Extended and Embedding Formulations for MINLP

Juan Pablo Vielma
Massachusetts Institute of Technology


Supported by NSF grant CMMI-1351619
Nonlinear Mixed 0-1 Integer Formulations

- Modeling Finite Alternatives = Unions of Convex Sets

\[ x \in \bigcup_{i=1}^{n} C_i \subseteq \mathbb{R}^d \]
Extended and Non-Extended Formulations for $\bigcup_{i=1}^{n} C_i$

$C_i = \{ x \in \mathbb{R}^d : f_i(x) \leq 0 \}$

Extended

- $f_i(x^i, y_i) \leq 0 \quad \forall i \in [n]$  
- $\sum_{i=1}^{n} x^i = x$  
- $\sum_{i=1}^{n} y_i = 1$  
- $y \in \{0, 1\}^n$  
- $x, x^i \in \mathbb{R}^d \quad \forall i \in [n]$  

Strong, but large

Non-Extended

- $f_i(x) \leq M_i (1 - y_i) \quad \forall i \in [n]$  
- $\sum_{i=1}^{n} y_i = 1$  
- $y \in \{0, 1\}^n$  
- $x \in \mathbb{R}^d \quad \forall i \in [n]$  

Small, but weak?
Outline

• Extended Formulations
  – Conic formulations, stability and outer-approximation

• Non-Extended Formulations
  – Embedding formulations = Strong non-extended
Extended Formulations:

Birdmen: Or (The Unexpected Virtue of Discipline)
Extended Formulations: Perspective “v/s” Cones

* e.g. Ceria and Soares ‘99

\[ C_i = \{ x \in \mathbb{R}^d : f_i(x) \leq 0 \} \]

\[ \tilde{f}(x, y) = \begin{cases} 
  y f(x/y) & \text{if } y > 0 \\
  \lim_{\alpha \downarrow 0} \alpha f(x' - x + x/\alpha) & \text{if } y = 0 \\
  +\infty & \text{if } y < 0
\end{cases} \]

\[ \tilde{f}_i(x^i, y_i) \leq 0 \quad \forall i \in [n] \]

\[ \sum_{i=1}^{n} x^i = x \]

\[ \sum_{i=1}^{n} y_i = 1 \]

\[ y \in \{0, 1\}^n \]

\[ x, x^i \in \mathbb{R}^d \quad \forall i \in [n] \]

* e.g. Ben-tal and Nemirovski ’01, Helton and Nie ‘09

\[ C_i = \left\{ x \in \mathbb{R}^d : \exists u \in \mathbb{R}^{p_i} \text{ s.t. } A^i x + D^i u - b \in K^i \right\} \]

\[ K^i \text{ closed convex cone} \]

\[ A^i x^i + D^i u^i - b y_i \in K^i \quad \forall i \in [n] \]

\[ \sum_{i=1}^{n} x^i = x \]

\[ \sum_{i=1}^{n} y_i = 1 \]

\[ y \in \{0, 1\}^n \]

\[ x, x^i \in \mathbb{R}^d \quad \forall i \in [n] \]

\[ u^i \in \mathbb{R}^{p_i} \quad \forall i \in [n] \]

Both formulations are ideal (extreme points of continuous relaxation satisfy integrality constraints)
Cones Can Mitigate Unintended Numerical Issues

• Let \( C_i = \{ x \in \mathbb{R}^2 : f_i(x) \leq 0 \} \)
where \( f_i(x) = x_1^2 - x_2 - 1 \)

\[
\tilde{f}_i(x, y) = \begin{cases} 
  y(x_1/y)^2 - x_2 - y & \text{if } y > 0 \\
  -x_2 & \text{if } y = x_1 = 0 \\
  +\infty & \text{if o.w.}
\end{cases}
\]

• Conic (SOCP) representation

\[
C_i = \left\{ x \in \mathbb{R}^2 : \sqrt{x_2^2 + 4x_1^2} \leq 2 + x_2 \right\}
\]

\[
\sqrt{(x_2^i)^2 + 4(x_1^i)^2} \leq 2y_i + x_2
\]
Conic = Really Extended

• Conic representation = additional auxiliary variables

\[ C_i = \left\{ x \in \mathbb{R}^d : \exists u \in \mathbb{R}^{p_i} \text{ s.t. } A^i x + D^i u - b \in K^i \right\} \]

• Bad for NLP solvers, but good for polyhedral approximations in MINLP solvers:
  – Separable = (Tawarmalani and Sahinidis ’05, Hijazi et al. ’12)
  – SOCP = (V. et al. ‘15) = If SOCP representable use MI-SOCP solver!
  – General MINLP? = Disciplined Convex Programming (DCP) (Grant et al. 06): Implemented in CVX
    • Systematic way to prove convexity
    • Yields extended conic representation (MINLPLib2 -1)
Solving Mixed Integer Disciplined Convex Programs

- **Pajarito** solver!
  - Lubin, Yamangil, Bent and V. ’15
  - SLay10M (it / time):
    - Bonmim = 69 it / 1,379 s
    - Hijazi et al. = 23 it / 14 s
    - Pajarito = 5 it / 12 s
      (automatic separability)
  - Solved gams01 from MINLPLIB2 (prev 91% GAP)
  - Solved tls5-6 (prev 25%-29% GAP) = Just SOCP + Gurobi!
  - ~200 lines of Julia code

Miles Lubin and Emre Yamangil
Strong Non-Extended Formulations:
Minkowski Sums, Good or Evil
Constructing Non-extended Ideal Formulations

- Pure Integer:
  \[ Q := \text{conv} \left( \{ p_i \}_{i=1}^n \right) \]

- Mixed Integer:
  \[ Q \cap \mathbb{Z}^2 \]

![Diagram](image_url)
Embedding Formulation = Ideal non-Extended

\[ Q(H) := \text{conv} \left( \bigcup_{i=1}^{n} P_i \times \{h^i\} \right) \]

\[ (x, y) \in Q \cap (\mathbb{R}^d \times \mathbb{Z}^k) \iff y = h^i \land x \in P_i \]

\[ \text{ext}(Q) \subseteq \mathbb{R}^d \times \mathbb{Z}^k \]

\[ H := \{h^i\}_{i=1}^{n} \subseteq \{0, 1\}^k, \quad h^i \neq h^j \]
Alternative Encodings

- “Only” use 0-1 encodings

- Options for 0-1 encodings:
  - Traditional or **Unary** encoding
    \[ H = \left\{ y \in \{0, 1\}^n : \sum_{i=1}^{n} y_i = 1 \right\} \]
    \[ = \left\{ e^i \right\}_{i=1}^{n} \]
    - Binary encodings: \( H \equiv \{0, 1\}^{\log_2 n} \)
    - Others (e.g. **incremental** encoding \( \equiv \) unary)
Unary Encoding, Minkowski Sum and Cayley Trick

For traditional or unary encoding:

\[ Q \cap (\mathbb{R}^2 \times \{0.5\}) \equiv P_1 + P_2 = \]

\[ Q(H) \cap (\mathbb{R}^d \times \{\frac{1}{n} \sum_{i=1}^{n} e^i\}) \equiv \sum_{i=1}^{n} P_i \]
Encoding Selection Matters: Evil Minkowski Sum

• Size of unary formulation is:
  (Lee and Wilson ’01)

\[
\frac{2\sqrt{n/2}}{\sqrt{n/2}} + \left(\sqrt{n/2} + 1\right)^2
\]

• Size of one binary formulation:
  (V. and Nemhauser ’08)

\[
4\log_2\sqrt{n/2} + 2 + \left(\sqrt{n/2} + 1\right)^2
\]

• Right embedding = significant computational advantage over alternatives (Extended, Big-M, etc.)
Complexity of Family of Polyhedra $\mathcal{P} := \{P_i\}_{i=1}^n$

- Embedding complexity = smallest **ideal** formulation

$$mc(\mathcal{P}) := \min_H \{\text{size}(Q(H))\}$$

- Relaxation complexity = smallest formulation

$$rc(\mathcal{P}) := \min_{Q,H} \{\text{size}(Q)\}$$
Complexity Results

• Lower and Upper bounds for special structures:
  – e.g. for Special Order Sets of Type 2 (SOS2) on \( n \) variables
    • Embedding complexity (ideal)
      \[
      2\lceil\log_2 n\rceil \quad \text{General Inequalities}
      \]
      \[
      n + 1 \leq \ldots \leq n + 1 + 2\lceil\log_2 n\rceil \quad \text{Total}
      \]
    • Relaxation complexity (non-ideal)
      \[
      2 \leq \ldots \leq 4 \quad \text{General Inequalities}
      \]
      \[
      2 \leq \ldots \leq 5 + 2n \quad \text{Total}
      \]

• Relation to other complexity measures
  \[
  \text{hc} \left( \mathcal{P} \right) := \text{size} \left( \text{conv} \left( \bigcup_{i=1}^{n} P_i \right) \right)
  \]
  \[
  \text{xc} \left( \mathcal{P} \right) := \min_{R} \left\{ \text{size} \left( R \right) : \text{proj}_x \left( R \right) = \text{conv} \left( \bigcup_{i=1}^{n} P_i \right) \right\}
  \]

• Still open questions (see V. 2015)
Faces for Unary Encoding: Good Minkowski Sum

- Two types of facets (or faces):
  - \( P_1 \times \{0\} \equiv y_i \geq 0 \)
  - \( \text{conv} \left( (F_1 \times 0) \cup (F_2 \times 1) \right) \)
    - \( F_i \) proper face of \( P_i \)
  - Not all combinations of faces
    - Which ones are valid?
      - Minkowski to the rescue!
Valid Combinations = Common Normals

\[ N(F_1) \cap N(F_2) \neq \emptyset \]
\[ \Leftrightarrow \]
\( \text{conv} ((F_1 \times 0) \cup (F_2 \times 1)) \)
\( \text{is face of } Q(H) \)
• Description of boundary of $Q(H)$ is easy if “normals condition” yields convex hull of 1 nonlinear constraint and point(s)
Easy to Recover and Generalize Existing Results

• Isotone function results from Hijazi et al. ‘12 and Bonami et al. ’15 (n=1, 2):
  \[
  C_i = \{ x \in \mathbb{R}^d : l^i \leq x \leq u^i, \quad f_i(x) \leq 0 \}
  \]
  
• Can generalize to \( n \geq 3 \) and two functions per set:

• Other special cases (previous slide)
Also “Non-isotone” Results: Pizza Slices

\[ \sqrt{\frac{x_4^2}{4} + \left( \frac{x_2}{3} - \frac{1}{3} y_1 + \frac{1}{3} y_3 \right)^2} \leq 2 y_4 + x_1 + \frac{4}{3} y_1 + y_2 + \frac{4}{3} y_3 \]

\[ \text{conv} \left( \bigcup_{i=1}^{4} \left( C_i \times \{ e^i \} \right) \right) = 4 \text{ conic} + 4 \text{ linear inequalities} \]
Bad Example: Representability Issues

Embedding Formulations

\[ Q := \text{conv} ((C_1 \times \{0\}) \cup (C_2 \times \{1\})) \]

can fail to be basic semi-algebraic

Zariski closure of boundary?

Description with finite number of (quadratic) polynomial inequalities?
Final Positive Results

• Unions of Homothetic Convex Bodies

\[ C_i = \lambda_i C + b_i \]

(all extreme points exposed)

\[
\text{conv} \left( \bigcup_{i=1}^{n} (C_i \times \{e^i\}) \right) = \\
\gamma_C \left( x - \sum_{i=1}^{n} y_i b^i \right) \leq \sum_{i=1}^{n} \lambda_i y_i \\
\sum_{i=1}^{n} y_i = 1 \\
y \geq 0 \\
\forall i \in [n]
\]

\[ \gamma_C(x) := \inf \{ \lambda > 0 : x \in \lambda C \} \]

• Generalizes polyhedral results from Balas ‘85, Jeroslow ‘88 and Blair ‘90
Summary

• Extended formulations
  – Really extended formulations through SOCP or DCP
  – Pajarito = MIDCP / MINLP extended polyhedral solver
• Embedding formulations = systematic procedure for ideal non-extended formulations
  – Polyhedral case = Formulations and complexity
  – Non-Polyhedral 1 = Simplified proofs, extensions and new formulations
  – Non-Polyhedral 2 = Representability issues