Branch-and-Refine for Global Optimization of Mixed-Integer Nonconvex Optimization Problems

Sven Leyffer, leyffer@mcs.anl.gov
Mathematics & Computer Science Division
Argonne National Laboratory

joint work with Annick Sartenaer and **Emilie Wanufelle**

University of Namur

MIP 2008, Columbia University, August 6, 2008

Overview

1. Applications & Motivation
   - Tertiary Voltage Control
   - Blackout Prevention in National Power Grid
   - Problem Characteristics
2. Branch-and-Refine for Nonconvex MINLPs
   - Decomposition of Nonlinear Functions
   - Piecewise Polyhedral Envelopes
   - Branch-and-Refine
3. Preliminary Numerical Results
   - Preliminary Numerical Experience
   - Conclusions & Future Research

Tertiary Voltage Control

Efficient management of electricity grid

- optimal power flow problem in network
- predict effects of network expansion
- special case: tertiary voltage control (TVC)

Variables (complex for alternating current)

- voltage: difference in electric potential between nodes
- current: flow of electricity
- power: quantity of energy transferred
  $S = P + iQ$ real/reactive power use polar coordinates

$\Rightarrow$ nonconvex mixed-integer nonlinear program (MINLP)

Tertiary Voltage Control

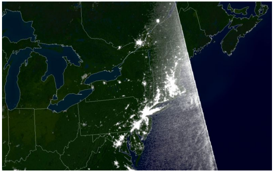
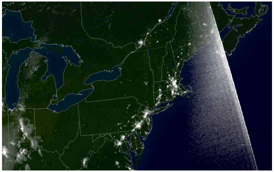
Example: expression for real power (reactive power similar):

$$P_{ij} = \nu_i^2 \, (y_{ij} \cos(\zeta_{ij}) + g_{ij}) - \nu_i \nu_j y_{ij} \cos(\zeta_{ij} + \theta_i - \theta_j)$$

if branch $ij \in S_i^e$ etc. where

- g_{ij} conductance
- ν_i is modulus of voltage at node i
- θ_j argument of voltage at node i
- y_{ij} is modulus of admittance on branch ij
- ζ_{ij} argument of admittance on branch ij

Blackout Prevention in National Power Grid

area affected by 2003 blackout: before and during

- 2003 blackout cost $4-10 billion and affected 50 million people
- prevent with contingency analysis
  - find least number of lines whose removal results in failure
  - binary variables model removal of lines
  - nonlinearities model power flow
  - results in large integer optimization problem
- current limit: 10s of lines

Grand Challenge Problem

Power Grid Problem Characteristics

TVC & Blackout Prevention are nonconvex MINLP

$$(P) \begin{cases} \underset{x,y}{\text{minimize}} & g_0(x,y), \\ \text{subject to} & g_i(x,y) = 0, \; i = 1,..,m, \\ & (x,y) \in P, \; y \in \mathbb{Z}^t \end{cases}$$

where

- nonconvex equality constraints
  nonlinear components: $\sin(w)$, $\cos(w)$, w^2, $w_1 w_2$
- P is polyhedral set: bounds & network constraints
- discrete variables:
  - TVC: transformer ratios, capacitor banks $z_i \nu_i^2 Q_{0_i}$
  - Blackout Prevention: removal of lines

Experience with BARON

BARON: branch-and-reduce [Sahinidis & Tawarmalani, 2002]

- outer approximation from global under- and over-estimators
- library of under-estimators
- branch & reduce range of variables
- cannot be applied to sin or cos:
  ... only trivial under- over-estimators ± 1

Abuse BARON: Taylor expand sin-terms and cos-terms . . .

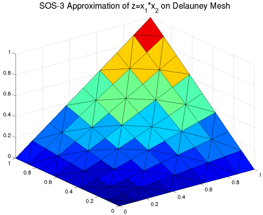
$\Rightarrow$ BARON fails: no improvement of lower bound after 10 hours!

. . . not really a fair comparison!

Experience with SOS Approximation

[Beale & Tomlin, 1970]: special-ordered sets (SOS)

- piecewise linear approx. of $h(x)$
- choose breakpoints $\bar{x}_k$, $k \in I_M$
- partition $\otimes_{q=1}^{n}[l_q, u_q]$ into simplices
- approximation $\tilde{h}(x)$ with $\lambda_k \geq 0$



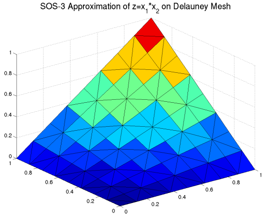
$$\tilde{h}(x) = \sum_{k \in I_M} \lambda_k h(\bar{x}_k), \quad x = \sum_{k \in I_M} \lambda_k \bar{x}_k, \quad 1 = \sum_{k \in I_M} \lambda_k$$

- SOS condition: at most $n+1$ λ_k non-zero on single simplex

... see [Martin et al, 2006] for recent implementation

Pitfall 1: Exponential Complexity of SOS

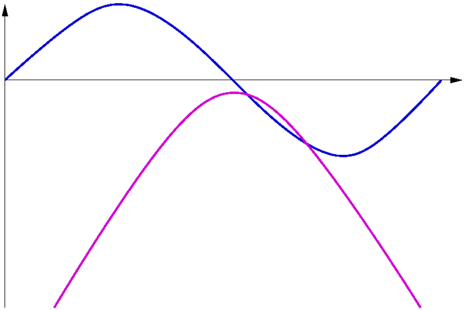
SOS-approximation has exponential complexity:



- approximate $h(x, y)$ for $(x, y) \in \mathbb{R}^n$
- p breakpoints in each dimension
  $\Rightarrow$ p^n SOS-variables λ_i

e.g. expression for real power has $n = 8$ variables ... impractical

Pitfall 2: Infeasible SOS Approximations

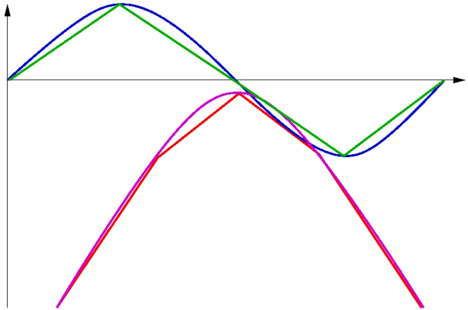


SOS approx infeasible

$\sin(x) = 0$
$-0.35\ (x - \pi)^2 - 0.3 = 0$

...observed infeasible SOS on early examples

Pitfall 2: Infeasible SOS Approximations

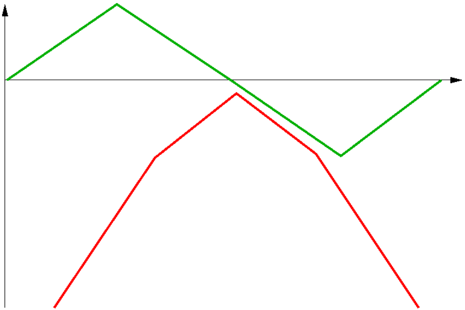


SOS approx infeasible

$\sin(x) = 0$
$-0.35\,(x-\pi)^2 - 0.3 = 0$

...observed infeasible SOS on early examples

Pitfall 2: Infeasible SOS Approximations



SOS approx infeasible

$\sin(x) = 0$
$-0.35\ (x - \pi)^2 - 0.3 = 0$

...observed infeasible SOS on early examples

Overview

Remedy 1: Decomposition of Nonlinear Functions

SOS-approximation needs p^n SOS-variables λ_k

Idea: decompose $h(x, y)$ into simpler functions:

$$\begin{array}{rcll} w_j & = & x_j & j = 1, \ldots, s, \\ w_{s+j} & = & y_j & j = 1, \ldots, t, \\ w_{s+t+j} & = & h_j(w_{j_1}\{, w_{j_2}\}) & j = 1, \ldots, K, \\ h(x, y) & = & w_{s+t+K}, & \end{array}$$

where h_j are univariate or bivariate and $j_1, j_2 < s + t + j$

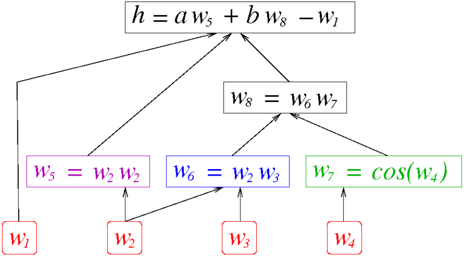
Remedy 1: Decomposition of Nonlinear Functions

Consider

$$h(x_1, x_2, x_3, x_4) = ax_2^2 + bx_2x_3\cos(x_4) - x_1$$

where a and b constants.

- $w_j = x_j \quad j = 1, \ldots,$
- $w_5 = w_2^2$
- $w_6 = w_2 w_3$
- $w_7 = \cos(w_4)$
- $w_8 = w_6 w_7$
- $h = aw_5 + bw_8 - w_1$



Decomposition not unique: e.g. $w_6 = \cos(w_4)$ etc.

Remedy 1: Decomposition of Nonlinear Functions

Example: Expression for active power

$$P_{ij} = \nu_i^2 \; (y_{ij} \cos(\zeta_{ij}) + g_{ij}) - \nu_i \nu_j y_{ij} \cos(\zeta_{ij} + \theta_i - \theta_j)$$

Simple functions:

- ν_i^2
- $\cos(\zeta_{ij})$
- $\sin(w_{j_1})$, where $w_{j_1} = \zeta_{ij} + \theta_i - \theta_j$
- 5 bilinear terms like $\nu_i \nu_j$

$\Rightarrow$ need only $5p^2 + 3p$ SOS variables, λ ... much smaller p^8

Decomposed Nonconvex MINLP

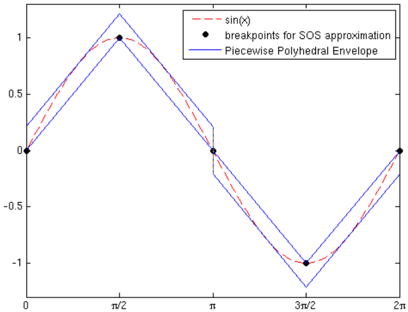
Decomposing every function gives

$$(D) \begin{cases} \underset{x,y,w}{\text{minimize}} & w_{0,K_0} \\ \text{subject to} & w_{ij} = x_j & \forall i,j \\ & w_{i,s+j} = y_j & \forall i,j \\ & w_{i,s+t+j} = g_{ij}(w_{i,j_1}\{,w_{i,j_2}\}) & \forall i,j \\ & w_{i,s+t+K_i} = 0 & i = 1,\dots,m \\ & (x,y) \in P,\ y \in \mathbb{Z}^t. & \end{cases}$$

…equivalent to MINLP (P) ... related to automatic differentiation
where $g_{ij}(w_{i,j_1}\{,w_{i,j_2}\})$ univariate/bivariate component of $g_i(x,y)$

Remedy 2: Piecewise Polyhedral Envelopes

Idea: Outer approximation by piecewise polyhedral envelopes



Univariate $w_h = h(w)$ becomes

$$\sum_{k \in I} \lambda_k \left(h(w_k) - L_k \right) \leq w_h \leq \sum_{k \in I} \lambda_k \left(h(w_k) + U_k \right)$$

Remedy 2: Piecewise Polyhedral Envelopes

Obtain bound L_k by solving

$$L_k = \max_{w \in [w^k, w^{k+1}], \lambda^k + \lambda^{k+1} = 1} \left(0, \lambda^k h(w^k) + \lambda^{k+1} h(w^{k+1}) - h(w)\right)$$

...similar for U_k

Bounds L_k, U_k pre-computed on $[w_k, w_{k+1}]$, e.g. $h(w) = w^2$:

$$L_k = (w_{k+1} - w_k)^2/4, \qquad U_k = 0$$

See Emilie's thesis for other functions ...

Piecewise Polyhedral Envelopes for $h = x\ y$

Theorem: Every (x, y, xy) with $l_x \leq x \leq u_x$ and $l_y \leq y \leq u_y$ is unique convex combination of $(l_x, l_y, l_x l_y)$, $(l_x, u_y, l_x u_y)$, $(u_x, l_y, u_x l_y)$ and $(u_x, u_y, u_x u_y)$, i.e. $\exists \lambda_i \geq 0,\ i = 1, \ldots, 4$:

$$\begin{pmatrix} x \\ y \\ xy \\ 1 \end{pmatrix} = \begin{bmatrix} l_x & l_x & u_x & u_x \\ l_y & u_y & l_y & u_y \\ l_x l_y & l_x u_y & u_x l_y & u_x u_y \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{pmatrix}$$

Implies $L_k = U_k = 0$, and equality (tighter relaxation):

$$w_{xy} = \sum_{(i,j) \in I} \lambda_{ij} x_i y_j$$

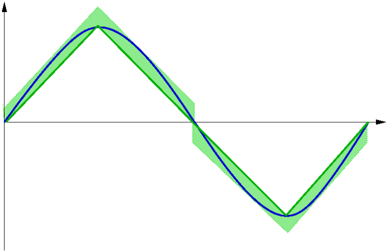
Piecewise Envelope Problem

Proposition: (E) is an outer approximation of (D) and hence (P).

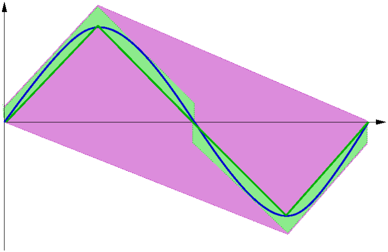
$$
(E) \left\{
\begin{array}{ll}
\underset{x,y,w,\lambda}{\text{minimize}} & w_{0,K_0} \\
\text{subject to} & w_{ij} = x_j, \qquad w_{i,s+j} = y_j \\
& x_j = \sum_{k \in I_j} \lambda_{j_k} x_{j_k}, \qquad 1 = \sum_{k \in I_j} \lambda_{j_k} \\
& w_{i,s+t+j} \geq \sum_{k \in I_{ij}} \lambda_{ij}^k \left(g_{ij}(w_{i,j_1}^k \{, w_{i,j_2}^k\}) - L_{ijk} \right) \\
& w_{i,s+t+j} \leq \sum_{k \in I_{ij}} \lambda_{ij}^k \left(g_{ij}(w_{i,j_1}^k \{, w_{i,j_2}^k\}) + U_{ijk} \right) \\
& w_{i,s+t+K_i} = 0 \\
& (x,y) \in P, \; y\mathbb{Z}^t, \text{ and } w \in W,
\end{array}
\right.
$$

where W bounds deduced from (x, y) bounds.

Piecewise Envelope Problem: Illustration



SOS Outer Approximation



Convex Hull

Branch-and-Refine: Outline

Classical Branch-and-Bound:

Solve envelope problem (E) branch on SOS-condition or $y \in \mathbb{Z}^t$

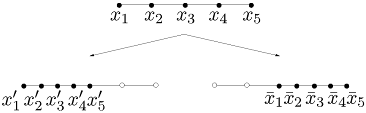
$\Rightarrow$ large discretization error or large number of λ_k variables

Idea: Instead refine discretization after branching:

- tighten envelope as we go down tree: refine
- exploit exactness of bilinear terms $w_1 w_2$
- better numerical results

Branch-and-Refine: Branching

1D SOS



2D SOS

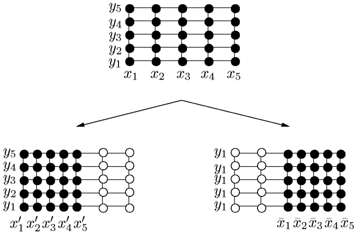


Illustration of branching and refinement

Branch-and-Refine: Fathoming Rules

Also solve NLP(X_k, Y_k):

$$\begin{cases} z_{\text{NLP}_k} := \underset{x,y}{\text{minimize}} & g_0(x,y) \\ \text{subject to} & g_i(x,y) = 0, \ i = 1,..,m \\ & x \in X_k, \ y \in Y_k \end{cases}$$

...upper bound on node (X_k, Y_k).

Fathoming Rules:

1. infeasible LP relaxation
2. NLP(X_k, Y_k) solution same as LP(X_k, Y_k) relaxation
3. LP relaxation dominated by incumbent

Branch-and-Refine: Algorithm

set $U = \infty$, $k = 1$ & put $\mathrm{LP}(X_k, Y_k)$ on stack
while stack is not empty
    solve $\mathrm{LP}(X_k, Y_k)$ ... solution (x^k, y^k)
    if $\mathrm{LP}(X_k, Y_k)$ infeasible or $z_{\mathrm{LP}_k} \geq U - \epsilon$ **then**
        fathom node (case 1. or 3.)
    else
        solve $\mathrm{NLP}(X_k, Y_k)$ ... solution $(\hat{x}^k, \hat{y}^k)$
        if $z_{\mathrm{NLP}_k} < U - \epsilon$ & $\hat{y}^k$ integer **then**
            update $U := z_{\mathrm{NLP}_k}$ & incumbent $(x^*, y^*) := (\hat{x}^k, \hat{y}^k)$
        if $|z_{\mathrm{NLP}_k} - z_{\mathrm{LP}_k}| \leq \epsilon$ **then**
            fathom node (case 2.)
        else
            branch creating two new LPs

Theorem: If $(x, y) \in P$ is bounded $\Rightarrow$ get ϵ-optimal solution.

Overview

1. Applications & Motivation
   - Tertiary Voltage Control
   - Blackout Prevention in National Power Grid
   - Problem Characteristics
2. Branch-and-Refine for Nonconvex MINLPs
   - Decomposition of Nonlinear Functions
   - Piecewise Polyhedral Envelopes
   - Branch-and-Refine
3. Preliminary Numerical Results
   - Preliminary Numerical Experience
   - Conclusions & Future Research

Test Problems (Generic)

prob	#var	#cons	#var OA	#cons OA	#sets λ	#disc
pb0	4	2	44	32	6	1
pb1	4	2	44	32	6	1
pb2	6	2	41	30	5	1
pb3	6	2	41	30	5	1
pb4	12	4	97	71	11	2
pb5	12	4	97	71	11	2
pb6	12	4	143	97	19	3
pb7	12	4	143	97	19	3
pb8	12	4	119	77	14	2
pb9	12	4	119	77	14	2
pb10	10	4	111	72	13	2
pb11	10	4	111	72	13	2
pb12	24	8	275	187	40	6
pb13	24	8	275	187	40	6

Test Problems (Tertiary Voltage Control)

prob	#var	#cons	#var OA	#cons OA	#sets λ	#disc
TVC1	16	9	269	200	39	6
TVC2	18	9	275	204	40	6
TVC3	27	15	422	315	61	9
TVC4	27	15	422	315	61	9
TVC5	37	21	602	449	87	13
TVC6	38	21	635	472	92	14

... moderately sized problems

Complexity of nonconvex MINLPs depends on
terms in computational graph $\simeq$ #sets λ

Do We Need Global Solvers?

Comparison with NLP solvers

solver	# Problems Solved	# Global Solutions
BnR	20	20
Filter	12	8
IPOPT	17	14
KNITRO	17	13

Comparison with MINLP solvers

solver	# Problems Solved	# Global Solutions
BnR	20	20
BONMIN	15	11
MINLPBB	11	9

Implementation Details & Tricks

- LPs solved with CPLEX
- decomposition hand-coded by Emilie (yikes!)
  - exploit common sub-expressions
  - can be automated, similar to automatic differentiation (AD)
- NLPs solved with FilterSQP (AD for gradients/Hessian)
- propagate & strengthen bounds through computational graph
- pre-solve (LP) to reduce range of variables (like BARON)
  - adaptive presolve is best: tail-off factor
- pseudo-cost branching (generalized to nonconvex)
- best-estimate node selection (generalized to nonconvex)

Numerical Results (# LPs solved)

prob	basic	+presolve	+var-select	+node-select
pb0	63	63	68	68
pb1	133	131	79	68
pb2	2115	3237	194	260
pb3	135	197	121	97
pb4	15389	11388	120	120
pb5	3009	257	145	145
pb6	65800	6145	348	292
pb7	377	1353	1235	1121
pb8	fail	198817	263	241
pb9	62149	33668	442	442
pb10	113846	51816	205	197
pb11	3806	7349	558	258
pb12	fail	33407	1503	1056
pb13	fail	8093	17388	3885

Numerical Results (# LPs solved)

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

Conclusions & Future Research

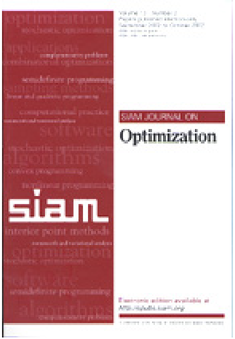
Branch-and-Refine

- three key ingredients:
  1. decompose functions into 1D and 2D components
  2. construct piecewise polyhedral envelope
  3. branch on variables not on SOS condition
- favorable theoretical properties (see Emilie's thesis)
- encouraging numerical results

Future Work

- exploit expression tree ... use AD tricks for better OA
- what decomposition is best ... non-unique
- avoid λ variables ... work with OA directly
- avoid SOS for convex functions ... NLP subproblems
- efficient implementation & support for `AMPL`, `GAMS`

Two Shameless Commercials



Submit your papers to SIOPT!

- electronic submission
- fast review
- top impact factor

... looking for MIP papers

IMA
Institute for Mathematics and Its Applications

IMA Hot Topics Workshop (November 17-21, 2008)

Mixed-Integer Nonlinear Optimization:
Algorithmic Advances and Applications