u_2$$
$$\lambda_1 + \lambda_2 = 1$$
$$\lambda_1 \geq 0, \lambda_2 \geq 0$$
$$f_1(u_1) \leq 0$$
$$f_2(u_2) \leq 0$$

- If we impose $\lambda_1, \lambda_2 \in \{0, 1\}$ then have conditions for the union, but not convex

$$
\begin{gathered}
x = \lambda_1 u_1 + \lambda_2 u_2 \\
\lambda_1 + \lambda_2 = 1 \\
\lambda_1 \geq 0, \lambda_2 \geq 0 \\
f_1(u_1) \leq 0 \\
f_2(u_2) \leq 0
\end{gathered}
$$

- If we impose $\lambda_1, \lambda_2 \in \{0, 1\}$ then have conditions for the union, but not convex
- Instead, introduce $z_1 = \lambda_1 u_1$, $z_2 = \lambda_2 u_2$

$$x = z_1 + z_2$$
$$\lambda_1 + \lambda_2 = 1$$
$$\lambda_1 \geq 0, \lambda_2 \geq 0$$
$$f_1(z_1/\lambda_1) \leq 0$$
$$f_2(z_2/\lambda_2) \leq 0$$

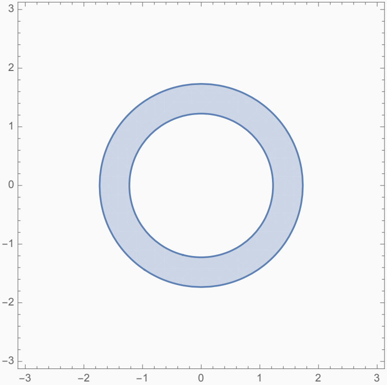
- Still not convex

$$
\begin{gathered}
x = z_1 + z_2 \\
\lambda_1 + \lambda_2 = 1 \\
\lambda_1 \geq 0, \lambda_2 \geq 0 \\
\lambda_1 f_1(z_1/\lambda_1) \leq 0 \\
\lambda_2 f_2(z_2/\lambda_2) \leq 0 \\
\lambda_1, \lambda_2 \in \{0, 1\}
\end{gathered}
$$

- *Perspective* function of any convex function is convex.

- Which nonconvex sets can be modeled with MICP? **Finite unions of compact, convex sets**

Conjecture: the following set is not MICP representable



State of the art for solving MICP

$$\begin{aligned} \min_{x} \quad & f(x) \\ \text{subject to} \quad & g_j(x) \leq 0, \forall j \\ & x_i \in \mathbb{Z}, \forall i \in I \end{aligned}$$

- f, g_j convex, **smooth**
- Bonmin, KNITRO, α-ECP, DICOPT, FilMINT, MINLP_BB, SBB, ...

Idea: Solve an alternating sequence of MILP and convex problems using existing, good solvers.

Given points $x_1^*, \ldots, x_R^*$, solve mixed-integer linear relaxation:

$$\begin{aligned} \min_x \quad & t \\ \text{subject to} \quad & g_j(x_r^*) + \nabla g_j(x_r^*)^T(x - x_r^*) \leq 0, \quad \forall j, r = 1, \ldots, R \\ & f(x_r^*) + \nabla f(x_r^*)^T(x - x_r^*) \leq t, \quad r = 1, \ldots, R \\ & x_i \in \mathbb{Z}, \quad \forall i \in I \end{aligned}$$

Then solve convex problem with integer values fixed to get a new feasible solution, and repeat with $R = R + 1$.

> *[...] solution algorithms for [MICP] have benefit from the technological progress made in solving MILP and NLP. However, in the realm of [MICP], the progress has been far more modest, and the dimension of solvable [MICP] by current solvers is small when compared to MILPs and NLPs.*

– Bonami, Kilinç, and Linderoth (2012)

Classical approaches form **polyhedral outer approximations** by a finite collection of "gradient linearizations". But a good polyhedral outer approximation might need **too many linear constraints**.

Recall:

$$\mathcal{B}_1 := \{x : ||x||_1 \le 1\} = \left\{ x : \sum_{i=1}^{n} s_i x_i \le 1, s \in \{-1, +1\}^n \right\}$$

$$\mathcal{B}_1 := \{x : ||x||_1 \leq 1\} = \left\{x : \sum_{i=1}^{n} s_i x_i \leq 1, s \in \{-1, +1\}^n\right\}$$

Standard trick in linear programming is to **introduce extra variables:**

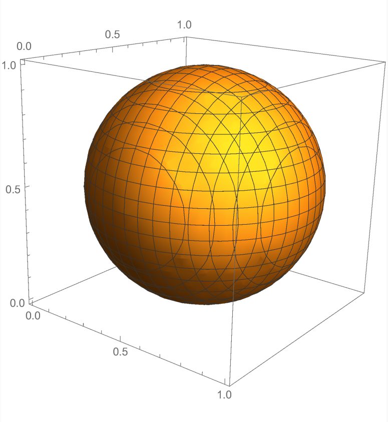
$$\mathcal{B}_1 = \left\{x \in \mathbb{R}^n : \exists y \in \mathbb{R}^n \text{ s.t. } x \leq y, x \geq -y, \sum_i y_i \leq 1\right\}$$

The new formulation has $2n$ variables and $2n + 1$ constraints versus n variables and 2^n constraints. Not bad.

We call this an **extended formulation**.

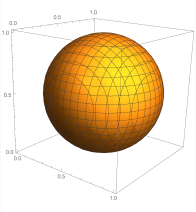
Hijazi et al. (2014):

$$B^n := \left\{ x \in \{0,1\}^n : \sum_{j=1}^{n} \left(x_j - \frac{1}{2} \right)^2 \leq \frac{n-1}{4} \right\}$$



Hijazi et al. (2014):

$$B^n := \left\{ x \in \{0,1\}^n : \sum_{j=1}^{n} \left(x_j - \frac{1}{2} \right)^2 \leq \frac{n-1}{4} \right\}$$



- B^n is empty
- No outer approximating hyperplane can exclude more than one integer lattice point
- So any polyhedral outer approximation which contains no integer vertices needs **at least** 2^n **hyperplanes**

$$B^n := \left\{ x \in \{0,1\}^n : \sum_{j=1}^{n} \left(x_j - \frac{1}{2} \right)^2 \le \frac{n-1}{4} \right\}$$

Consider the equivalent formulation in an extended set of variables:

$$\hat{B}^n := \left\{ x \in \{0,1\}^n, z \in \mathbb{R}^n : \sum_{j=1}^{n} z_j \le \frac{n-1}{4}, \left(x_j - \frac{1}{2} \right)^2 \le z_j \, \forall j \right\}$$

- $B^n = \text{proj}_x \, \hat{B}^n$

$$B^n := \left\{ x \in \{0,1\}^n : \sum_{j=1}^{n} \left(x_j - \frac{1}{2} \right)^2 \leq \frac{n-1}{4} \right\}$$

Consider the equivalent formulation in an extended set of variables:

$$\hat{B}^n := \left\{ x \in \{0,1\}^n, z \in \mathbb{R}^n : \sum_{j=1}^{n} z_j \leq \frac{n-1}{4}, \left(x_j - \frac{1}{2} \right)^2 \leq z_j \, \forall j \right\}$$

- $B^n = \operatorname{proj}_x \hat{B}^n$
- If you form a polyhedral relaxation of each $\left(x_j - \frac{1}{2}\right)^2 \leq z_j$ individually, then $2n$ hyperplanes in $\mathbb{R}^{2n}$ is sufficient to exclude all integer points

HOW TO FIX IT?

$$B^n := \left\{ x \in \{0,1\}^n : \sum_{j=1}^{n} \left(x_j - \frac{1}{2} \right)^2 \leq \frac{n-1}{4} \right\}$$

Consider the equivalent formulation in an extended set of variables:

$$\hat{B}^n := \left\{ x \in \{0,1\}^n, z \in \mathbb{R}^n : \sum_{j=1}^{n} z_j \leq \frac{n-1}{4}, \left(x_j - \frac{1}{2} \right)^2 \leq z_j \; \forall j \right\}$$

- $B^n = \operatorname{proj}_x \hat{B}^n$
- If you form a polyhedral relaxation of each $\left(x_j - \frac{1}{2}\right)^2 \leq z_j$ individually, then $2n$ hyperplanes in $\mathbb{R}^{2n}$ is sufficient to exclude all integer points
- Outer approximation algorithm now converges in 2 iterations

HOW TO FIX IT?

$$B^n := \left\{ x \in \{0,1\}^n : \sum_{j=1}^{n} \left(x_j - \frac{1}{2} \right)^2 \leq \frac{n-1}{4} \right\}$$

Consider the equivalent formulation in an extended set of variables:

$$\hat{B}^n := \left\{ x \in \{0,1\}^n, z \in \mathbb{R}^n : \sum_{j=1}^{n} z_j \leq \frac{n-1}{4}, \left(x_j - \frac{1}{2} \right)^2 \leq z_j \, \forall j \right\}$$

- $B^n = \operatorname{proj}_x \hat{B}^n$
- If you form a polyhedral relaxation of each $\left(x_j - \frac{1}{2}\right)^2 \leq z_j$ individually, then $2n$ hyperplanes in $\mathbb{R}^{2n}$ is sufficient to exclude all integer points
- Outer approximation algorithm now converges in 2 iterations
- Why? **Small polyhedron in higher dimension can have exponentially many facets in lower dimension**

“EXTENDED FORMULATION FOR SECOND-ORDER CONE”

$$SOC_n := \{(t, x) \in \mathbb{R}^n : ||x||_2 \leq t\}$$

Second order cone programming (SOCP) generalizes convex quadratic programming (QP).

"EXTENDED FORMULATION FOR SECOND-ORDER CONE"

$$SOC_n := \{(t, x) \in \mathbb{R}^n : ||x||_2 \leq t\}$$

Second order cone programming (SOCP) generalizes convex quadratic programming (QP).

Vielma, Dunning, Huchette, Lubin (2015):

- For MISOCP: Rewrite

  $$\sum_i x_i^2 \leq t^2$$

  as

  $$\sum_i z_i \leq t, \text{ where } z_i \geq x_i^2/t$$

- Implemented by CPLEX and Gurobi within months of publication

Hijazi et al. (2014) propose to reformulate constraints of the form

$$\sum_{i=1}^{n} g_i(x_i) \leq 0$$

to

$$\sum_{i=1}^{n} z_i \leq 0 \text{ and } g_i(x_i) \leq z_i \; i = 1, \ldots, n,$$

where g_i univariate, smooth, convex.

- Very impressive computational results
- Paper submitted in 2011, but there's still no solver which automates this transformation

Why has nobody implemented this?

Why has nobody implemented this?

$$\begin{aligned} \min_{x} \quad & f(x) \\ \text{subject to} \quad & g_j(x) \leq 0, \quad \forall j \\ & x_i \in \mathbb{Z}, \quad \forall i \in I \end{aligned}$$

- MICP solvers view f and g_j through “black-box” oracles for evaluations of values and derivatives
- Summation structure not accessible through the “black box”

Mixed-integer disciplined convex programming (MIDCP),

$$\begin{aligned} \min_{x} \quad & f(x) \\ \text{subject to} \quad & g_j(x) \leq 0, \quad \forall j \\ & x_i \in \mathbb{Z}, \quad \forall i \in I \end{aligned}$$

Objective function f and constraints g_j **disciplined convex**

Forget about derivatives

“Models with integer and binary variables must still obey all of the same disciplined convex programming rules that CVX enforces for continuous models. For this reason, we are calling these models ***mixed-integer disciplined convex programs*** *or* ***MIDCPs****.“*

What is a disciplined convex function? (Grant, Boyd, Ye, 2006)

- Expressed as composition of basic atoms. Rules of composition prove convexity.
- Example $\sqrt{x^2 + y^2}$ versus $||[x, y]||_2$, $\sqrt{xy}$ versus geomean(x, y).

Why disciplined convex?

- Functions expressed in this form can be automatically converted to conic optimization problems. We have good solvers and theory for conic optimization.

$$\begin{aligned} \min_x \quad & c^T x \\ \text{s.t.} \quad & Ax = b \\ & x \in \mathcal{K} \end{aligned}$$

- $\mathcal{K}$ is a cone if $x \in \mathcal{K}$ implies $\alpha x \in \mathcal{K} \, \forall \alpha \geq 0$
- Ususally we consider $\mathcal{K} = \mathcal{K}_1 \times \cdots \times \mathcal{K}_s$ where each $\mathcal{K}_j$ is one of a small number of cones like $\mathbb{R}^n_+$, SOC_n, the cone of positive semidefinite matrices...

DCP

$$\min_x \quad f(x) \qquad \text{s.t.} \quad g_j(x) \le 0, \quad \forall j$$

CVX
CVXPY
Convex.jl

Conic

$$\min_x \quad c^T x \qquad \text{s.t.} \quad Ax = b, \quad x \in \mathcal{K}$$

ECOS
SCS
Mosek

Optimal solution x^*

Example:

$$\sum_i \exp(c_i^T x + b_i) \le t$$

is expressed in conic form as

$$\sum_i z_i \le t \text{ and } (c_i^T x + b_i, 1, z_i) \in EXP.$$

where

$$EXP = \mathrm{cl}\{(x, y, z) \in \mathbb{R}^3 : y \exp(x/y) \le z, y > 0\}$$

is the **exponential cone**.

Example:

$$\sum_i \exp(c_i^T x + b_i) \leq t$$

is expressed in conic form as

$$\sum_i z_i \leq t \text{ and } (c_i^T x + b_i, 1, z_i) \in EXP.$$

where

$$EXP = \mathrm{cl}\{(x, y, z) \in \mathbb{R}^3 : y \exp(x/y) \leq z, y > 0\}$$

is the **exponential cone**.

The translation to conic form already produces a formulation with additional variables ideal for polyhedral outer approximation.

Example:

$$\exp(x^2 + y^2) \leq t$$

is DCP compliant. CVX will generate equivalent formulation

$$\exp(z) \leq t \text{ where } x^2 + y^2 \leq z$$

which is then translated to EXP and second-order cones.

- Tawarmalani and Sahinidis (2005) show that polyhedral outer approximations that exploit composition structure have extra strength. This comes for “free” with conic representations

- Mixed-integer second order cone programming (MISOCP) has rapidly improving commercial support by Gurobi/CPLEX
    - Gurobi claims 2.8x improvement from version 6.0 to 6.5 (1 year)
- Otherwise (exponential cones, SDP), mainly research codes (`ecos_bb` by Han Wang, 2014; `SCIP-SDP` by Mars, 2012)
- For MISOCP, no established open-source codes

How general is MISOCP?

MINLPLIB2 benchmark library designed for derivative-based mixed-integer nonlinear solvers. We classified all of the convex instances by conic representability.

Of the 333 convex instances, **65%** are MISOCP representable.

- `tls6` instance previously unsolved. We translated it to Convex.jl and it was solved by Gurobi!

What about the remaining 35% of problems?

Of remaining 116 instances, in 107 the only nonlinear expressions involve exp and log.

What about the remaining 35% of problems?

Of remaining 116 instances, in 107 the only nonlinear expressions involve exp and log.

Recall the exponential cone:

$$EXP = \mathrm{cl}\{(x, y, z) \in \mathbb{R}^3 : y \exp(x/y) \leq z, y > 0\}$$

- Closed, convex, nonsymmetric cone. Theoretically tractable and supported by ECOS (Akle, 2015)

What about the remaining 9 of 333 instances?

- Seven representable by combination of *EXP* and second-order cones (*SOC*)
- Two representable by the power cone

$$POW_\alpha = \{(x, y, z) \in \mathbb{R}^3 : |z| \leq x^a y^{1-a}, x \geq 0, y \geq 0\}$$

Folklore: **Almost all convex optimization problems of practical interest can be represented as conic programming problems using second-order, positive semidefinite, exponential, and power cones.**

So cones are general and provide a good extended formulation of the original convex problem.

How do we apply polyhedral outer approximation to mixed-integer conic problems?

- We'll state the first finite-time outer approximation algorithm for general mixed-integer conic optimization

MIDCP

$$
\begin{aligned}
\min_x \quad & f(x) \\
\text{s.t.} \quad & g_j(x) \leq 0, \quad \forall j \\
& x_i \in \mathbb{Z}, \quad \forall i \in I
\end{aligned}
$$

CVX
CVXPY
Convex.jl

MICONE

$$
\begin{aligned}
\min_x \quad & c^T x \\
\text{s.t.} \quad & Ax = b \\
& x \in \mathcal{K} \\
& x_i \in \mathbb{Z}, \quad \forall i \in I
\end{aligned}
$$

???

Optimal solution x^*

$$
\begin{aligned}
\min_{x,z} \quad & c^T z \\
\text{s.t.} \quad & A_x x + A_z z = b \\
& L \leq x \leq U, \quad x \in \mathbb{Z}^n, \quad z \in \mathcal{K},
\end{aligned}
\tag{MICONE}
$$

- $\mathcal{K} = \mathcal{K}_1 \times \mathcal{K}_2 \times \cdots \times \mathcal{K}_r$, where each $\mathcal{K}_i$ is a standard cone: nonnegative orthant, second-order cone, exponential cone, power cone, or even positive semidefinite cone
- WLOG, integer variables do not belong to cones, have zero objective coefficients

Need to outer approximate cone $\mathcal{K}$ with finite number of linear constraints...

Need to outer approximate cone $\mathcal{K}$ with finite number of linear constraints...

Given a closed, convex cone $\mathcal{K}$, the **dual cone** $\mathcal{K}^*$ is the set such that $z \in \mathcal{K}$ iff $z^T\beta \geq 0 \ \forall \beta \in \mathcal{K}^*$.

- Nonnegative orthant, second-order and positive semidefinite cone are *self dual*. Exponential and power cones are not
- Any finite subset of $\mathcal{K}^*$ yields a polyhedral outer approximation of $\mathcal{K}$!

$$\begin{aligned} \min_{x,z} \quad & c^T z \\ \text{s.t.} \quad & A_x x + A_z z = b \\ & L \leq x \leq U, \quad x \in \mathbb{Z}^n, \\ & \beta^T z \geq 0 \quad \forall \beta \in T, \end{aligned} \tag{MIOA(T)}$$

- If $T = \mathcal{K}^*$, then equivalent to MICONE
- If $T \subseteq \mathcal{K}^*$ and $|T| < \infty$, then *polyhedral outer approximation* $\rightarrow$ MILP relaxation of MICONE
- **Claim:** $\exists T \subseteq \mathcal{K}^*, |T| < \infty$ such that MIOA(T) is equivalent to *MICONE*

Consider the subproblem with integer values $x = \hat{x}$ fixed:

$$
\begin{aligned}
v_{\hat{x}} = \min_{z} \quad & c^T z \\
\text{s.t.} \quad & A_z z = b - A_x \hat{x}, \\
& z \in \mathcal{K}.
\end{aligned}
\tag{$CP(\hat{x})$}
$$

The dual of $CP(\hat{x})$ is

$$
\begin{aligned}
\max_{\beta,\lambda} \quad & \lambda^T (b - A_x \hat{x}) \\
\text{s.t.} \quad & \beta = c - A_z^T \lambda \\
& \beta \in \mathcal{K}^*.
\end{aligned}
$$

- Start with $T = \emptyset$ (or small)
- Until the optimal objective of MIOA(T) is $\geq$ the optimal objective of the best feasible solution:
  - Let $\hat{x}$ be the optimal integer solution of the MILP problem MIOA(T)
  - Solve conic problem $CP(\hat{x})$, add dual solution $\beta_{\hat{x}}$ to T
  - If $CP(\hat{x})$ was feasible, record new feasible solution to MICONE

- Start with $T = \emptyset$ (or small)
- Until the optimal objective of MIOA(T) is $\geq$ the optimal objective of the best feasible solution:
    - Let $\hat{x}$ be the optimal integer solution of the MILP problem MIOA(T)
    - Solve conic problem $CP(\hat{x})$, add dual solution $\beta_{\hat{x}}$ to T
    - If $CP(\hat{x})$ was feasible, record new feasible solution to MICONE

Finite termination assuming strong duality holds at subproblems

- Easy to implement, requires black box MILP and conic solvers
- Very often, number of iterations is < 10, sometimes 10-30. Worst observed is 171.
- Each MILP subproblem could be NP-Hard, but takes advantage of fast solvers

$$
\begin{aligned}
v_{\hat{x}} = \min_{z} \quad & c^T z \\
\text{s.t.} \quad & A_z z = b - A_x \hat{x}, \\
& z \in \mathcal{K}.
\end{aligned}
\qquad\qquad
\begin{aligned}
\max_{\beta,\lambda} \quad & \lambda^T (b - A_x \hat{x}) \\
\text{s.t.} \quad & \beta = c - A_z^T \lambda \\
& \beta \in \mathcal{K}^*.
\end{aligned}
$$

Suppose $CP(\hat{x})$ is feasible and strong duality holds at the optimal primal-dual solution $(z_{\hat{x}}, \beta_{\hat{x}}, \lambda_{\hat{x}})$. Then for any z with $A_z z = b - A_x \hat{x}$ and $\beta_{\hat{x}}^T z \geq 0$, we have $c^T z \geq v_{\hat{x}}$.

$$v_{\hat{x}} = \min_{z} \quad c^T z \qquad\qquad \max_{\beta,\lambda} \quad \lambda^T(b - A_x\hat{x})$$

$$\text{s.t.} \quad A_z z = b - A_x\hat{x}, \qquad\qquad \text{s.t.} \quad \beta = c - A_z^T\lambda$$

$$z \in \mathcal{K}. \qquad\qquad \beta \in \mathcal{K}^*.$$

Suppose $CP(\hat{x})$ is feasible and strong duality holds at the optimal primal-dual solution $(z_{\hat{x}}, \beta_{\hat{x}}, \lambda_{\hat{x}})$. Then for any z with $A_z z = b - A_x\hat{x}$ and $\beta_{\hat{x}}^T z \geq 0$, we have $c^T z \geq v_{\hat{x}}$.

Including $\beta_{\hat{x}}^T z \geq 0$ in polyhedral relaxation implies that if $x = \hat{x}$ is the optimal (integer part of the) solution of the relaxation, then the objective value of the relaxation is at least $v_{\hat{x}}$.

Proof.

$$\beta_{\hat{x}}^T z = (c - A_z^T\lambda_{\hat{x}})^T z = c^T z - \lambda_{\hat{x}}^T(b - A_x\hat{x}) = c^T z - v_{\hat{x}} \geq 0.$$

□

$$
\begin{aligned}
v_{\hat{x}} = \min_{z} \quad & c^T z \\
\text{s.t.} \quad & A_z z = b - A_x \hat{x}, \\
& z \in \mathcal{K}.
\end{aligned}
\qquad\qquad
\begin{aligned}
\max_{\beta,\lambda} \quad & \lambda^T (b - A_x \hat{x}) \\
\text{s.t.} \quad & \beta = c - A_z^T \lambda \\
& \beta \in \mathcal{K}^*.
\end{aligned}
$$

Given $\hat{x}$, assume $CP(\hat{x})$ is infeasible and its dual is unbounded, such that we have a ray $(\beta_{\hat{x}}, \lambda_{\hat{x}})$ satisfying $\beta_{\hat{x}} \in \mathcal{K}^*$, $\beta = -A_z^T \lambda_{\hat{x}}$, and $\lambda_{\hat{x}}^T (b - A_x \hat{x}) > 0$. Then for any z satisfying $A_z z = b - A_x \hat{x}$ we have $\beta_{\hat{x}}^T z < 0$.

Including $\beta_{\hat{x}}^T z \geq 0$ in polyhedral relaxation implies $x = \hat{x}$ is infeasible to the relaxation.

$$
\begin{aligned}
v_{\hat{x}} = \min_{z} \quad & c^T z \\
\text{s.t.} \quad & A_z z = b - A_x \hat{x}, \\
& z \in \mathcal{K}.
\end{aligned}
\qquad\qquad
\begin{aligned}
\max_{\beta, \lambda} \quad & \lambda^T (b - A_x \hat{x}) \\
\text{s.t.} \quad & \beta = c - A_z^T \lambda \\
& \beta \in \mathcal{K}^*.
\end{aligned}
$$

Given $\hat{x}$, assume $CP(\hat{x})$ is infeasible and its dual is unbounded, such that we have a ray $(\beta_{\hat{x}}, \lambda_{\hat{x}})$ satisfying $\beta_{\hat{x}} \in \mathcal{K}^*$, $\beta = -A_z^T \lambda_{\hat{x}}$, and $\lambda_{\hat{x}}^T (b - A_x \hat{x}) > 0$. Then for any z satisfying $A_z z = b - A_x \hat{x}$ we have $\beta_{\hat{x}}^T z < 0$.

Including $\beta_{\hat{x}}^T z \geq 0$ in polyhedral relaxation implies $x = \hat{x}$ is infeasible to the relaxation.

Proof.

$$
\beta_{\hat{x}}^T z = -\lambda_{\hat{x}}^T A_z z = -\lambda_{\hat{x}}^T (b - A_x \hat{x}) < 0.
$$

$\square$

Claim: $\exists T \subseteq \mathcal{K}^*, |T| < \infty$ such that MIOA(T) is equivalent to *MICONE*

Claim: $\exists T \subseteq \mathcal{K}^*, |T| < \infty$ such that MIOA(T) is equivalent to *MICONE*

Proof.

Solve $CP(\hat{x})$ for all possible $\hat{x}$ (finite). Then include all corresponding dual vectors $\beta_{\hat{x}}$ in T. If x^* is the integer solution of MIOA(T), then $CP(x^*)$ is feasible from the second lemma. From first lemma, objective of MIOA(T) is at least the optimal value of $CP(x^*)$, but it's also a relaxation so x^* must be optimal. $\square$

- Polyhedral approximations can be made *much* better by introducing extra variables

- Polyhedral approximations can be made *much* better by introducing extra variables
- Summation structure and compositions of convex functions should be broken up

- Polyhedral approximations can be made *much* better by introducing extra variables
- Summation structure and compositions of convex functions should be broken up
- Nobody automated this before! (Derivative-based view prevents implementation)

- Polyhedral approximations can be made *much* better by introducing extra variables
- Summation structure and compositions of convex functions should be broken up
- Nobody automated this before! (Derivative-based view prevents implementation)
- DCP to the rescue. Conic form is general and encodes the information we need as a solver

- Polyhedral approximations can be made *much* better by introducing extra variables
- Summation structure and compositions of convex functions should be broken up
- Nobody automated this before! (Derivative-based view prevents implementation)
- DCP to the rescue. Conic form is general and encodes the information we need as a solver
- Proposed first outer approximation algorithm for MICONE, based on conic duality

Computational results

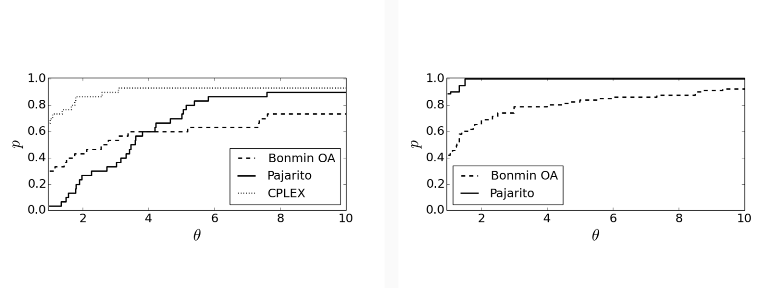
Pajarito

New solver!

- 1000 lines of Julia
- Input is in conic form, accessible through Convex.jl DCP language
- Works with any MILP and conic solvers supported in Julia
- For tough conic problems, also supports traditional nonlinear solvers
- Released last week!

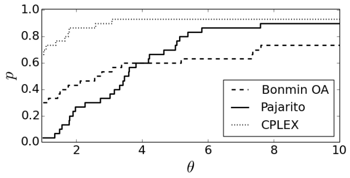
Translated 194 convex instances from MINLPLIB2 from AMPL into Convex.jl (heroic parsing work by E. Yamangil).

Compare with Bonmin's outer approximation algorithm using CPLEX as the MILP solver. We use the same version of CPLEX plus KNITRO as NLP (conic) solver.

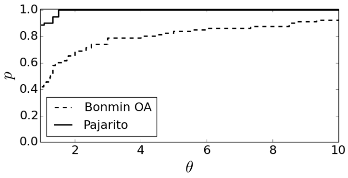


(a) Solution time (b) Number of OA iterations

Figure: Comparison performance profiles for SOC representable instances (Bonmin >30 sec)



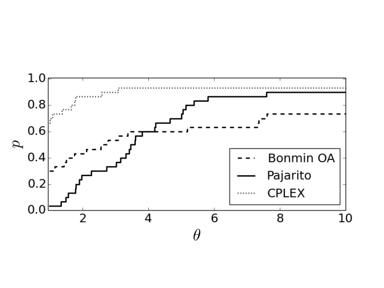
(a) Solution time



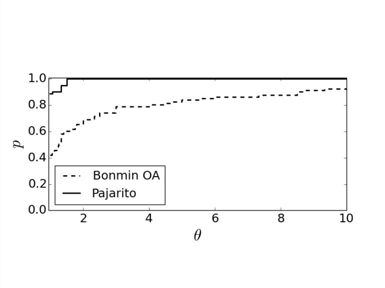
(b) Number of OA iterations

Figure: Comparison performance profiles for SOC representable instances (Bonmin >30 sec)

- Dominates on iteration count. Haven't optimized Pajarito so Bonmin is faster on the easier problems.



(a) Solution time



(b) Number of OA iterations

Figure: Comparison performance profiles for SOC representable instances (Bonmin >30 sec)

- Dominates on iteration count. Haven't optimized Pajarito so Bonmin is faster on the easier problems.
- CPLEX is usually the fastest on MISOCPs. Use that instead.

Almost none of the non-MISOCP instances in MINLPLIB2 are hard!

`gams01` unsolved instance which is EXP+SOC representable

- 145 variables. 1268 constraints (1158 are linear)
- Previous best bound: 1735.06. Best known solution: 21516.83
- Pajarito solved in 6 iterations ($<$ 10 hours). Optimal solution is 21380.20.

- Extensions to geometric programming and semidefinite programming
- What happens when strong duality fails?
- Lots of room for improvement in reliability of conic solvers (especially exponential cone)

Pajarito Mountain, New Mexico

Thanks! Questions?

https://github.com/mlubin/Pajarito.jl

http://arxiv.org/abs/1511.06710

DOE CSGF