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

1. Challenges of Nonconvex MINLP & General Approach
   - Challenges of Nonconvex MINLP
   - General Approach to Nonconvex MINLP

2. Generic Relaxation Strategies

3. Spatial Branch-and-Bound

4. Tightening Bounds and Relaxations

5. Exploiting Structure, Structure, and Structure

6. Summary and Conclusions
Outline

1. Challenges of Nonconvex MINLP & General Approach
   - Challenges of Nonconvex MINLP
   - General Approach to Nonconvex MINLP

2. Generic Relaxation Strategies

3. Spatial Branch-and-Bound

4. Tightening Bounds and Relaxations

5. Exploiting Structure, Structure, and Structure

6. Summary and Conclusions
Challenges of Nonconvex MINLP

Mixed-Integer Nonlinear Program (MINLP)

\[
\min_{x} f(x) \quad \text{subject to } c(x) \leq 0, \ x \in X, \ x_i \in \mathbb{Z} \ \forall \ i \in I
\]

... 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

\[
\min_{x} f(x) \quad \text{subject to } x \in \mathcal{F} := \{x : c(x) \leq 0, \ x \in X\}
\]

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

\(x^*\) is a global minimum iff \(f(x) \geq f(x^*)\) for all \(x \in \mathcal{F}\)

NB: Neighborhood \(\mathcal{N}(x^*)\) makes no sense for MINLPs!
Challenges of Nonconvex MINLP

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

Nonconvex \( f(x) \) with three local and one global min
Challenges of Nonconvex MINLP

\[ \min_x f(x) \quad \text{subject to } c(x) \leq 0, \ x \in X, \ x_i \in \mathbb{Z} \ \forall \ i \in I \]

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

1. Challenges of Nonconvex MINLP & General Approach
   - Challenges of Nonconvex MINLP
   - General Approach to Nonconvex MINLP

2. Generic Relaxation Strategies

3. Spatial Branch-and-Bound

4. Tightening Bounds and Relaxations

5. Exploiting Structure, Structure, and Structure

6. Summary and Conclusions
General Approach to Nonconvex MINLP

\[
\min_{x} f(x) \quad \text{subject to } c(x) \leq 0, \ x \in X, \ x_i \in \mathbb{Z} \ \forall \ i \in I
\]

Use our old MIP trick: \textit{convex relaxation}!

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

![Diagram showing convex relaxation](image)
General Approach to Nonconvex MINLP

\[
\begin{align*}
\min_{x} & \quad f(x) \quad \text{subject to} \quad c(x) \leq 0, \ x \in X, \ x_i \in \mathbb{Z} \ \forall \ i \in I \\
\text{Relaxation} & \quad \min_{x} \quad \tilde{f}(x) \quad \text{subject to} \quad \tilde{c}(x) \leq 0, \ x \in X \\
\end{align*}
\]

... 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

1. Challenges of Nonconvex MINLP & General Approach
   - Challenges of Nonconvex MINLP
   - General Approach to Nonconvex MINLP

2. Generic Relaxation Strategies

3. Spatial Branch-and-Bound

4. Tightening Bounds and Relaxations

5. Exploiting Structure, Structure, and Structure

6. Summary and Conclusions
Factorable Functions and MINLP

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

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

**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 $\mathcal{O} = \{+, \times, /, ^, \sin, \cos, \exp, \log, | \cdot |\}$.
- Excludes integrals $\int_{\xi=x_0}^{x} h(\xi) d\xi$ and black-box functions
- Represented as expression trees
Expression tree 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 \( c(x) \)

\[
\minimize_{x} f(x) \quad \text{subject to } c(x) \leq 0, \ x \in X, \ x_i \in \mathbb{Z} \ \forall \ i \in I
\]

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

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

Three nodes without entering arcs for objective & constraints
Reformulation of Factorable MINLP

Reformulate factorable MINLP as

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

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

- $q$ new auxiliary variables, $x_{n+1}, \ldots, x_{n+q}$
- $\vartheta_k$ is operator from $\mathcal{O}\{+, \times, /, ^, \sin, \cos, \exp, \log\}$
- Bounds on variables written explicitly
Example of Reformulation of Factorable MINLP

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

Reformulation

\[
\begin{align*}
\min \quad & x_9 \\
\text{s.t.} \quad & x_3 = \sin x_2 \\
& x_4 = x_1 + x_3 - 4 \\
& x_5 = x_1 x_2 \\
& x_6 = x_2^3 \\
& x_7 = x_5 + x_6 - 5 \\
& x_8 = x_2^2 \\
& x_9 = x_1 + x_8 \\
& -4 \leq x_1 \leq 4 \\
& -9 \leq x_4 \leq 0 \\
& -40 \leq x_5 \leq 40 \\
& -45 \leq x_7 \leq 0 \\
& 0 \leq x_6 \leq 1000 \\
& 0 \leq x_8 \leq 100 \\
& 0 \leq x_9 \leq 104 \\
& x_1, x_2, x_5, x_6, x_7, x_8, x_9 \in \mathbb{Z}.
\end{align*}
\]

- 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, $k = n + 1, n + 2, \ldots, n + q$
  \[
  \Theta_k = \{ x \in \mathbb{R}^{n+q} : x_k = \vartheta_k(x), x \in X, l \leq x \leq u, x_i \in \mathbb{Z}, i \in I \}
  \]

  ... nonconvex due to nonlinear equality

- Let $\tilde{\Theta}_k \supset \Theta_k$ convex relaxation

  \[
  \begin{aligned}
  &\text{minimize} & & x_{n+q} \\
  &\text{subject to} & & x \in \tilde{\Theta}_k \\
  & & & k = n + 1, n + 2, \ldots, n + q \\
  & & & l_i \leq x_i \leq u_i \\
  & & & i = 1, 2, \ldots, n + q \\
  & & & x \in X.
  \end{aligned}
  \]

  ... convex relaxation ... only look at simple sets!
Reformulation of Factorable MINLP

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

\[
\begin{aligned}
\text{minimize} & \quad x_{n+q} \\
\text{subject to} & \quad x \in \tilde{\Theta}_k \\
& \quad k = n + 1, n + 2, \ldots, n + q \\
& \quad l_i \leq x_i \leq u_i \\
& \quad i = 1, 2, \ldots, n + q \\
& \quad x \in X.
\end{aligned}
\]

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

\[
\tilde{\Theta}_k = \{ x \in \mathbb{R}^{n+q} : a^k x_k + B^k x \geq d^k, x \in X, l \leq x \leq u \},
\]

Gives lower bounding LP relaxation for MINLP solvers:

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

Now just need to construct polyhedral sets, see e.g. Lecture IV
Examples of Polyhedral Relaxations

Construct relaxation for each operator
∈ \( O\{+, \times, /, ^{\hat{\text{\,}}}, \sin, \cos, \exp, \log\} \)

- 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 \( \Theta_k = \{x : x_k = x_i x_j, L \leq x \leq U\} : \)

\[
\begin{align*}
  x_k &\geq l_j x_i + l_i x_j - l_i l_j \\
  x_k &\geq u_j x_i + u_i x_j - u_i u_j
\end{align*} \quad \begin{align*}
  x_k &\leq l_j x_i + u_i x_j - u_i l_j \\
  x_k &\leq u_j x_i + l_i x_j - l_i u_j
\end{align*}
\]

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, $\tilde{\Theta}_k$, 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, $\tilde{\Theta}_k$, of $x_k = x_i^3$ and $x_k = x_ix_j$
Alternative Relaxation Approach

[Androulakis et al., 1995] propose $\alpha$-convexification for

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

Lower bound obtained from:

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

which can be written as convex quadratic

$$\tilde{f}(x) = x^T P x + d^T x,$$

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

1. Challenges of Nonconvex MINLP & General Approach
   - Challenges of Nonconvex MINLP
   - General Approach to Nonconvex MINLP

2. Generic Relaxation Strategies

3. Spatial Branch-and-Bound

4. Tightening Bounds and Relaxations

5. Exploiting Structure, Structure, and Structure

6. Summary and Conclusions
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 (⇒ lower bounds)
  ... and nonconvex NLPs (⇒ 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: \( \text{NLP}(l^-, u^-) \) and \( \text{NLP}(l^+, u^+) \)

... generates tree almost like integer BnB

NLP node is subproblem: \( \text{NLP}(l, u) \)

\[
\begin{align*}
\text{minimize} & \quad f(x), \\
\text{subject to} & \quad c(x) \leq 0, \\
& \quad x \in X \\
& \quad l_i \leq x_i \leq u_i \quad \forall i = 1, 2, \ldots, n \\
& \quad x_i \in \mathbb{Z}, \quad \forall i \in I
\end{align*}
\]

... restriction of original MINLP
Spatial Branch-and-Bound (BnB)

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

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

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: \( x_i \leq \lfloor \hat{x}_i \rfloor \lor x_i \geq \lceil \hat{x}_i \rceil \) 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 \( x_i \leq \hat{x}_i \lor x_i \geq \hat{x}_i \) ... 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 \( \text{NLP}(l, u) \) into \( \text{NLP}(l^-, u^-) \) and \( \text{NLP}(l^+, u^+) \)

... based on \( x_i \leq b \lor x_i \geq b \)

- Good performance depends on good choice of \( i \) and \( b \)
- Ideal choice balances three goals
  1. Increase both bounds \( \text{LP}(l^-, u^-) \) and \( \text{LP}(l^+, u^+) \)
  2. Shrink both feasible sets \( \text{NLP}(l^-, u^-) \) and \( \text{NLP}(l^+, u^+) \)
  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 example $x_k = \vartheta_k(x_i) = (x_i)^2$ violated.
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 $x_i \leq b \lor x_i \geq b$
- Ensure that $\hat{x}$ infeasible in both $LP(l^-, u^-)$ and $LP(l^+, u^+)$
  ... 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 (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($l$, $u$) ⇒ 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

1. Challenges of Nonconvex MINLP & General Approach
   - Challenges of Nonconvex MINLP
   - General Approach to Nonconvex MINLP

2. Generic Relaxation Strategies

3. Spatial Branch-and-Bound

4. Tightening Bounds and Relaxations

5. Exploiting Structure, Structure, and Structure

6. Summary and Conclusions
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 $U$:

$$l'_i = \min\{x_i : x \in \mathcal{F}, f(x) \leq U\}; \quad u'_i = \max\{x_i : x \in \mathcal{F}, f(x) \leq U\}.$$  

... 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 ...

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

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!
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:
  $$\min x_1 \text{ s.t. } x_1 = \alpha x_2, \ x_2 = \alpha x_1, \ x_1 \in [-1, 1]$$
  - Solution is $(0, 0)$
  - FBBT does not terminate in finite number of steps
  - Sequence of tighter bounds for $l = 1, 2, \ldots$ with
    $$\left[-\frac{1}{\alpha^l}, \frac{1}{\alpha^l}\right], \ l \to (0, 0)$$
    ... 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

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

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

\[
l'_i = \min\{x_i : x \in \mathcal{F}(l, u)\}
\]

\[
u'_i = \max\{x_i : x \in \mathcal{F}(l, u)\}
\]

... only apply at root node, or small number of nodes
### Numerical Results for Branch-and-Refine

<table>
<thead>
<tr>
<th>prob</th>
<th>basic</th>
<th>+presolve</th>
<th>+var-select</th>
<th>+node-select</th>
</tr>
</thead>
<tbody>
<tr>
<td>TVC1</td>
<td>108861</td>
<td>40446</td>
<td>7756</td>
<td>8031</td>
</tr>
<tr>
<td>TVC2</td>
<td>fail</td>
<td>72270</td>
<td>5792</td>
<td>5547</td>
</tr>
<tr>
<td>TVC3</td>
<td>62045</td>
<td>861</td>
<td>627</td>
<td>627</td>
</tr>
<tr>
<td>TVC4</td>
<td>fail</td>
<td>38792</td>
<td>1396</td>
<td>1582</td>
</tr>
<tr>
<td>TVC5</td>
<td>fail</td>
<td>7369</td>
<td>5619</td>
<td>4338</td>
</tr>
<tr>
<td>TVC6</td>
<td>fail</td>
<td>12131</td>
<td>6096</td>
<td>5503</td>
</tr>
</tbody>
</table>

(\# LPs solved)
Outline

1. Challenges of Nonconvex MINLP & General Approach
   - Challenges of Nonconvex MINLP
   - General Approach to Nonconvex MINLP

2. Generic Relaxation Strategies

3. Spatial Branch-and-Bound

4. Tightening Bounds and Relaxations

5. Exploiting Structure, Structure, and Structure

6. Summary and Conclusions
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
  ⇒ 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: \( x^T Ax + cx + d \leq 0, \quad x \in \mathbb{R}^n \)

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

\[
\begin{align*}
\text{min} \quad & 4x_0 + x_1 \\
\text{s.t.} \quad & 7x_0^2 - 2x_1^2 + 26x_2^2 \\
& -12x_0x_1 - 8x_1x_2 \\
& +16x_0x_2 \leq -100 \\
& x \geq 0
\end{align*}
\]

Eigenvalues: -5, 6, 30

<table>
<thead>
<tr>
<th>Solver</th>
<th># Iterations</th>
</tr>
</thead>
<tbody>
<tr>
<td>BARON</td>
<td>321</td>
</tr>
<tr>
<td>Couenne</td>
<td>701</td>
</tr>
<tr>
<td>MINOTAURO</td>
<td>2</td>
</tr>
</tbody>
</table>
Identifying SOC Structure in Quadratic Constraints

- Factorize $A = QDQ^T$, $Q$ orthogonal & $D$ diagonal matrix.
- Let $D = RER$ with $E$ a diagonal $\{0, \pm1\}$

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

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

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

- Separate/branch on absolute value:

$$\left\| \ldots \right\|_2 \leq y_j - \frac{b_j}{2} \quad \text{and} \quad \left\| \ldots \right\|_2 \leq -y_j + \frac{b_j}{2}$$
## Results: Small Quadratic Instances

<table>
<thead>
<tr>
<th>Inst.</th>
<th>Var</th>
<th>Con</th>
<th>BARON</th>
<th>Couenne</th>
<th>MINOTAURO</th>
</tr>
</thead>
<tbody>
<tr>
<td>q1d2</td>
<td>2</td>
<td>2</td>
<td>39</td>
<td>14</td>
<td>2</td>
</tr>
<tr>
<td>q1d3</td>
<td>3</td>
<td>2</td>
<td>321</td>
<td>701</td>
<td>2</td>
</tr>
<tr>
<td>q2d6</td>
<td>6</td>
<td>4</td>
<td>107505</td>
<td>3868500</td>
<td>4</td>
</tr>
<tr>
<td>q3d6</td>
<td>6</td>
<td>6</td>
<td>301</td>
<td>2001</td>
<td>8</td>
</tr>
<tr>
<td>q3d9</td>
<td>9</td>
<td>6</td>
<td>&gt;1250100</td>
<td>&gt;1844800</td>
<td>8</td>
</tr>
<tr>
<td>q4d8</td>
<td>8</td>
<td>8</td>
<td>3715</td>
<td>29301</td>
<td>16</td>
</tr>
<tr>
<td>q5d10</td>
<td>10</td>
<td>10</td>
<td>1532839</td>
<td>3125701</td>
<td>32</td>
</tr>
<tr>
<td>q5d10b</td>
<td>10</td>
<td>10</td>
<td>&gt;1033800</td>
<td>&gt;2818700</td>
<td>32</td>
</tr>
<tr>
<td>q5d15</td>
<td>15</td>
<td>10</td>
<td>557905</td>
<td>&gt;1321800</td>
<td>32</td>
</tr>
<tr>
<td>q6d12</td>
<td>12</td>
<td>12</td>
<td>&gt;1358100</td>
<td>&gt;3377600</td>
<td>64</td>
</tr>
</tbody>
</table>
## Results: Small Quadratic Instances

<table>
<thead>
<tr>
<th>Inst.</th>
<th>Var</th>
<th>Con</th>
<th>BARON [s]</th>
<th>Couenne</th>
<th>MINOTAUR</th>
</tr>
</thead>
<tbody>
<tr>
<td>q1d2</td>
<td>2</td>
<td>2</td>
<td>0.1</td>
<td>0.2</td>
<td>0.02</td>
</tr>
<tr>
<td>q1d3</td>
<td>3</td>
<td>2</td>
<td>0.50</td>
<td>0.7</td>
<td>0.03</td>
</tr>
<tr>
<td>q2d6</td>
<td>6</td>
<td>4</td>
<td>158.2</td>
<td>2498</td>
<td>0.2</td>
</tr>
<tr>
<td>q3d6</td>
<td>6</td>
<td>6</td>
<td>0.7</td>
<td>1.3</td>
<td>0.7</td>
</tr>
<tr>
<td>q3d9</td>
<td>9</td>
<td>6</td>
<td>16.7%</td>
<td>574.0%</td>
<td>0.3</td>
</tr>
<tr>
<td>q4d8</td>
<td>8</td>
<td>8</td>
<td>6.98</td>
<td>16.6</td>
<td>2.4</td>
</tr>
<tr>
<td>q5d10</td>
<td>10</td>
<td>10</td>
<td>2261.8</td>
<td>2259.9</td>
<td>1.8</td>
</tr>
<tr>
<td>q5d10b</td>
<td>10</td>
<td>10</td>
<td>145.1%</td>
<td>54.5%</td>
<td>1.8</td>
</tr>
<tr>
<td>q5d15</td>
<td>15</td>
<td>10</td>
<td>2519.7</td>
<td>27.8%</td>
<td>18.4</td>
</tr>
<tr>
<td>q6d12</td>
<td>12</td>
<td>12</td>
<td>0.4%</td>
<td>3.5%</td>
<td>9.0</td>
</tr>
</tbody>
</table>
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:

\[
2x_1 + 2x_2 \geq x_3 + x_4 \geq 1 \iff x_1 + x_2 \geq 1/2
\]

... weaker than convex hull
Consider quadratically constrained quadratic program (QCQPs)

\[
\begin{align*}
\text{minimize } \quad & \quad x^T Q_0 x + c_0^T x, \\
\text{subject to } \quad & \quad x^T Q_k x + c_k^T x \leq b_k, \quad k = 1, \ldots, q \\
& \quad A x \leq b, \\
& \quad 0 \leq x \leq u, \quad x_i \in \mathbb{Z}, \quad \forall i \in I,
\end{align*}
\]

... 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

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

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

- Get nonconvex constraints by multiplying nonnegative pairs \( x_i, x_j, u_i - x_i, \text{ and } u_j - x_j \):
  \[
  x_i x_j \geq 0, \quad (u_i - x_i)(u_j - x_j) \geq 0, \quad x_i (u_j - x_j) \geq 0, \quad (u_i - x_i)x_j \geq 0
  \]

- 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!
Semi-Definite Programming (SDP) relaxations of QCQPs

1. Relax $X - xx^T = 0$ to $X - xx^T \succeq 0$

2. $X - xx^T \succeq 0 \iff \begin{bmatrix} 1 & x^T \\ x & X \end{bmatrix} \succeq 0$

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 ($x^T H x \geq 0, \forall x$)

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 $f_i : \mathbb{R}^{n_i} \rightarrow \mathbb{R}$ 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

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

Decompose as ($g[2]$ exercise)

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

- 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 \( r > 0 \)

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

For any \( i \), convex hull of two variable set

\[
B^I_i := \left\{ (x_i, y_i) \in \mathbb{Z}_+ \times \mathbb{Z}_+ \mid x_i y_i \geq r \right\}
\]

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

\[
\text{conv}(B^I_i) = \left\{ x : a^k x_i + b^k y_i \geq 1, \quad k = 1, \ldots, d \right\}
\]

wlog (scale) right-hand-side = 1
Structure of \( \text{conv}(B^I_i) \)

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

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

The convex hull of $B_i^I$ 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]

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

1. Challenges of Nonconvex MINLP & General Approach
   - Challenges of Nonconvex MINLP
   - General Approach to Nonconvex MINLP

2. Generic Relaxation Strategies

3. Spatial Branch-and-Bound

4. Tightening Bounds and Relaxations

5. Exploiting Structure, Structure, and Structure

6. Summary and Conclusions
Summary and Key Points

Key Points

- General approach to nonconvex MINLP based on
  - Decomposition of nonlinear functions $\rightarrow$ computational graph
  - 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


Deterministic global optimization using interval constraint propagation techniques.


GloMIQO: Global mixed-integer quadratic optimizer. 

BARON: A general purpose global optimization software package. 
Journal of Global Optimization, 8:201–205.

Preprocessing and probing techniques for mixed integer programming problems. 

Global optimization of nonconvex MINLPs. 

Strong valid inequalities for orthogonal disjunctions and bilinear covering sets. 
10.1007/s10107-010-0374-6.

Kluwer Academic Publishers, Boston MA.