Mixed-Integer Nonlinear Optimization: Applications, Algorithms, and Computation V

Sven Leyffer

Mathematics & Computer Science Division
Argonne National Laboratory

Graduate School in
Systems, Optimization, Control and Networks
Université catholique de Louvain
February 2013

Collaborators





Pietro Belotti, Ashutosh Mahajan, Christian Kirches,
Jeff Linderoth, and Jim Luedtke

Outline

Outline

Challenges of Nonconvex MINLP

Mixed-Integer Nonlinear Program (MINLP)

$$\underset{x}{\text{minimize}} \ f(x) \quad \text{subject to } c(x) \leq 0, \ x \in X, \ x_i \in \mathbb{Z} \ \forall \ i \in I$$

... now drop assumption that $f(x)$ and $c(x)$ are convex

Challenges of nonconvex MINLP

- Objective function $f(x)$ can have many local minimizers
- Continuous relaxation of constraint set

$$\{x | c(x) \leq 0, \ x \in X\}$$

... can be disjoint, may have no interior

Challenges of Nonconvex MINLP

Definition (Local/Global Minimum)

Consider nonconvex optimization problem

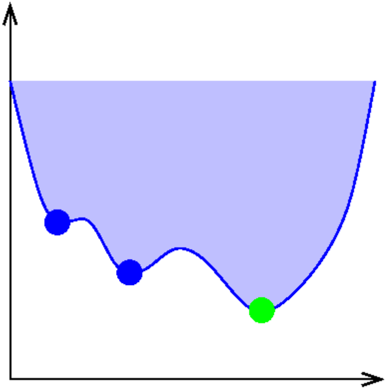
$$\underset{x}{\text{minimize}}\ f(x) \quad \text{subject to } x \in \mathcal{F} := \{x \ : \ c(x) \le 0,\ x \in X\}$$

- x^* is a local minimum iff $\exists\, \mathcal{N}(x^*)$ such that $f(x) \ge f(x^*)$ for all $x \in \mathcal{N}(x^*) \cap \mathcal{F}$
- x^* is a global minimum iff $f(x) \ge f(x^*)$ for all $x \in \mathcal{F}$

NB: Neighborhood $\mathcal{N}(x^*)$ makes no sense for MINLPs!

Challenges of Nonconvex MINLP

$$\underset{x}{\text{minimize}}\ f(x) \quad \text{subject to } c(x) \le 0,\ x \in X,\ x_i \in \mathbb{Z}\ \forall\ i \in I$$



Nonconvex $f(x)$ with three local and one global min

Challenges of Nonconvex MINLP

$$\underset{x}{\text{minimize}}\ f(x) \quad \text{subject to } c(x) \leq 0,\ x \in X,\ x_i \in \mathbb{Z}\ \forall\ i \in I$$

Remarks:

- NLP solvers are not guaranteed to find even local minima
  ... though they work remarkably well in practice!
- BnB, Benders, OA, ECP not guaranteed to find optimum
- Finding a global min is difficult ... proving it is even harder

There are many important applications of nonconvex MINLPs!

Real-Life Nonconvex Stairs

... at Hotel Les Tanneurs, Namur, Belgium

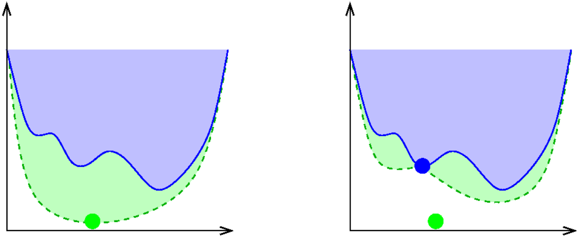
Outline

General Approach to Nonconvex MINLP

$$\underset{x}{\text{minimize}} \; f(x) \quad \text{subject to } c(x) \le 0,\; x \in X,\; x_i \in \mathbb{Z}\; \forall\, i \in I$$

Use our old MIP trick: convex relaxation!

- Relax integrality as before: $x_i \in \mathbb{R}\; \forall\, i \in I$
- New: relax $f(x) \ge \breve{f}(x)$ and constraints $c(x) \ge \breve{c}(x)$
- Ensure relaxation is tractable: e.g. $\breve{f}(x)$, $\breve{c}(x)$ convex



General Approach to Nonconvex MINLP

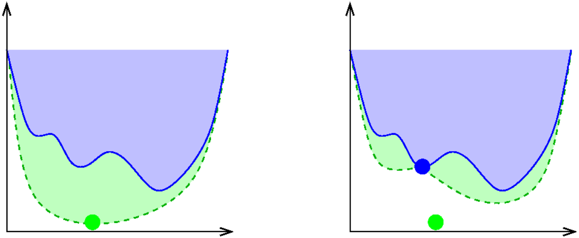
$$\underset{x}{\text{minimize}} \; f(x) \quad \text{subject to } c(x) \le 0,\; x \in X,\; x_i \in \mathbb{Z}\ \forall\, i \in I$$

Relaxation $\quad \underset{x}{\text{minimize}} \; \breve{f}(x) \quad \text{subject to } \breve{c}(x) \le 0,\; x \in X$

... gives lower bound; but solution typically infeasible in MINLP

Need constraint enforcement to guarantee convergence

- Branching on integer variables or convex underestimators
- Relaxation refinement tightens the relaxation over subdomain



Outline

Factorable Functions and MINLP

Consider MINLP with nonconvex, factorable $f(x)$ and $c(x)$

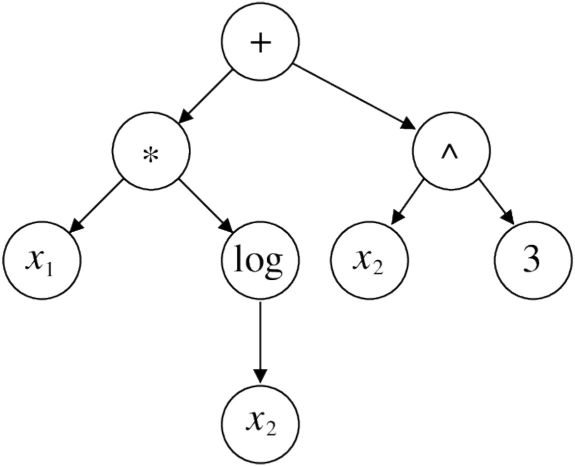
$$\underset{x}{\text{minimize}} \;\; f(x) \quad \text{subject to } c(x) \le 0,\; x \in X,\; x_i \in \mathbb{Z}\; \forall\, i \in I$$

Definition (Factorable Function)

$g(x)$ is factorable iff expressed as sum of products of unary functions of a finite set $\mathcal{O}_{\text{unary}} = \{\sin, \cos, \exp, \log, |\cdot|\}$ whose arguments are variables, constants, or other functions, which are factorable.

- Combination of functions from set of operators $\mathcal{O} = \{+, \times, /, \hat{}, \sin, \cos, \exp, \log, |\cdot|\}$.
- Excludes integrals $\int_{\xi=x_0}^{x} h(\xi) d\xi$ and black-box functions
- Represented as expression trees

Expression Tree Example



Expression tree of $f(x_1, x_2) = x_1 \log(x_2) + x_2^3$

Relaxations of Factorable Functions

MINLP with nonconvex, factorable $f(x)$ and $c(x)$

$$\underset{x}{\text{minimize}} \ f(x) \quad \text{subject to } c(x) \leq 0, \ x \in X, \ x_i \in \mathbb{Z} \ \forall \ i \in I$$

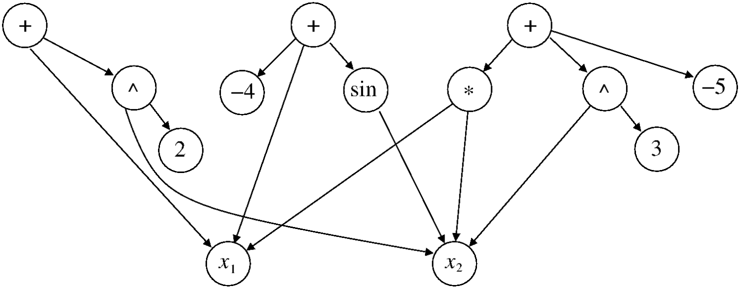
Combine expression trees of objective and constraints

- Root of each expression is $c_1(x), c_2(x), \ldots, c_m(x)$, or $f(x)$
- Associated bounds: $[-\infty, 0]$ for $c_i(x)$, and $[-\infty, \bar{\eta}]$ for $f(x)$
- Leaf nodes of all trees represent variables $x_1, x_2, \ldots, x_n$

$\Rightarrow$ gives directed acyclic graph (DAG)

Modeling languages (e.g. AMPL, GAMS) have DAG & "API"

Example of DAG



$$\begin{aligned} &\min\ x_1 + x_2^2 \\ &\text{s.t. } x_1 + \sin x_2 \leq 4, \qquad x_1 x_2 + x_2^3 \leq 5 \\ &\qquad x_1 \in [-4, 4] \cap \mathbb{Z}, \qquad x_2 \in [0, 10] \cap \mathbb{Z}. \end{aligned}$$

Three nodes without entering arcs for objective & constraints

Reformulation of Factorable MINLP

Reformulate factorable MINLP as

$$\begin{cases} \underset{x}{\text{minimize}} & x_{n+q} \\ \text{subject to} & x_k = \vartheta_k(x) \quad k = n+1, n+2, \ldots, n+q \\ & l_i \le x_i \le u_i \quad i = 1, 2, \ldots, n+q \\ & x \in X, \\ & x_i \in \mathbb{Z}, \ \forall i \in I, \end{cases}$$

see e.g. [Smith and Pantelides, 1997]

- q new auxiliary variables, $x_{n+1}, \ldots, x_{n+q}$
- ϑ_k is operator from $\mathcal{O}\{+, \times, /, \hat{}, \sin, \cos, \exp, \log\}$
- Bounds on variables written explicitly

Example of Reformulation of Factorable MINLP

$$\begin{aligned} \min\ & x_1 + x_2^2 \\ \text{s.t.}\ & x_1 + \sin x_2 \le 4, \qquad x_1 x_2 + x_2^3 \le 5 \\ & x_1 \in [-4, 4] \cap \mathbb{Z}, \qquad x_2 \in [0, 10] \cap \mathbb{Z}. \end{aligned}$$

Reformulation

$$\begin{array}{l|l|l|l|l} \min & x_9 & & & \\ \text{s.t.} & x_3 = \sin x_2 & x_7 = x_5 + x_6 - 5 & 0 \le x_2 \le 10 & 0 \le x_6 \le 1000 \\ & x_4 = x_1 + x_3 - 4 & x_8 = x_2^2 & -1 \le x_3 \le 1 & -45 \le x_7 \le 0 \\ & x_5 = x_1 x_2 & x_9 = x_1 + x_8 & -9 \le x_4 \le 0 & 0 \le x_8 \le 100 \\ & x_6 = x_2^3 & -4 \le x_1 \le 4 & -40 \le x_5 \le 40 & -4 \le x_9 \le 104 \\ & x_1, x_2, x_5, x_6, x_7, x_8, x_9 \in \mathbb{Z}. & & & \end{array}$$

- Integrality inherited from function
- Bounds inherited from function

Reformulation of Factorable MINLP

Theorem (Equivalence of Factorable Formulation)

MINLP and factorable MINLP are equivalent, i.e. optimal solutions to one can be transformed into optimal solution of the other.

Factorable form makes it easier to get convex relaxation:

- Nonconvex sets, $k = n+1, n+2, \ldots, n+q$

  $$\Theta_k = \{x \in \mathbb{R}^{n+q} : x_k = \vartheta_k(x), x \in X, l \le x \le u, x_i \in \mathbb{Z}, i \in I\}$$

  ... nonconvex due to nonlinear equality
- Let $\breve{\Theta}_k \supset \Theta_k$ convex relaxation

$$\left\{\begin{array}{lll} \underset{x}{\text{minimize}} & x_{n+q} & \\ \text{subject to} & x \in \breve{\Theta}_k & k = n+1, n+2, \ldots, n+q \\ & l_i \le x_i \le u_i & i = 1, 2, \ldots, n+q \\ & x \in X. & \end{array}\right.$$

... convex relaxation ... only look at simple sets!

Reformulation of Factorable MINLP

General convex relaxation with polyhedral sets $\breve{\Theta}_k$:

$$
\begin{cases}
\underset{x}{\text{minimize}} & x_{n+q} \\
\text{subject to } & x \in \breve{\Theta}_k \qquad k = n+1, n+2, \dots, n+q \\
& l_i \le x_i \le u_i \qquad i = 1, 2, \dots, n+q \\
& x \in X.
\end{cases}
$$

Polyhedral set $\breve{\Theta}_k$ defined by $a^k \in \mathbb{R}^{m_k}$, $B^k \in \mathbb{R}^{m_k \times (n+q)}$, and $d^k \in \mathbb{R}^{m_k}$:

$$
\breve{\Theta}_k = \{x \in \mathbb{R}^{n+q} : a^k x_k + B^k x \ge d^k, x \in X, l \le x \le u\},
$$

Gives lower bounding LP relaxation for MINLP solvers:

$$
\begin{cases}
\underset{x}{\text{minimize}} & x_{n+q} \\
\text{subject to } & a^k x_k + B^k x \ge d^k \quad k = n+1, n+2, \dots, n+q \\
& l_i \le x_i \le u_i \qquad i = 1, 2, \dots, n+q \\
& x \in X.
\end{cases}
$$

Now just need to construct polyhedral sets, see e.g. Lecture IV

Examples of Polyhedral Relaxations

Construct relaxation for each operator $\in \mathcal{O}\{+, \times, /, \hat{}, \sin, \cos, \exp, \log\}$

- Odd-degree monomials, $x_k = x_i^{2p+1}$, see [Liberti and Pantelides, 2003]
- Bilinear functions $x_k = x_i x_j$, [McCormick, 1976]
  Let $x = (x_i, x_j, x_k), L = (l_i, l_j, l_k), U = (u_i, u_j, u_k)$
  get convex hull of $\Theta_k = \{x : x_k = x_i x_j, L \le x \le U\}$:

$$x_k \ge l_j x_i + l_i x_j - l_i l_j \qquad x_k \le l_j x_i + u_i x_j - u_i l_j$$

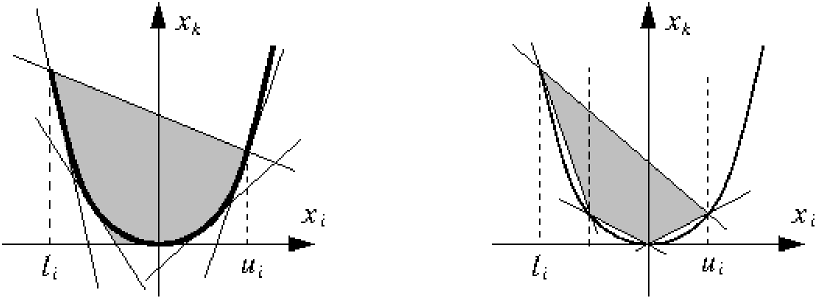
$$x_k \ge u_j x_i + u_i x_j - u_i u_j \qquad x_k \le u_j x_i + l_i x_j - l_i u_j$$

Remark

Note that tightness of convex hull depends on bounds $l_i, l_j, l_k, u_i, u_j, u_k$

Examples of Polyhedral Relaxations

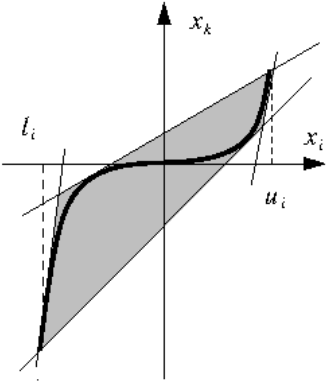
Polyhedral relaxation, $\breve{\Theta}_k$, of $x_k = x_i^2$ with x_i continuous/integer

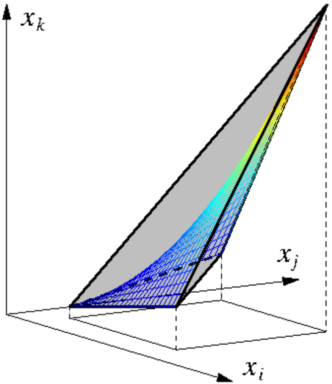


... if $x_i \in \mathbb{Z}$ then add inequalities violated at $x_i' \notin \mathbb{Z}$

Examples of Polyhedral Relaxations

Polyhedral relaxation, $\breve{\Theta}_k$, of $x_k = x_i^3$ and $x_k = x_i x_j$





Alternative Relaxation Approach

[Androulakis et al., 1995] propose α-convexification for

$$f(x) = x^T Qx + c^T x \quad \text{with} \quad x \in [l, u]$$

Lower bound obtained from:

$$\breve{f}(x) = x^T Qx + c^T x + \alpha \sum_{i=1}^{n} (x_i - l_i)(x_i - u_i).$$

which can be written as convex quadratic

$$\breve{f}(x) = x^T Px + d^T x,$$

where $P = Q + \alpha I \succeq 0$ iff $\alpha \geq -\lambda_{\min}(Q)$

Can be extended to non-quadratic functions

Solver GloMIQO [Misener and Floudas, 2012]

Outline

Spatial Branch-and-Bound (BnB)

To separate solution of relaxation use spatial BnB

- Implicit enumeration technique like integer BnB
- Recursively define partitions of feasible set into two sets
- Use reformulation outlined above
- Solve LP relaxations ($\Rightarrow$ lower bounds)
  ... and nonconvex NLPs ($\Rightarrow$ upper bound if feasible)

Classic references & Solvers:

- [Sahinidis, 1996, Tawarmalani and Sahinidis, 2002]
  BARON solver
- [Smith and Pantelides, 1997]
- [Belotti et al., 2009]
  Couenne solver ... open-source in COIN-OR

Spatial Branch-and-Bound (BnB)

Key ingredients of spatial BnB

1. Procedure to compute lower bound for subproblem
2. Procedure for partitioning feasible set of subproblem: NLP(l^-, u^-) and NLP(l^+, u^+)

... generates tree almost like integer BnB

NLP node is subproblem: NLP(l, u)

$$
\begin{cases}
\underset{x}{\text{minimize}} & f(x), \\
\text{subject to} & c(x) \le 0, \\
& x \in X \\
& l_i \le x_i \le u_i \; \forall i = 1, 2, \ldots, n \\
& x_i \in \mathbb{Z}, \; \forall i \in I
\end{cases}
$$

... restriction of original MINLP

Spatial Branch-and-Bound (BnB)

Lower bounding problem at NLP(l,u), e.g. LP(l,u)

$$\begin{cases} \underset{x}{\text{minimize}} & x_{n+q} \\ \text{subject to} & a^k x_k + B^k x \geq d^k \quad k = n+1, n+2, \ldots, n+q \\ & l_i \leq x_i \leq u_i \qquad\quad i = 1, 2, \ldots, n+q \\ & x \in X. \end{cases}$$

If LP(l,u) infeasible, then prune node.

Otherwise, $\hat{x}$ optimal solution of LP(l,u):

- If $\hat{x}$ feasible in NLP(l,u) (hence MINLP), then fathom node (new incumbent)
- If $\hat{x}$ not feasible in NLP(l,u) then ... branch ...
  1. $\hat{x}$ not integral, i.e., $\exists i \in I : \hat{x}_i \notin \mathbb{Z}$
  2. Nonconvex constraint is violated, i.e.

$$\exists k \in \{n+1, n+2, \ldots, n+q\} : \hat{x}_k \neq \vartheta_k(\hat{x}).$$

Branching for Spatial Branch-and-Bound

Two possible ways to branch (integer / nonlinear):

1. $\hat{x}$ not integral: $x_i \leq \lfloor \hat{x}_i \rfloor \vee x_i \geq \lceil \hat{x}_i \rceil$ like integer BnB
2. $\exists k : \hat{x}_k \neq \vartheta_k(\hat{x})$ nonlinear infeasible:
   - Choose branching variable x_i from arguments of $\vartheta_k(x)$
   - Branch $x_i \leq \hat{x}_i \vee x_i \geq \hat{x}_i$ ... two subproblems
   - Refine convex relaxation in each branch ... tighter bounds

Remark

Branching on $\hat{x}_k \neq \vartheta_k(\hat{x})$ leaves $\hat{x}$ feasible in both branches spatial BnB no longer finite ... different from integer BnB

Theorem (Finite Termination Smokescreen)

Spatial BnB is finite if spatial branching process is finite.

... interval arithmetic helps eliminate subproblems

Branching for Spatial Branch-and-Bound

Partition NLP(l, u) into NLP(l^-, u^-) and NLP(l^+, u^+)
... based on $x_i \leq b \vee x_i \geq b$

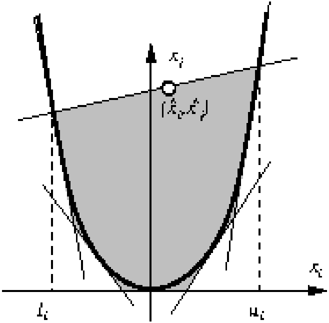
- Good performance depends on good choice of i and b
- Ideal choice balances three goals
  1. Increase both bounds LP(l^-, u^-) and LP(l^+, u^+)
  2. Shrink both feasible sets NLP(l^-, u^-) and NLP(l^+, u^+)
  3. Provide a balanced BnB tree

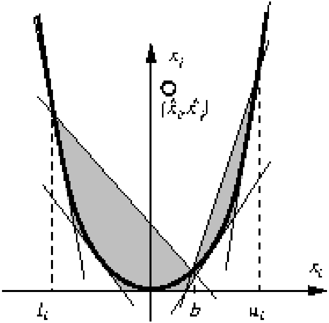
Finding continuous branching candidates x_i:

- x_i not fixed in parent problem
- x_i is argument of violated function $\hat{x}_k \neq \vartheta_k(x)$

Branching for Spatial Branch-and-Bound

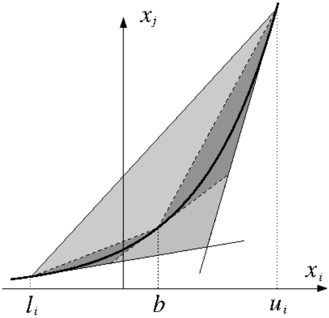
Branching example $x_k = \vartheta_k(x_i) = (x_i)^2$ violated





Branching for Spatial Branch-and-Bound

Branching example $x_k = \vartheta_k(x_i) = e^{x_i}$



Branching for Spatial Branch-and-Bound

Variable selection techniques

- Strong branching, pseudocost branching, and reliability branching generalized from MINLP
- Violation transfer:
  - Find variable x_i with largest impact on constraint violation
  - Look at all $x_k \neq \vartheta_k(x)$ for all $k = 1, 2, \ldots, n + q$

Choice of branching point b:

- Matters more than for integer branching
  ... because branch is $x_i \leq b \vee x_i \geq b$
- Ensure that $\hat{x}$ infeasible in both $\mathrm{LP}(l^-, u^-)$ and $\mathrm{LP}(l^+, u^+)$
  ... ensure refinement is good enough $\Rightarrow$ convergence "proof"

Nonconvex Branch-and-Bound

Branch-and-bound for Nonconvex MINLP

Choose tol $\epsilon > 0$, set $U = \infty$, add $(\text{NLP}(-\infty, \infty))$ to heap $\mathcal{H}$.

while $\mathcal{H} \neq \emptyset$ **do**

  Remove NLP(l, u) from heap: $\mathcal{H} = \mathcal{H} - \{\text{NLP}(l, u)\}$.

  Solve relaxation LP(l, u) $\Rightarrow$ solution $x^{(l,u)}$

  Possibly solve NLP(l, u) for an upper bound

  if *LP(l, u) is infeasible* **then**

    Prune node: infeasible

  else if $f(x^{(l,u)}) > U$ **then**

    Prune node; dominated by bound U

  else if $x_I^{(l,u)}$ *integral and* $x_k = \vartheta_k(x), \ \forall k$ **then**

    Update incumbent : $U = f(x^{(l,u)}), \ x^* = x^{(l,u)}$.

  else

    BranchOnVariable$(x_i^{(l,u)}, l, u, \mathcal{H})$

Outline

Tightening Bounds and Relaxations

Bound tightening to reduce range of bounds $x_i \in [l_i, u_i]$
... because tighter bounds $\Rightarrow$ tighter relaxations $\Rightarrow$ smaller trees

Conceptual bound-tightening procedure:
Feasible set $\mathcal{F} = \{x \in [l, u] : c(x) \leq 0,\ x \in X, x_I \in \mathbb{Z}^p\}$
Solve $2n$ (global) optimization problems, given upper bound U:

$$l_i' = \min\{x_i : x \in \mathcal{F}, f(x) \leq U\}; \qquad u_i' = \max\{x_i : x \in \mathcal{F}, f(x) \leq U\}.$$

... nonconvex MINLPs just as hard $\Rightarrow$ use relaxations:

1. FBBT: feasibility-based bound tightening
2. OBBT: optimality-based bound tightening

FBBT: Feasibility-Based Bound Tightening

FBBT broadly used:

- Artificial intelligence community & constraint programming
- NLP solvers [Messine, 2004]
- MILP solvers [Savelsbergh, 1994]

Basic Principle of FBBT

Infer bounds on x_i from tighter bounds on x_j for $j \neq i$.

Example 1: $x_j = x_i^3$ and $x_i \in [l_i, u_i]$

- Tighten interval of x_j to $[l_j, u_j] \cap [l_i^3, u_i^3]$
- Tightened l_j' on $x_j \Rightarrow$ tighter $l_i' = \sqrt[3]{l_j}$ for x_i

Example 2: $x_k = x_i x_j$ with $(1, 1, 0) \leq (x_i, x_j, x_k) \leq (5, 5, 2)$

- $l_i = l_j = 1 \Rightarrow l_k = l_i l_j = 1 > 0$
- $u_k = 2 \Rightarrow x_i \leq \frac{u_k}{l_j}$ and $x_j \leq \frac{u_k}{l_i} \Rightarrow u_i' = u_j' = 2 < 5$

FBBT: Feasibility-Based Bound Tightening

FBBT for affine functions $x_k = a_0 + \sum_{j=1}^{n} a_j x_j$ for $k > n$

- $J^+ = \{j = 1, 2, \ldots, n : a_j > 0\}$ positive coefficients
- $J^- = \{j = 1, 2, \ldots, n : a_j < 0\}$ negative coefficients

$\Rightarrow$ valid bounds are ...

$$a_0 + \sum_{j \in J^-} a_j u_j + \sum_{j \in J^+} a_j l_j \le x_k \le a_0 + \sum_{j \in J^-} a_j l_j + \sum_{j \in J^+} a_j u_j$$

Bounds $[l_k, u_k]$ on x_k give new bounds on x_j, e.g. for $j \in J^+$

$$l'_j = \frac{1}{a_j} \left(l_k - \left(a_0 + \sum_{i \in J^+ \setminus \{j\}} a_i u_i + \sum_{i \in J^-} a_i l_i \right) \right)$$

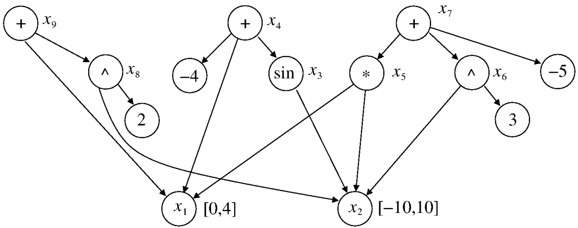
... similar for u'_j and $j \in J^-$

Better bounds from convex combination of inequalities, ...

... or solving more equations!

FBBT: Feasibility-Based Bound Tightening

For nonlinear functions, propagate bounds through DAG:



Assume solution $\hat{x}$ found with $f(\hat{x}) = 10$:

1. $10 \geq x_9 := x_1 + x_8$ and $x_1 \geq -4$ imply $x_8 \leq 14 < 100$ tighter
2. Propagate to $x_8 = x_2^2$ implies $-\sqrt{14} \leq x_2 \leq \sqrt{14}$ tightens x_2

... no more tightening

In general propagate bounds until improvement tails off.

FBBT: Feasibility-Based Bound Tightening

Properties of FBBT

- Efficient and fast implementation for large-scale MINLP
- Can exhibit poor convergence, e.g. for $\alpha > 1$ consider:
  $\min x_1$ s.t. $x_1 = \alpha x_2,\ x_2 = \alpha x_1,\ x_1 \in [-1, 1]\}$
  - Solution is $(0, 0)$
  - FBBT does not terminate in finite number of steps
  - Sequence of tighter bounds for $l = 1, 2, \ldots$ with $\{[-\frac{1}{\alpha^l}, \frac{1}{\alpha^l}]\}_l \to (0, 0)$

... hence combine with other techniques

OBBT: Optimality-Based Bound Tightening

Solving $\min / \max x_i$ s.t. $x \in \mathcal{F}$ (nonconvex MINLP) not practical
Instead, define (linear) relaxation

$$\mathcal{F}(l,u) = \left\{ x \in \mathbb{R}^{n+q} : \begin{array}{ll} a^k x_k + B^k x \geq d^k & k = n+1, n+2, \ldots, n+q \\ l_i \leq x_i \leq u_i & i = 1, 2, \ldots, n+q \\ x \in X & \end{array} \right\}.$$

Now get bounds on x_i for $i = 1, \ldots, n$ by solving $2n$ LPs:

$$\begin{aligned} l'_i &= \min\{x_i : x \in \mathcal{F}(l,u)\} \\ u'_i &= \max\{x_i : x \in \mathcal{F}(l,u)\} \end{aligned} \qquad (1)$$

... only apply at root node, or small number of nodes

Numerical Results for Branch-and-Refine

prob	basic	+presolve	+var-select	+node-select
TVC1	108861	40446	7756	8031
TVC2	fail	72270	5792	5547
TVC3	62045	861	627	627
TVC4	fail	38792	1396	1582
TVC5	fail	7369	5619	4338
TVC6	fail	12131	6096	5503

(# LPs solved)

Outline

Relaxations of Structured Nonconvex Sets

Spatial BnB for nonconvex MINLP is broadly applicable

- Branch-and-refine is one example (see Lecture IV)
- Generality of approach means that bounds can be weak
- In general, may not get convex hull of feasible set

$\Rightarrow$ search enormous trees without solving the problem

Example: Try solving nonlinear power flow with BARON!

Important to exploit structure in spatial BnB

- Look for special structure within problems
- Design tight relaxations for classes of nonconvex constraints
- Implement problem/structure specific branching rules

Nonconvex Quadratic Constraints [Mahajan and Munson, 2010]

Quadratic constraint: $x^T A x + cx + d \leq 0, \qquad x \in \mathbb{R}^n$

Applications: reactor core-reloading; power networks

- All Eigenvalues of A positive $\Rightarrow$ region is convex
- Otherwise, region is nonconvex
- Other solvers create outer approximation of feasible region:
  1. create McCormick outer approximation of terms $x_i x_j$, $\forall i \neq j$
  2. solve relaxation and branch on individual x_i

Small Example

$$\begin{aligned} \min_{x \geq 0} \; & 4x_0 + x_1 \\ \text{s.t.} \; & 7x_0^2 - 2x_1^2 + 26x_2^2 \\ & -12x_0x_1 - 8x_1x_2 \\ & +16x_0x_2 \leq -100 \end{aligned}$$

Solver	# Iterations
BARON	321
Couenne	701
MINOTAUR	2

Eigenvalues: -5, 6, 30

Identifying SOC Structure in Quadratic Constraints

- Factorize $A = QDQ^T$, Q orthogonal & D diagonal matrix
- Let $D = RER$ with E a diagonal $\{0, \pm 1\}$

$$y^T E y + b^T y + d \quad \text{where } y = RQ^T x, b = R^{-1} Q^T c$$

- If no negative eigenvalues, then convex constraint!
- If exactly one negative and no zero eigenvalues, then equivalent to two convex SOCs:

$$\Rightarrow \left\| \begin{matrix} (y_i + \frac{b_i}{2})_{i \in I_+} \\ \sqrt{\tilde{z}} \end{matrix} \right\|_2 \le \left| y_j - \frac{b_j}{2} \right| \quad \text{(of the form } \sum_{i=0}^{n-1} x_i^2 \le x_n^2) \quad (2)$$

- Separate/branch on absolute value:

$$\| \dots \|_2 \le y_j - \frac{b_j}{2} \quad \text{and} \quad \| \dots \|_2 \le -y_j + \frac{b_j}{2}$$

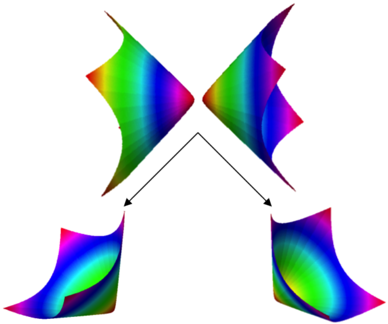
Results: Small Quadratic Instances

			# Nodes		
Inst.	Var	Con	BARON	Couenne	MINOTAUR
q1d2	2	2	39	14	2
q1d3	3	2	321	701	2
q2d6	6	4	107505	3868500	4
q3d6	6	6	301	2001	8
q3d9	9	6	>1250100	>1844800	8
q4d8	8	8	3715	29301	16
q5d10	10	10	1532839	3125701	32
q5d10b	10	10	>1033800	>2818700	32
q5d15	15	10	557905	>1321800	32
q6d12	12	12	>1358100	>3377600	64

Results: Small Quadratic Instances

			Time[s] or gap% after 1h		
Inst.	Var	Con	BARON	Couenne	MINOTAUR
q1d2	2	2	0.1	0.2	0.02
q1d3	3	2	0.50	0.7	0.03
q2d6	6	4	158.2	2498	0.2
q3d6	6	6	0.7	1.3	0.7
q3d9	9	6	16.7%	574.0%	0.3
q4d8	8	8	6.98	16.6	2.4
q5d10	10	10	2261.8	2259.9	1.8
q5d10b	10	10	145.1%	54.5%	1.8
q5d15	15	10	2519.7	27.8%	18.4
q6d12	12	12	0.4%	3.5%	9.0

Illustration of Branching on Cones



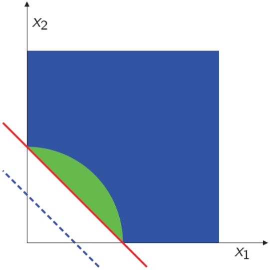
Branch on second-order cones (indefinite with one negative e.v.):

- Eigenvalue decomposition to expose structure
- Convex substructures (solved as NLPs)
- Better than thousands of little boxes (branch-and-bound)

See also "animated pdf files" ...

Another Example of Importance of Structure

Nonconvex set: $x_1^2 + x_2^2 \geq 1$ and $x_1, x_2 \in [0,2]$
Convex hull $\{X : x_1, x_2 \in [0,2] \text{ and } x_1 + x_2 \geq 1\}$



Relaxation introduces x_3 and x_4 with $x_3 \leq x_1^2$ and $x_4 \leq x_2^2$

1. Replace $x_1^2 + x_2^2 \geq 1$ by $x_3 + x_4 \geq 1$ (linear)
2. Relax nonconvex constraints $x_3 \leq x_1^2$ and $x_4 \leq x_2^2$:
   $x_3 \leq 2x_1$ and $x_4 \leq 2x_2$

Eliminate variables x_3 and x_4 and get:

$$2x_1 + 2x_2 \geq x_3 + x_4 \geq 1 \quad \Leftrightarrow \quad x_1 + x_2 \geq 1/2$$

... weaker than convex hull

General Nonconvex Quadratic Functions

Consider quadratically constrained quadratic program (QCQPs)

$$\begin{cases} \underset{x}{\text{minimize}} & x^T Q_0 x + c_0^T x, \\ \text{subject to} & x^T Q_k x + c_k^T x \le b_k, \quad k = 1, \ldots, q \\ & Ax \le b, \\ & 0 \le x \le u, \; x_i \in \mathbb{Z}, \; \forall i \in I, \end{cases}$$

... can also include integer variables and linear constraints

- Q_k $n \times n$ symmetric matrix
- A is an $m \times n$ matrix
- Q_k not necessarily convex $\Rightarrow$ nonconvex problem

General Nonconvex Quadratic Functions

Equivalent reformulation of QCQP: introduce X_{ij} for all i, j pairs

$$\begin{cases} \underset{x,X}{\text{minimize}} & Q_0 \bullet X + c_0^T x, \\ \text{subject to} & Q_k \bullet X + c_k^T x \le b_k, \quad k = 1, \ldots, q \\ & Ax \le b, \\ & 0 \le x \le u, \; x_i \in \mathbb{Z}, \; \forall i \in I, \\ & X = xx^T, \end{cases}$$

where $X = [X_{ij}]$ matrix, $Q_k \bullet X = \sum_{ij} [Q_k]_{ij} X_{ij} = \sum_{ij} [Q_k]_{ij} x_i x_j$.

- $X = xx^T$ represents nonconvex constraint $X_{ij} = x_i x_j$
  ... otherwise problem is linear!
- Relaxing equality $X = xx^T$ gives convex (linear) relaxation

... next show two approaches: RLT and SDP

General Nonconvex Quadratic Functions

Reformulation-Linearization Technique (RLT)
[Adams and Sherali, 1986]

- Get nonconvex constraints by multiplying nonnegative pairs $x_i, x_j, u_i - x_i$, and $u_j - x_j$:

$$x_i x_j \geq 0, \quad (u_i - x_i)(u_j - x_j) \geq 0, \quad x_i(u_j - x_j) \geq 0, \quad (u_i - x_i)x_j \geq 0$$

- Linearize constraints, replacing $x_i x_j$ by X_{ij}:

$$X_{ij} \geq 0, \quad X_{ij} \geq u_i x_j + u_j x_i - u_i u_j, \quad X_{ij} \leq u_j x_i, \quad X_{ij} \leq u_i x_j$$

- Replace nonconvex $X = xx^T$ by linear inequalities

$\Rightarrow$ polyhedral relaxation ... same as earlier relaxation of $x_i x_j$.

General Nonconvex Quadratic Functions

Strengthening RLT

1. Exploiting binary variables ... similar for integers
   - If $x_i \in \{0,1\}$ then $x_i^2 = x_i$ for all feasible points
   - Add linear constraints $X_{ii} = x_i$
2. Multiply linear constraints to improve RLT relaxation
   - Multiplying $x_i \geq 0$ and $b_t - \sum_{j=1}^n a_{tj} x_j \geq 0$ gives

$$b_t x_i - \sum_{j=1}^n a_{tj} x_i x_j \geq 0$$

   - Again linearize $X_{ij} = x_i x_j$ to get inequality

$$b_t x_i - \sum_{j=1}^n a_{tj} X_{ij} \geq 0$$

... generalizes to products between linear constraints

Snag: results in potentially huge LP relaxation!

General Nonconvex Quadratic Functions

Semi-Definite Programming (SDP) relaxations of QCQPs

1. Relax $X - xx^T = 0$ to $X - xx^T \succeq 0$
2. $$X - xx^T \succeq 0 \quad \Leftrightarrow \quad \begin{bmatrix} 1 & x^T \\ x & X \end{bmatrix} \succeq 0$$
3. Improve by adding $X_{ii} = x_i$ for binary $x_i \in \{0, 1\}$
4. Additional constraints by squaring & linearizing constraints

... here $H \succeq 0$ means H positive semi-definite ($x^T Hx \geq 0, \ \forall x$)

Both RLT and SDP good in practice ... RLT re-starts better!

Partial Separability and SDP Relaxations

Often Hessians Q_k have more structure, e.g. partially separable

$$q(x) = \sum_{i=1}^{l} \left(\frac{1}{2} x_{[i]}^T H_{[i]} x_{[i]} + g_{[i]}^T x_{[i]} \right)$$

Definition (Partially Separable Function)

A nonlinear function $f(x)$ is partially separable, iff

$$f(x) = \sum_{i=1}^{l} f_i(x_{[i]})$$

where $f_i : \mathbb{R}^{n_i} \to \mathbb{R}$ depends on subvector $x_{[i]}$ of x where $n_i \ll n$.

- SDP relaxation works with $n \times n$ SDP matrix
- Partially separable SDP has l matrices of size $n_i \times n_i$
- Smaller cones $\Rightarrow$ faster linear algebra

Partial Separability and SDP Relaxations

Using partial separability, we can make some $H_{[i]}$ convex

$$H = \begin{bmatrix} 4 & -1 & & \\ -1 & 4 & -1 & \\ & -1 & 4 & -1 \\ & & -1 & -1 \end{bmatrix} \qquad g = \begin{pmatrix} -1 \\ -1 \\ -1 \\ -1 \end{pmatrix}$$

Decompose as ($g_{[2]}$ exercise)

$$H_{[1]} = \begin{bmatrix} 4 & -1 & \\ -1 & 4 & -1 \\ & -1 & 4 \end{bmatrix} \quad H_{[2]} = \begin{bmatrix} 0 & -1 \\ -1 & -1 \end{bmatrix} \quad g_{[1]} = \begin{pmatrix} -1 \\ -1 \\ -1 \end{pmatrix}$$

- Get $H_{[1]} \succeq 0$ convex
- Solve order of magnitude faster

Bilinear Covering Sets

Framework for valid inequalities with "orthogonal disjunction"

Pure integer covering set for $r > 0$

$$B^{\mathrm{I}} := \left\{(x,y) \in \mathbb{Z}_+^n \times \mathbb{Z}_+^n \mid \sum_{i=1}^{n} x_i y_i \geq r\right\}$$

For any i, convex hull of two variable set

$$B_i^{\mathrm{I}} := \{(x_i, y_i) \in \mathbb{Z}_+ \times \mathbb{Z}_+ \mid x_i y_i \geq r\}$$

is polyhedron defined by $d \leq \lceil r \rceil + 1$ linear inequalities:

$$\mathrm{conv}(B_i^{\mathrm{I}}) = \left\{x \;:\; a^k x_i + b^k y_i \geq 1, \quad k = 1, \ldots, d\right\}$$

wlog (scale) right-hand-side $= 1$

Structure of $\mathrm{conv}(B_i^{\mathrm{I}})$

- Includes $x_i \geq 1$ and $y_i \geq 1$
- Other inequalities constructed as $ax_i + by_i \geq 1$:
  - Do not cut off any $(x_i^t, y_i^t) = (t, \lceil r/t \rceil)$
  - Satisfied by exactly two (x_i^t, y_i^t), for $t = 1, \ldots, \lceil r \rceil$

Bilinear Covering Sets

Let $\Pi := \left\{\pi : \{1, \ldots, n\} \to \{1, \ldots, d\}\right\}$: i.e. $\pi \in \Pi$ then $\pi(i)$ selects an inequality in $\text{conv}(B_i^{\mathrm{I}})$

Theorem (Characterization of Convex Hull $\text{conv}(B_i^{\mathrm{I}})$)

The convex hull of B^{I} is given by the set of $x \in \mathbb{R}_+^n, y \in \mathbb{R}_+^n$ that satisfy the inequalities

$$\sum_{i=1}^{n} (a^{\pi(i)} x_i + b^{\pi(i)} y_i) \geq 1, \quad \forall \pi \in \Pi$$

See [Tawarmalani et al., 2010]

$\text{conv}(B_i^{\mathrm{I}})$ has exponential number of inequalities, but have ...
... efficient separation: $\pi(i)$ index of most violated constraint

Outline

Summary and Key Points

Key Points

- General approach to nonconvex MINLP based on
  1. Decomposition of nonlinear functions $\rightarrow$ computational graph
  2. Construction of under-estimators of simple functions
- Exploiting structure is key to success
- Must exploit structure of nonconvex MINLP
- Three pillars of nonconvex MINLP:
  structure, structure, structure

Final Exam for Course Credit: Have a beer with Sven on Friday!

Office Hours: Today after the course in room 115

Adams, W. and Sherali, H. (1986).
A tight linearization and an algorithm for zero-one quadratic programming problems.
Management Science, 32(10):1274–1290.

Androulakis, I. P., Maranas, C. D., and Floudas, C. A. (1995).
αBB : A global optimization method for general constrained nonconvex problems.
Journal of Global Optimization, 7:337–363.

Belotti, P., Lee, J., Liberti, L., Margot, F., and Wächter, A. (2009).
Branching and bounds tightening techniques for non-convex MINLP.
Optimization Methods and Software, 24(4-5):597–634.

Liberti, L. and Pantelides, C. C. (2003).
Convex envelopes of monomials of odd degree.
Journal of Global Optimization, 25(2):157–168.

Mahajan, A. and Munson, T. (2010).
Exploiting second-order cone structure for global optimization.
Preprint ANL/MCS-P1801-101, Argonne National Laboratory.
Submitted to Mathematical Programming.

McCormick, G. P. (1976).
Computability of global solutions to factorable nonconvex programs: Part I — Convex underestimating problems.
Mathematical Programming, 10:147–175.

Messine, F. (2004).

Deterministic global optimization using interval constraint propagation techniques.
RAIRO-RO, 38(4):277–294.

Misener, R. and Floudas, C. (2012).
GloMIQO: Global mixed-integer quadratic optimizer.
Journal of Global Optimization, pages 1–48.

Sahinidis, N. V. (1996).
BARON: A general purpose global optimization software package.
Journal of Global Optimization, 8:201–205.

Savelsbergh, M. W. P. (1994).
Preprocessing and probing techniques for mixed integer programming problems.
ORSA Journal on Computing, 6:445–454.

Smith, E. M. B. and Pantelides, C. C. (1997).
Global optimization of nonconvex MINLPs.
Computers & Chemical Engineering, 21:S791–S796.

Tawarmalani, M., Richard, J.-P., and Chung, K. (2010).
Strong valid inequalities for orthogonal disjunctions and bilinear covering sets.
Mathematical Programming, 124:481–512.
10.1007/s10107-010-0374-6.

Tawarmalani, M. and Sahinidis, N. V. (2002).
Convexification and Global Optimization in Continuous and Mixed-Integer Nonlinear Programming: Theory, Algorithms, Software, and Applications.
Kluwer Academic Publishers, Boston MA.