Aspects on solving convex and nonconvex MINLP problems

TAPIO WESTERLUND

CENTER OF EXCELLENCE IN
OPTIMIZATION AND SYSTEMS ENGINEERING
ÅBO AKADEMI UNIVERSITY, FINLAND

MINLP 2014 WORKSHOP
CARNEGIE MELLON UNIVERSITY, PITTSBURGH, USA
JUNE 4, 2014

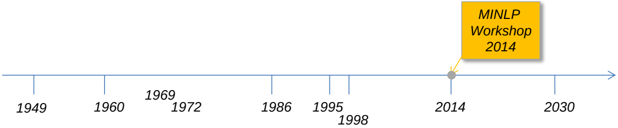
OSE
OPTIMIZATION AND SYSTEMS ENGINEERING

Åbo Akademi University

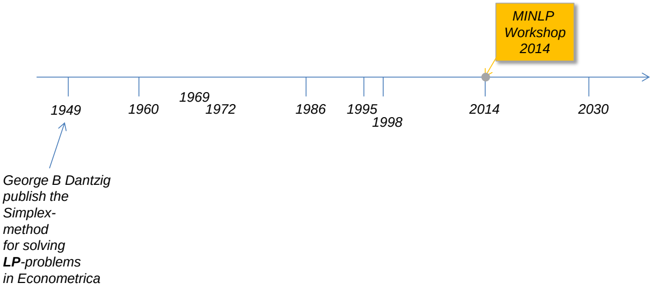
Contents

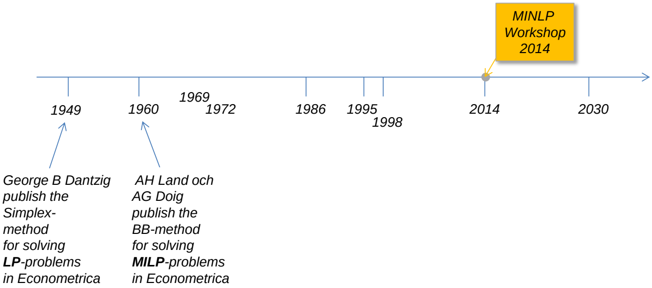
1. Introduction – a short background to MINLP

A short background to MINLP



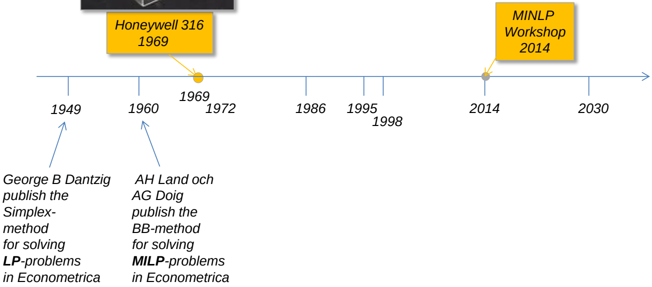
Some mile stones


MINLP Workshop 2014
1949
1960
1969
1972
1986
1995
1998
2014
2030
George B Dantzig publish the Simplex-method for solving **LP**-problems in Econometrica

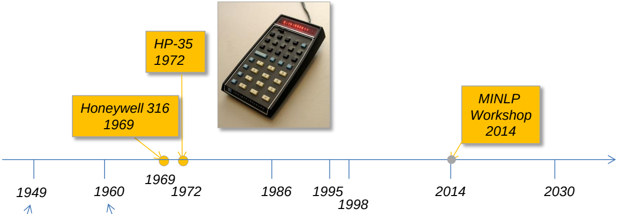

MINLP Workshop 2014
1949
1960
1969
1972
1986
1995
1998
2014
2030
George B Dantzig publish the Simplex-method for solving **LP**-problems in Econometrica
AH Land och AG Doig publish the BB-method for solving **MILP**-problems in Econometrica

(Typical laptop 2014: 16GB RAM & 64-bit operating system)

H-316, 16kB primary memory

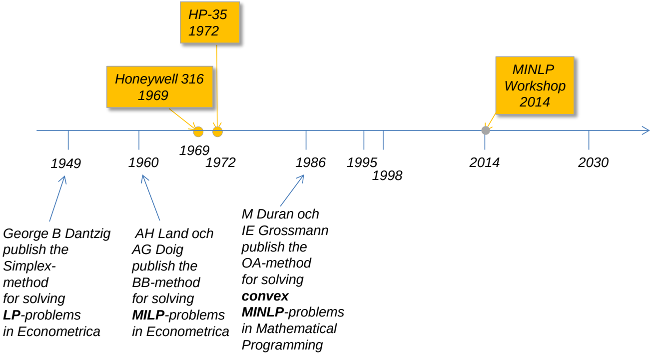


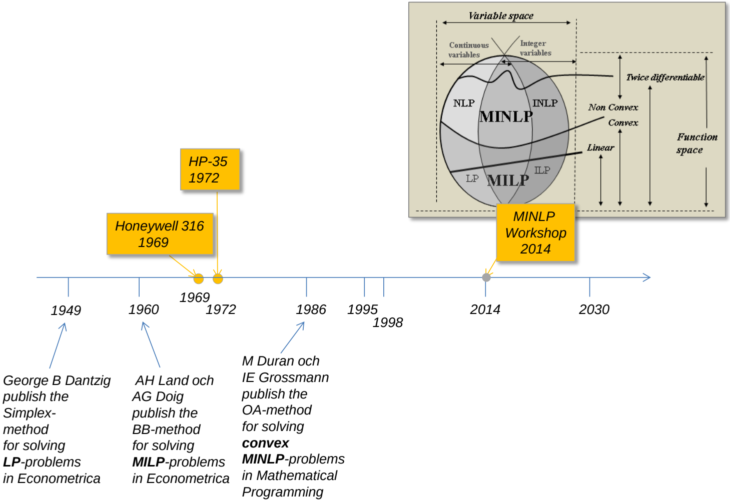
HP-35 (1 memory)
for sale between 1972-1975
as the first
"scientific pocket calculator"
in the world

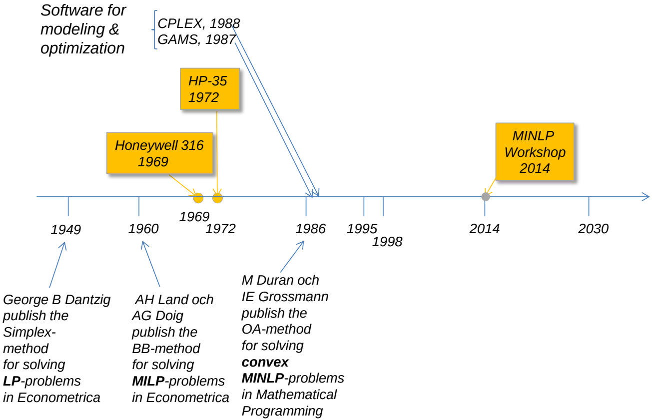


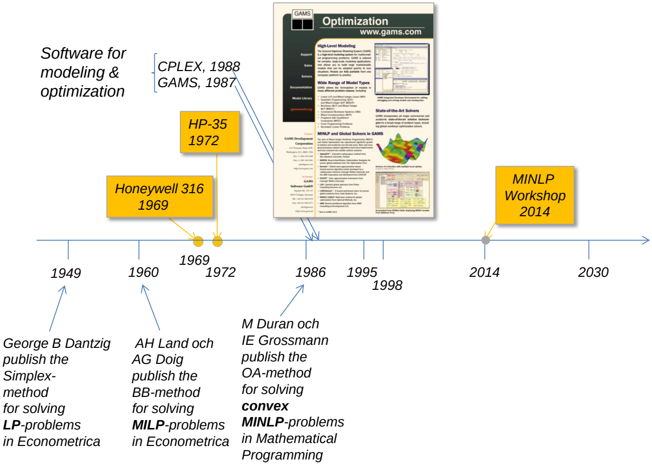
George B Dantzig
publish the
Simplex-
method
for solving
***LP**-problems*
in Econometrica

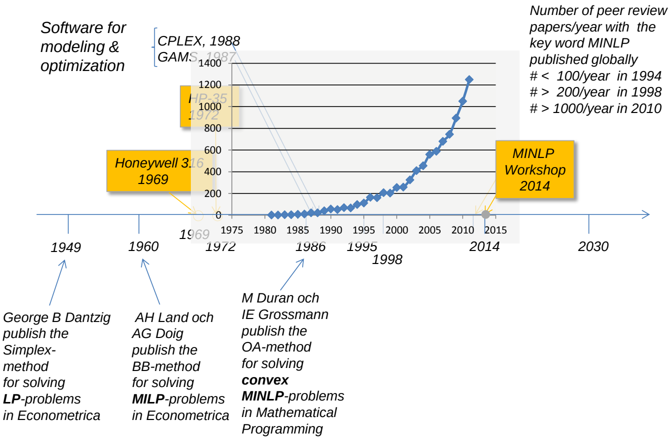
AH Land och
AG Doig
publish the
BB-method
for solving
***MILP**-problems*
in Econometrica


HP-35
1972
Honeywell 316
1969
MINLP
Workshop
2014
1949
1960
1969
1972
1986
1995
1998
2014
2030
George B Dantzig
publish the
Simplex-
method
for solving
LP-problems
in Econometrica
AH Land och
AG Doig
publish the
BB-method
for solving
MILP-problems
in Econometrica
M Duran och
IE Grossmann
publish the
OA-method
for solving
convex
MINLP-problems
in Mathematical
Programming

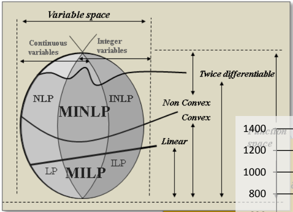

Variable space
Continuous variables
Integer variables
NLP
INLP
MINLP
LP
MILP
ILP
Twice differentiable
Non Convex
Convex
Linear
Function space
HP-35 1972
Honeywell 316 1969
MINLP Workshop 2014
1949
1960
1969
1972
1986
1995
1998
2014
2030
George B Dantzig publish the Simplex-method for solving **LP**-problems in Econometrica
AH Land och AG Doig publish the BB-method for solving **MILP**-problems in Econometrica
M Duran och IE Grossmann publish the OA-method for solving **convex** **MINLP**-problems in Mathematical Programming


Software for modeling & optimization
CPLEX, 1988
GAMS, 1987
HP-35 1972
Honeywell 316 1969
MINLP Workshop 2014
1949
1960
1969
1972
1986
1995
1998
2014
2030
George B Dantzig publish the Simplex-method for solving **LP**-problems in Econometrica
AH Land och AG Doig publish the BB-method for solving **MILP**-problems in Econometrica
M Duran och IE Grossmann publish the OA-method for solving **convex MINLP**-problems in Mathematical Programming

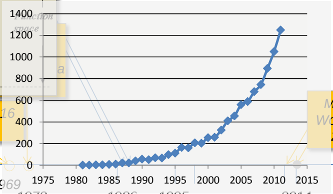

Software for modeling & optimization
CPLEX, 1988
GAMS, 1987
GAMS
Optimization
www.gams.com
HP-35
1972
Honeywell 316
1969
MINLP Workshop 2014
1949
1960
1969
1972
1986
1995
1998
2014
2030
George B Dantzig publish the Simplex-method for solving **LP**-problems in Econometrica
AH Land och AG Doig publish the BB-method for solving **MILP**-problems in Econometrica
M Duran och IE Grossmann publish the OA-method for solving **convex MINLP**-problems in Mathematical Programming



Software for modeling & optimization
CPLEX, 1988
GAMS, 1987
HP-35 1972
Honeywell 316 1969
Number of peer review papers/year with the key word MINLP published globally
< 100/year in 1994
> 200/year in 1998
> 1000/year in 2010
MINLP Workshop 2014
1400
1200
1000
800
600
400
200
0
1975 1980 1985 1990 1995 2000 2005 2010 2015
1949
1960
1969
1972
1986
1995
1998
2014
2030
George B Dantzig publish the Simplex-method for solving **LP**-problems in Econometrica
AH Land och AG Doig publish the BB-method for solving **MILP**-problems in Econometrica
M Duran och IE Grossmann publish the OA-method for solving **convex** **MINLP**-problems in Mathematical Programming



Honeywell 316
1969



Number of peer review papers/year with the key word MINLP published globally
 # < 100/year in 1994
 # > 200/year in 1998
 # > 1000/year in 2010

MINLP Workshop
2014

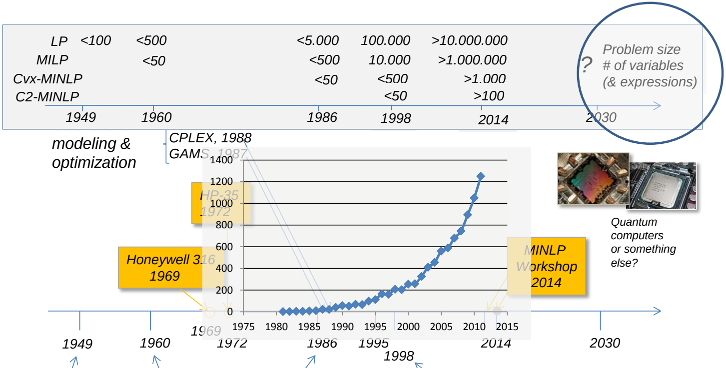
1949 *1960* *1969* *1972* *1986* *1995* *1998* *2014* *2030*

*George B Dantzig publish the Simplex-method for solving **LP**-problems in Econometrica*

*AH Land och AG Doig publish the BB-method for solving **MILP**-problems in Econometrica*

*M Duran och IE Grossmann publish the OA-method for solving **convex** **MINLP**-problems in Mathematical Programming*

*CS Adjiman, S Dallwig och CA Floudas publish the α-BB method for solving **C2** **MINLP**-problems in Computers & Chem. Engng.*

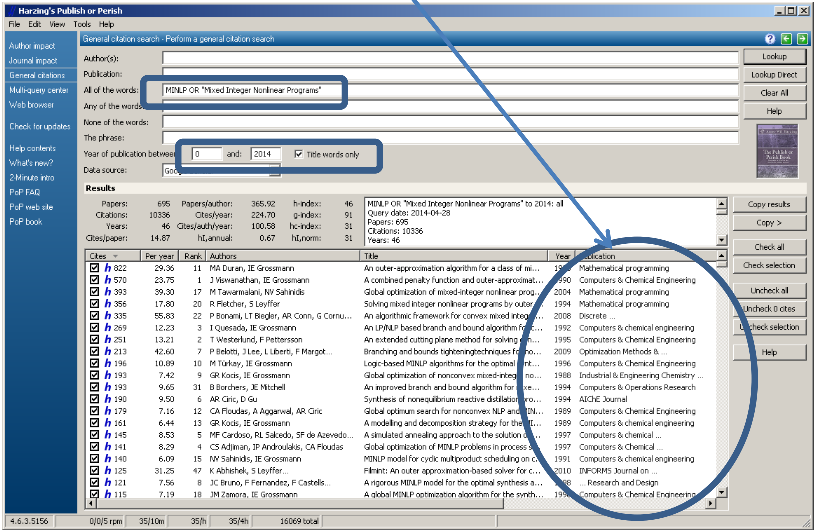


*George B Dantzig publish the Simplex-method for solving **LP**-problems in Econometrica*

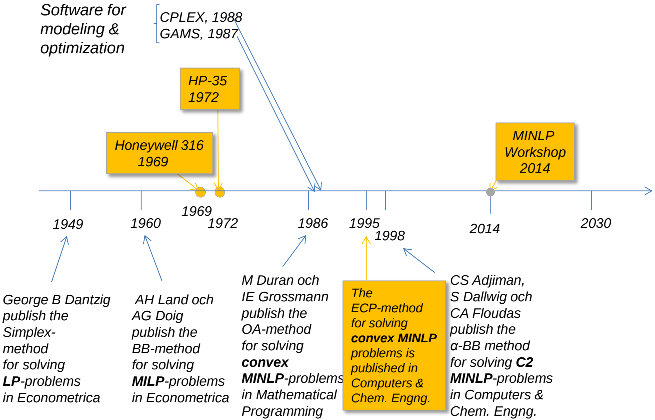
*AH Land och AG Doig publish the BB-method for solving **MILP**-problems in Econometrica*

*M Duran och IE Grossmann publish the OA-method for solving **convex MINLP**-problems in Mathematical Programming*

*CS Adjiman, S Dallwig och CA Floudas publish the α-BB method for solving **C2 MINLP**-problems in Computers & Chem. Engng.*



Harzing's Publish or Perish
File Edit View Tools Help
Author impact
Journal impact
General citations
Multi-query center
Web browser
Check for updates
Help contents
What's new?
2-Minute intro
PoP FAQ
PoP web site
PoP book
General citation search - Perform a general citation search
Author(s):
Publication:
All of the words: MINLP OR "Mixed Integer Nonlinear Programs"
Any of the words:
None of the words:
The phrase:
Year of publication between 0 and: 2014 Title words only
Data source:
Lookup
Lookup Direct
Clear All
Help
Results
Papers: 695 Papers/author: 365.92 h-index: 46
Citations: 10336 Cites/year: 224.70 g-index: 91
Years: 46 Cites/auth/year: 100.58 hc-index: 31
Cites/paper: 14.87 hI,annual: 0.67 hI,norm: 31
MINLP OR "Mixed Integer Nonlinear Programs" to 2014: all
Query date: 2014-04-28
Papers: 695
Citations: 10336
Years: 46
Copy results
Copy >
Check all
Check selection
Uncheck all
Uncheck 0 cites
Help
Cites Per year Rank Authors Title Year Publication
822 29.36 11 MA Duran, IE Grossmann An outer-approximation algorithm for a class of mi... Mathematical programming
570 23.75 1 J Viswanathan, IE Grossmann A combined penalty function and outer-approximat... Computers & Chemical Engineering
393 39.30 17 M Tawarmalani, NV Sahinidis Global optimization of mixed-integer nonlinear prog... 2004 Mathematical programming
356 17.80 20 R Fletcher, S Leyffer Solving mixed integer nonlinear programs by outer 1994 Mathematical programming
335 55.83 22 P Bonami, LT Biegler, AR Conn, G Cornu... An algorithmic framework for convex mixed integ... 2008 Discrete ...
269 12.23 3 I Quesada, IE Grossmann An LP/NLP based branch and bound algorithm f... 1992 Computers & chemical engineering
251 13.21 2 T Westerlund, F Pettersson An extended cutting plane method for solving ... 1995 Computers & Chemical Engineering
213 42.60 7 P Belotti, J Lee, L Liberti, F Margot... Branching and bounds tighteningtechniques f... 2009 Optimization Methods & ...
196 10.89 10 M Türkay, IE Grossmann Logic-based MINLP algorithms for the optimal ... 1996 Computers & Chemical Engineering
193 7.42 9 GR Kocis, IE Grossmann Global optimization of nonconvex mixed-integ... 1988 Industrial & Engineering Chemistry ...
193 9.65 31 B Borchers, JE Mitchell An improved branch and bound algorithm for ... 1994 Computers & Operations Research
190 9.50 6 AR Ciric, D Gu Synthesis of nonequilibrium reactive distillatio... 1994 AIChE Journal
179 7.16 12 CA Floudas, A Aggarwal, AR Ciric Global optimum search for nonconvex NLP and ... 1989 Computers & Chemical Engineering
161 6.44 13 GR Kocis, IE Grossmann A modelling and decomposition strategy for the ... 1989 Computers & chemical engineering
145 8.53 5 MF Cardoso, RL Salcedo, SF de Azevedo... A simulated annealing approach to the solution ... 1997 Computers & chemical ...
141 8.29 4 CS Adjiman, IP Androulakis, CA Floudas Global optimization of MINLP problems in process ... 1997 Computers & Chemical ...
140 6.09 15 NV Sahinidis, IE Grossmann MINLP model for cyclic multiproduct scheduling on ... 1991 Computers & chemical engineering
125 31.25 47 K Abhishek, S Leyffer... Filmint: An outer approximation-based solver for ... 2010 INFORMS Journal on ...
121 7.56 8 JC Bruno, F Fernandez, F Castells... A rigorous MINLP model for the optimal synthesis a... ... Research and Design
115 7.19 18 JM Zamora, IE Grossmann A global MINLP optimization algorithm for the synth... Computers & Chemical Engineering
4.6.3.5156 0/0/5 rpm 35/10m 35/h 35/4h 16069 total


Software for modeling & optimization
CPLEX, 1988
GAMS, 1987
HP-35 1972
Honeywell 316 1969
MINLP Workshop 2014
1949
1960
1969
1972
1986
1995
1998
2014
2030
George B Dantzig publish the Simplex-method for solving **LP**-problems in Econometrica
AH Land och AG Doig publish the BB-method for solving **MILP**-problems in Econometrica
M Duran och IE Grossmann publish the OA-method for solving **convex MINLP**-problems in Mathematical Programming
The ECP-method for solving **convex MINLP** problems is published in Computers & Chem. Engng.
CS Adjiman, S Dallwig och CA Floudas publish the α-BB method for solving **C2 MINLP**-problems in Computers & Chem. Engng.

Software for modeling & optimization

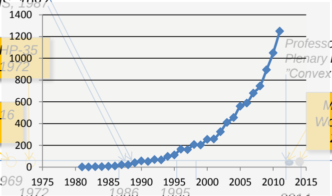
CPLEX, 1988
GAMS, 1987

HP-35 1972

Honeywell 316 1969

Professor Robert Weismantel, ETH, Zürich
Plenary lecture at ISMP-2012
"Convex MINLP starts to be a technology"

MINLP Workshop 2014



1949 *1960* *1969* *1972* *1986* *1995* *1998* *2014* *2030*

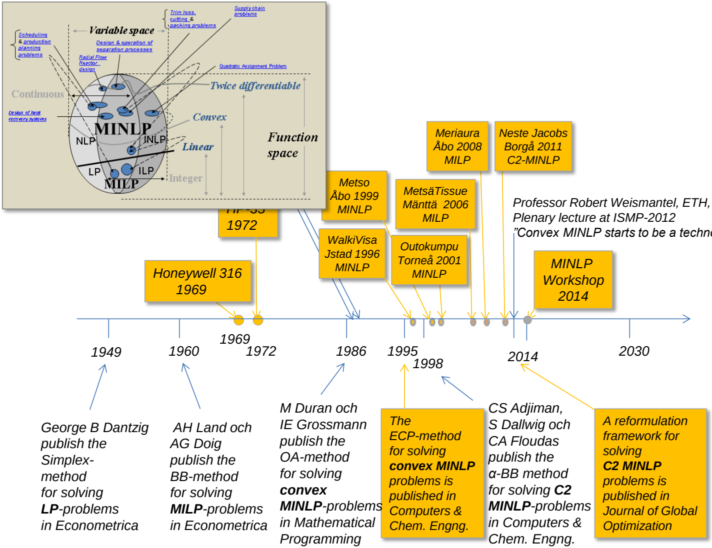
George B Dantzig publish the Simplex-method for solving ***LP****-problems in Econometrica*

AH Land och AG Doig publish the BB-method for solving ***MILP****-problems in Econometrica*

M Duran och IE Grossmann publish the OA-method for solving ***convex MINLP****-problems in Mathematical Programming*

The ECP-method for solving ***convex MINLP*** *problems is published in Computers & Chem. Engng.*

CS Adjiman, S Dallwig och CA Floudas publish the α-BB method for solving ***C2 MINLP****-problems in Computers & Chem. Engng.*


Scheduling & production planning problems
Variable space
Trim loss, cutting & packing problems
Supply chain problems
Design & operation of separation processes
Radial Flow Reactor design
Quadratic Assignment Problem
Continuous
Twice differentiable
Design of heat recovery systems
MINLP
NLP
INLP
Convex
Linear
Function space
LP
ILP
MILP
Integer
HP-35
1972
Metso Åbo 1999 MINLP
MetsäTissue Mänttä 2006 MILP
Meriaura Åbo 2008 MILP
Neste Jacobs Borgå 2011 C2-MINLP
Professor Robert Weismantel, ETH, Zürich
Plenary lecture at ISMP-2012
"Convex MINLP starts to be a technology"
WalkiVisa Jstad 1996 MINLP
Outokumpu Torneå 2001 MINLP
MINLP Workshop 2014
Honeywell 316 1969
1949
1960
1969
1972
1986
1995
1998
2014
2030
George B Dantzig publish the Simplex-method for solving **LP**-problems in Econometrica
AH Land och AG Doig publish the BB-method for solving **MILP**-problems in Econometrica
M Duran och IE Grossmann publish the OA-method for solving **convex MINLP**-problems in Mathematical Programming
The ECP-method for solving **convex MINLP** problems is published in Computers & Chem. Engng.
CS Adjiman, S Dallwig och CA Floudas publish the α-BB method for solving **C2 MINLP**-problems in Computers & Chem. Engng.
A reformulation framework for solving **C2 MINLP** problems is published in Journal of Global Optimization

2. Aspects on algorithms for convex MINLP problems

Convex functions

Problem (P1)

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & g(x) \leq 0, \end{array}$$

where f and g are convex functions.

Convex functions or convex sets

Problem (P1)

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & g(x) \leq 0, \end{array}$$

where f and g are convex functions.

Problem (P2)

$$\begin{array}{ll} \text{minimize} & f(x) \\ \text{subject to} & x \in C, \end{array}$$

where f is a convex function, $C = \{x | g(x) \leq 0\}$, and g are convex/quasiconvex functions.

Smooth or nonsmooth functions

- Does the convergence properties of a considered "convex MINLP" solver still hold true if the functions are not differentiable but convex/quasiconvex?

	convex	quasiconvex
smooth twice differentiable (C^2)	?	?
smooth once differentiable (C^1)	?	?
nonsmooth continuous	?	?
locally Lipschitz continuous	?	?

Nonsmooth functions in MINLP

Question: Is it possible to only replace gradients with subgradients in order to handle nonsmooth functions rigourously in algorithms for differentiable convex problems?

Answer: Not for all convex MINLP algorithms!

- Yes, *e.g.,* for ECP
- No, for certain versions of OA, *e.g.,* the linear OA[1]:

Algorithm 1 (Linear Outer Approximation).
Initialization: y^0 is given; set $i=0$, $T^{-1}=\emptyset$, $S^{-1}=\emptyset$ and UBD $=\infty$.
REPEAT
(1) Solve the subproblem NLP(y^i), or the feasibility problem F(y^i) if NLP(y^i) is infeasible, and let the solution be x^i.
(2) Linearize the objective and (active) constraint functions about (x^i, y^i). Set $T^i = T^{i-1} \cup \{i\}$ or $S^i = S^{i-1} \cup \{i\}$ as appropriate.
(3) IF (NLP(y^i) is feasible and $f^i <$ UBD) THEN
update current best point by setting $x^* = x^i$, $y^* = y^i$, UBD $= f^i$.
(4) Solve the current relaxation M^i of the master program M, giving a new integer assignment y^{i+1} to be tested in the algorithm. Set $i = i+1$.
UNTIL(M^i is infeasible).

[1] Fletcher, R. and Leyffer, S., Solving mixed integer nonlinear programs by outer approximation, Mathematical Programming 66, pp. 327–349, 1994.

A convex nonsmooth example where the gradient is replaced by a subgradient[2]

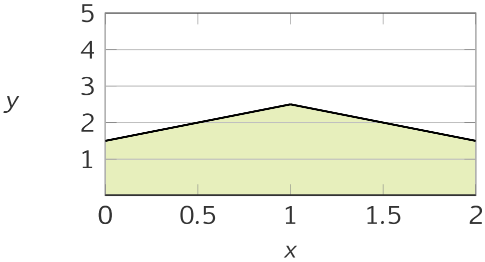
$$
\begin{array}{ll}
\text{minimize} & 2x - y \\
\text{subject to} & g(x,y) \leq 0 \\
& y - 4x - 1 \leq 0 \\
& 0 \leq x \leq 2, \quad y \in Y = \{0,1,2,3,4,5\},
\end{array}
\tag{E}
$$

where

$$
g(x,y) = \max\left\{-\frac{3}{2} - x + y,\ -\frac{7}{2} + y + x\right\}.
$$

[2] Eronen, V.-P., Mäkelä, M. M. and Westerlund, T., On the generalization of ECP and OA methods to nonsmooth convex MINLP problems, Optimization, pp. 1–17, iFirst, available online, 2012.

Solving with the linear outer approximation



Initialization: $y^0 = 3$

Step 1: Solve the subproblem NLP(y^0) or the feasibility problem $F(y^0)$ if NLP(y^0) is infeasible, and let the solution be x^0.

- There are no feasible points in the problem NLP(y^0), thus the feasibility problem F_{y^0} will be solved:

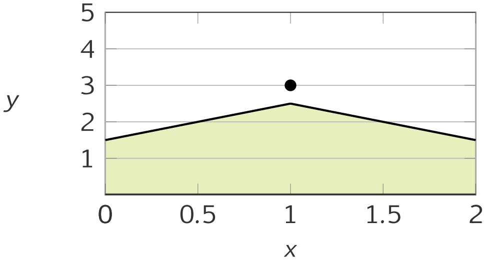
$$\begin{aligned} &\text{minimize} && \mu \\ &\text{subject to} && \max\left\{\frac{3}{2} - x,\ -\frac{1}{2} + x\right\} \le \mu \\ &&& 2 - 4x \le 0 \\ &&& 0 \le x \le 2. \end{aligned} \qquad (F_{y^0})$$

- The solution of F_{y^0} is $x^0 = 1$ with $\mu = 1/2$.

Step 2: Linearize g at the point $(x^0, y^0) = (1,3)$ for the next relaxed MILP master problem M^0.

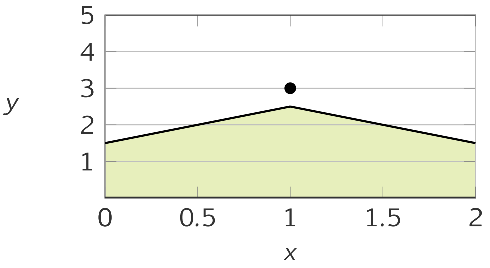
- Both the functions $-3/2 - x + y$ and $-7/2 + y + x$ have the same value $1/2$ at the point (x^0, y^0) and thus the subdifferential is

$$\partial g(1,3) = \left\{(\alpha, 1)^T | \alpha \in [-1,1]\right\}. \qquad (1)$$



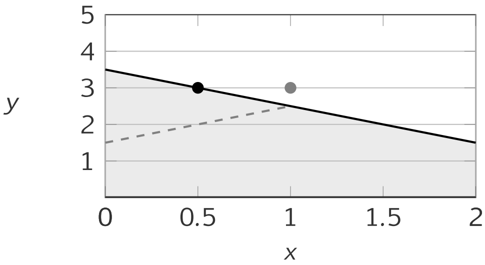
- Since we may select an abitrary subgradient we may choose, *e.g.*, $\xi(x^0, y^0) = (1,1)^T$. Thus the new linear constraint is

$$\frac{1}{2} + (1,1)(x-1, y-3)^T \leq 0 \quad \Rightarrow \quad x + y - \frac{7}{2} \leq 0.$$



Step 3: Update the current best point if NLP(y^0) is feasible, but since NLP(y^0) was not feasible go to Step 4.

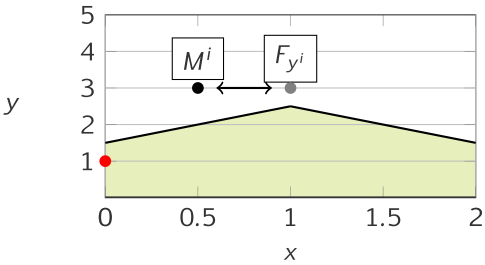
Step 4: Create and solve the current relaxation M^0 of the master program giving a new integer assignment y^1.



$$
\begin{array}{ll}
\text{minimize} & 2x - y \\
\text{subject to} & x + y - 7/2 \leq 0 \\
& y - 4x - 1 \leq 0 \\
& 0 \leq x \leq 2, \quad y \in Y.
\end{array}
\qquad (M^0)
$$

- The solution point of (M^0) is $(1/2, 3)$. Set $i = i + 1$, $y^1 = 3$.

Repeat steps 1–4: Until M^i is infeasible.



- Hence $y^1 = y^0$ and $F_{y^1} \equiv F_{y^0}$. Thus LOA may generate an infinite loop between points $(1, 3)$ and $(1/2, 3)$.
- Both of them are infeasible but the problem (E) has a feasible point $(0, 1)$ for example, where the objective function $2x - y$ has the value -1.

3. A new algorithm for solving convex MINLP problems

- A new interior point based algorithm for solving convex MINLP problems to global optimality is introduced.
- Roots:
  - Kelley's cutting plane algorithm 1960[3]
  - The extended cutting plane (ECP) algorithm 1995[4]
- Cutting planes are replaced with supporting hyperplanes using a line search procedure.
- Two LP preprocessing steps are utilized to quickly get a tight linear relaxation of the part of the feasible region defined by the convex/quasiconvex constraints.
- An interior point is required for the line search.

---

[3] Kelley, Jr., J., The cutting-plane method for solving convex programs, Journal of the SIAM, vol. 8(4), pp. 703–712, 1960.

[4] Westerlund, T. and Pettersson, F., An extended cutting plane method for solving convex MINLP problems, Computers & Chemical Engineering 19, pp. 131–136, 1995.

The MINLP problem

- The algorithm finds the optimal solution x^* to the following convex MINLP problem:

$$\boxed{x^* = \underset{x \in C \cap L \cap Y}{\arg\min}\, c^T x} \tag{P}$$

where $x = [x_1, x_2, \ldots, x_N]^T$ belongs to the compact set

$$X = \left\{ x \,\middle|\, \underline{x}_i \leq x_i \leq \overline{x}_i,\ i = 1, \ldots, N \right\} \subset \mathbb{R}^n,$$

the feasible region is defined by $C \cap L \cap Y$,

$$\begin{aligned} C &= \{x \,|\, g_m(x) \leq 0,\ m = 1, \ldots, M,\ x \in X\}, \\ L &= \{x \,|\, Ax \leq a,\ Bx = b,\ x \in X\}, \\ Y &= \{x \,|\, x_i \in \mathbb{Z},\ i \in I_{\mathbb{Z}},\ x \in X\}, \end{aligned}$$

and C is a convex set.

Steps in the interior point supporting hyperplane algorithm

NLP: If an interior point is not given, obtain a feasible, relaxed interior point (satisfying C) by solving a NLP problem.

LP1: Solve simple LP problems (initially in X) and conduct a line search procedure to obtain supporting hyperplanes giving a first linear relaxation of the convex set C. Optional.

LP2: Continue with a corresponding procedure as in LP1 but now also including the linear constraints in L. Optional.

MILP: Finally include the integer requirements and solve MILP problems using a corresponding procedure to find the optimal solution to (P).

NLP-step

- A point in C is required as an endpoint for the line searches to be conducted in the LP1-, LP2- and MILP-steps.
- Assuming that (P) has a solution, the internal point can be obtained from the following NLP problem:

$$\boxed{\begin{aligned} \tilde{x}_{\text{NLP}} &= \underset{x \in X}{\arg\min}\, F(x), \\ \text{where } F(x) &:= \max_{m=1,\ldots,M} \{g_m(x)\}. \end{aligned}} \tag{P-NLP}$$

- F is convex/quasiconvex since it is the maximum of convex/quasiconvex functions.
- (P-NLP) may be nonsmooth (if $M > 1$) even if g_m is smooth.
- The point $\tilde{x}_{\text{NLP}}$ need not be optimal but then fulfill $F(\tilde{x}_{\text{NLP}}) < 0$.
- Can be solved, *e.g.*, with the accelerated gradient method in[5].

[5] Nestorov, Y., Introductory lectures on convex optimization: A basic course, Kluwer Academic Publishers, 2004.

LP1-step

- Starting from $k=1$, $\Omega_0 = X$, the problem

$$\boxed{\tilde{x}_{\mathrm{LP}}^k = \operatorname*{arg\,min}_{\Omega_{k-1}} \; c^T x} \qquad \text{(P-LP1)}$$

  is repeatedly solved, and supporting hyperplanes (SHs)

$$l_k := F(x^k) + \xi_F(x^k)^T (x - x^k) \leq 0$$

  are generated and added to Ω_k. The point x^k is obtained by a line search for $F(x^k) = 0$ between the internal point $\tilde{x}_{\mathrm{NLP}}$ and the solution point to (P-LP1) $\tilde{x}_{\mathrm{LP}}^k$:

$$x^k = \lambda \tilde{x}_{\mathrm{NLP}} + (1-\lambda)\tilde{x}_{\mathrm{LP}}^k, \quad \lambda \in [0,1].$$

  $\xi_F(x^k)^T$ is a gradient or subgradient of F at x^k.
- If not $F(\tilde{x}_{\mathrm{LP}}^k) < \epsilon_{\mathrm{LP1}}$ or a maximum number of SHs have been generated, then k is increased and (P-LP1) resolved.

LP2-step

- This step is otherwise identical to LP1, with the exception that the linear constraints in L are now also included, *i.e.,*

$$\boxed{\tilde{x}_{\mathrm{LP}}^{k} = \operatorname*{arg\,min}_{\Omega_{k-1} \cap L} \; c^{T} x} \tag{P-LP2}$$

- (P-LP2) is repeatedly solved until $F(\tilde{x}_{\mathrm{LP}}^{k}) < \epsilon_{\mathrm{LP2}}$ or a maximum number of SHs have additionally been generated.

MILP-step

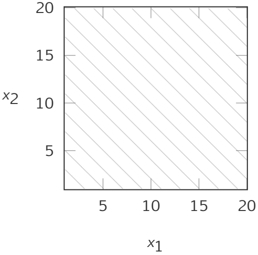
- Finally, in order to also fulfill the integer requirements of problem (P), a MILP step is performed.
- This step is otherwise identical to LP2, with the exception that the integer requirements in Y are now additionally considered, *i.e.,*

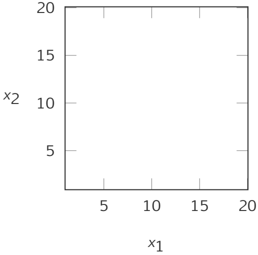
$$\boxed{\tilde{x}^k_{\text{MILP}} = \underset{\Omega_{k-1} \cap L \cap Y}{\arg\min} \; c^T x.} \tag{P-MILP}$$

- (P-MILP) is repeatedly solved until $F(\tilde{x}^k_{\text{MILP}}) < \epsilon_{\text{MILP}}$.
- Intermediate (P-MILP) problems do not need to be solved to optimality, but in order to guarantee an optimal solution of (P), the final MILP solution must be optimal.

An example

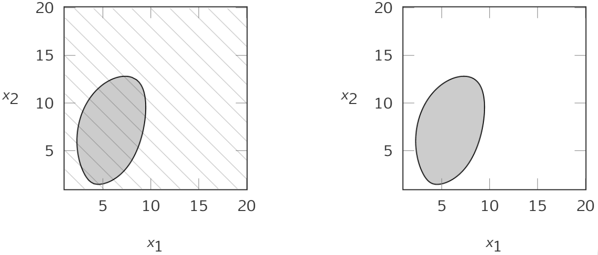
$$
\begin{aligned}
\text{minimize} \quad & c^T x = -x_1 - x_2 \\
\text{subject to} \quad & 1/x_1 + 1/x2 - x_1^{0.5} x_2^{0.5} + 4 \leq 0 \\
& 0.15(x_1 - 8)^2 + 0.1(x_2 - 6)^2 + 0.025 e^{x_1} x_2^{-3} - 5 \leq 0 \\
& 2x_1 - 3x_2 - 2 \leq 0 \\
& 1 \leq x_1 \leq 20, \quad 1 \leq x_2 \leq 20, \quad x_1 \in \mathbb{R}, \quad x_2 \in \mathbb{Z}.
\end{aligned}
$$





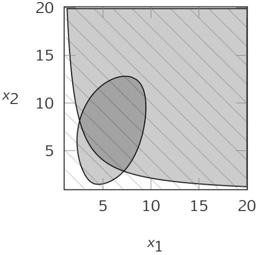
An example

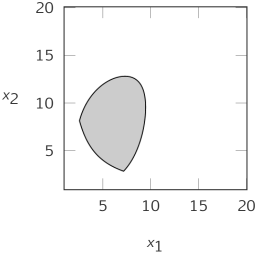
$$\begin{aligned}
\text{minimize}\quad & c^T x = -x_1 - x_2 \\
\text{subject to}\quad & 1/x_1 + 1/x2 - x_1^{0.5} x_2^{0.5} + 4 \le 0 \\
& 0.15(x_1-8)^2 + 0.1(x_2-6)^2 + 0.025 e^{x_1} x_2^{-3} - 5 \le 0 \\
& 2x_1 - 3x_2 - 2 \le 0 \\
& 1 \le x_1 \le 20, \quad 1 \le x_2 \le 20, \quad x_1 \in \mathbb{R}, \quad x_2 \in \mathbb{Z}.
\end{aligned}$$



An example

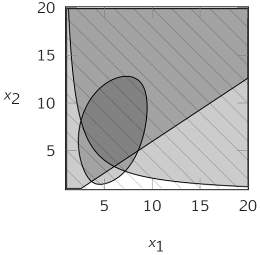
$$
\begin{aligned}
\text{minimize} \quad & c^T x = -x_1 - x_2 \\
\text{subject to} \quad & 1/x_1 + 1/x2 - x_1^{0.5} x_2^{0.5} + 4 \leq 0 \\
& 0.15(x_1 - 8)^2 + 0.1(x_2 - 6)^2 + 0.025 e^{x_1} x_2^{-3} - 5 \leq 0 \\
& 2x_1 - 3x_2 - 2 \leq 0 \\
& 1 \leq x_1 \leq 20, \quad 1 \leq x_2 \leq 20, \quad x_1 \in \mathbb{R}, \quad x_2 \in \mathbb{Z}.
\end{aligned}
$$

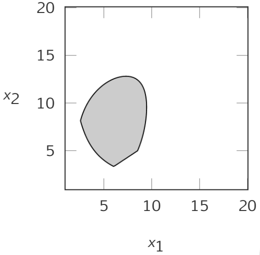




An example

$$
\begin{aligned}
\text{minimize} \quad & c^T x = -x_1 - x_2 \\
\text{subject to} \quad & 1/x_1 + 1/x2 - x_1^{0.5} x_2^{0.5} + 4 \leq 0 \\
& 0.15(x_1 - 8)^2 + 0.1(x_2 - 6)^2 + 0.025 e^{x_1} x_2^{-3} - 5 \leq 0 \\
& 2x_1 - 3x_2 - 2 \leq 0 \\
& 1 \leq x_1 \leq 20, \quad 1 \leq x_2 \leq 20, \quad x_1 \in \mathbb{R}, \quad x_2 \in \mathbb{Z}.
\end{aligned}
$$



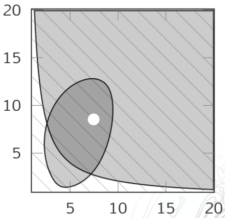


NLP step – find an interior point

$$\tilde{x}_{\mathrm{NLP}} = \operatorname*{arg\,min}_{(x_1,x_2)\in X} F(x_1,x_2),$$

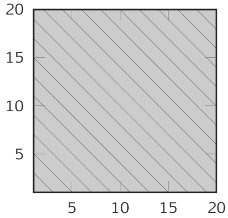
$$\text{where } F(x_1,x_2) := \max\{g_1(x_1,x_2),\ g_2(x_1,x_2)\}.$$

- The problem can be found using a suitable NLP solver.
- Not required to be the optimal point
- The optimal point here is $(7.45, 8.54)$



LP1 – Iteration 1

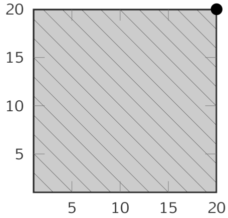
- Assume initially that $\Omega_0 = X$.



LP1 – Iteration 1

- Assume initially that $\Omega_0 = X$.
- $k = 1$, solve LP in Ω,

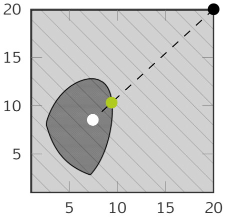
$$\tilde{x}_{\mathrm{LP}}^{k} = \underset{\Omega_{k-1}}{\arg\min}\; c^T x.$$



LP1 – Iteration 1

- Assume initially that $\Omega_0 = X$.
- $k = 1$, solve LP in Ω,

$$\tilde{x}_{\mathrm{LP}}^k = \arg\min_{\Omega_{k-1}} \ c^T x.$$



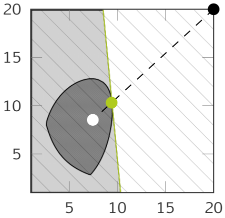
- Do line search

$$x^k = \lambda \tilde{x}_{\mathrm{NLP}} + (1-\lambda)\tilde{x}_{\mathrm{LP}}^k.$$

LP1 – Iteration 1

- Assume initially that $\Omega_0 = X$.
- $k = 1$, solve LP in Ω,

$$\tilde{x}_{\mathrm{LP}}^k = \underset{\Omega_{k-1}}{\arg\min}\ c^T x.$$



- Do line search

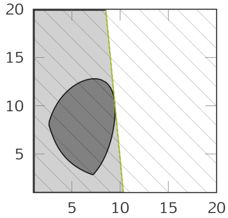
$$x^k = \lambda \tilde{x}_{\mathrm{NLP}} + (1-\lambda)\tilde{x}_{\mathrm{LP}}^k.$$

- Generate supporting hyperplane in x^k and add to Ω.

LP1 – Iteration 2

- $\Omega_1 = \{x | l_1(x) \leq 0,\ x \in X\}$.

  $$l_1(x) = 3.26x_1 + 0.313x_2 - 33.9$$



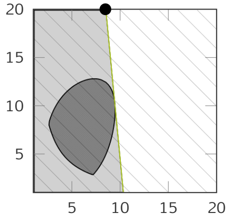
LP1 – Iteration 2

- $\Omega_1 = \{x | l_1(x) \leq 0,\ x \in X\}$.

$$l_1(x) = 3.26x_1 + 0.313x_2 - 33.9$$

- $k = 2$, solve LP in Ω,

$$\tilde{x}_{\mathrm{LP}}^k = \arg\min_{\Omega_{k-1}}\ c^T x.$$



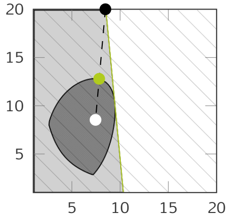
LP1 – Iteration 2

- $\Omega_1 = \{x | l_1(x) \leq 0,\ x \in X\}$.

$$l_1(x) = 3.26x_1 + 0.313x_2 - 33.9$$

- $k = 2$, solve LP in Ω,

$$\tilde{x}_{\mathrm{LP}}^k = \arg\min_{\Omega_{k-1}} \ c^T x.$$



- Do line search $x^k = \lambda \tilde{x}_{\mathrm{NLP}} + (1-\lambda)\tilde{x}_{\mathrm{LP}}^k$.

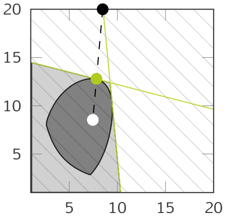
LP1 – Iteration 2

- $\Omega_1 = \{x | l_1(x) \leq 0,\ x \in X\}$.

$$l_1(x) = 3.26x_1 + 0.313x_2 - 33.9$$

- $k = 2$, solve LP in Ω,

$$\tilde{x}_{\text{LP}}^k = \arg\min_{\Omega_{k-1}}\ c^T x.$$

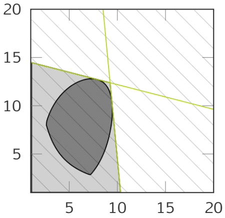


- Do line search $x^k = \lambda \tilde{x}_{\text{NLP}} + (1-\lambda)\tilde{x}_{\text{LP}}^k$.
- Generate supporting hyperplane in x^k and add to Ω.

LP1 – Iteration 3

- $\Omega_2 = \{x | l_j(x) \leq 0,\ j \in \{1,2\},\ x \in X\}$

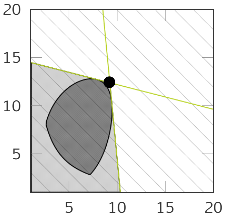
$$l_1(x) = 3.26x_1 + 0.313x_2 - 33.9$$
$$l_2(x) = 0.332x_1 + 1.30x_2 - 19.2$$



LP1 – Iteration 3

- $\Omega_2 = \{x | l_j(x) \leq 0,\ j \in \{1,2\},\ x \in X\}$

$$l_1(x) = 3.26x_1 + 0.313x_2 - 33.9$$
$$l_2(x) = 0.332x_1 + 1.30x_2 - 19.2$$



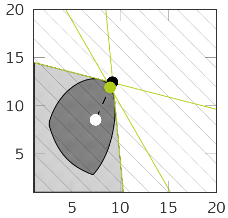
- $k = 3$, solve LP in Ω,

$$\tilde{x}^k_{\text{LP}} = \arg\min_{\Omega_{k-1}}\ c^T x.$$

LP1 – Iteration 3

- $\Omega_2 = \{x | l_j(x) \leq 0,\ j \in \{1,2\},\ x \in X\}$

$$l_1(x) = 3.26x_1 + 0.313x_2 - 33.9$$
$$l_2(x) = 0.332x_1 + 1.30x_2 - 19.2$$



- $k = 3$, solve LP in Ω,

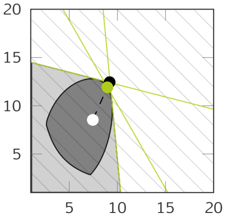
$$\tilde{x}_{\mathrm{LP}}^k = \arg\min_{\Omega_{k-1}}\ c^T x.$$

- Do line search, generate supporting hyperplane and add to Ω.

LP1 – Iteration 3

- $\Omega_2 = \{x | l_j(x) \leq 0,\ j \in \{1,2\},\ x \in X\}$

$$l_1(x) = 3.26x_1 + 0.313x_2 - 33.9$$
$$l_2(x) = 0.332x_1 + 1.30x_2 - 19.2$$



- $k = 3$, solve LP in Ω,

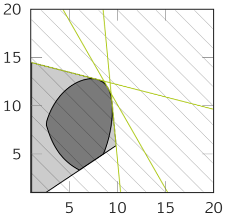
$$\tilde{x}_{\text{LP}}^k = \arg\min_{\Omega_{k-1}}\ c^T x.$$

- Do line search, generate supporting hyperplane and add to Ω.
- Terminate LP1-step since $F(\tilde{x}_{\text{LP}}^k) < \epsilon_{\text{LP1}}$.

LP2 – Iteration 4

- $\Omega_3 = \{x | l_j(x) \leq 0,\ j \in \{1,2,3\},\ x \in X\}$

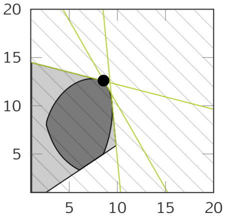
$$l_1(x) = 3.26x_1 + 0.313x_2 - 33.9$$
$$l_2(x) = 0.332x_1 + 1.30x_2 - 19.2$$
$$l_3(x) = 1.66x_1 + 0.951x_2 - 26.2$$



LP2 – Iteration 4

- $\Omega_3 = \{x | l_j(x) \leq 0,\ j \in \{1,2,3\},\ x \in X\}$

$$l_1(x) = 3.26x_1 + 0.313x_2 - 33.9$$
$$l_2(x) = 0.332x_1 + 1.30x_2 - 19.2$$
$$l_3(x) = 1.66x_1 + 0.951x_2 - 26.2$$



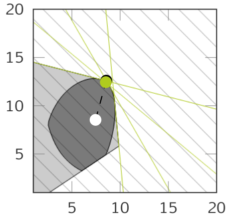
- $k = 4$, solve LP now in $\Omega \cap L$,

$$\tilde{x}_{\mathrm{LP}}^k = \arg\min_{\Omega_{k-1} \cap L}\ c^T x.$$

LP2 – Iteration 4

- $\Omega_3 = \{x | l_j(x) \leq 0,\ j \in \{1,2,3\},\ x \in X\}$

$$l_1(x) = 3.26x_1 + 0.313x_2 - 33.9$$
$$l_2(x) = 0.332x_1 + 1.30x_2 - 19.2$$
$$l_3(x) = 1.66x_1 + 0.951x_2 - 26.2$$



- $k = 4$, solve LP now in $\Omega \cap L$,

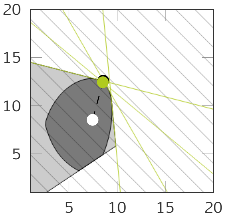
$$\tilde{x}_{\mathrm{LP}}^k = \arg\min_{\Omega_{k-1} \cap L}\ c^T x.$$

- Do line search, generate supporting hyperplane and add to Ω.

LP2 – Iteration 4

- $\Omega_3 = \{x | l_j(x) \leq 0,\ j \in \{1,2,3\},\ x \in X\}$

$$l_1(x) = 3.26x_1 + 0.313x_2 - 33.9$$
$$l_2(x) = 0.332x_1 + 1.30x_2 - 19.2$$
$$l_3(x) = 1.66x_1 + 0.951x_2 - 26.2$$

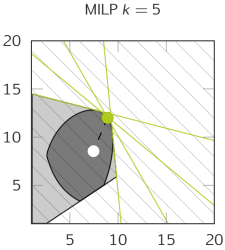


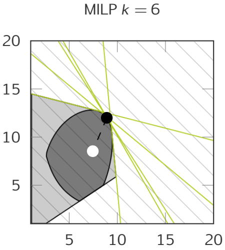
- $k = 4$, solve LP now in $\Omega \cap L$,

$$\tilde{x}^k_{\mathrm{LP}} = \arg\min_{\Omega_{k-1} \cap L}\ c^T x.$$

- Do line search, generate supporting hyperplane and add to Ω.
- Terminate LP2-step since $F(\tilde{x}^k_{\mathrm{LP}}) < \epsilon_{\mathrm{LP2}}$.

MILP step





- In this step the integer requirements in Y are also considered, *i.e.*, initially $k = 5, \Omega = \Omega_{k-1} \cap L \cap Y$.
- The MILP steps are required to guarantee an integer-feasible solution.

Solution and comparisons to other solvers

- Solving the MINLP problem with the supporting hyperplane algorithm gives the following solution

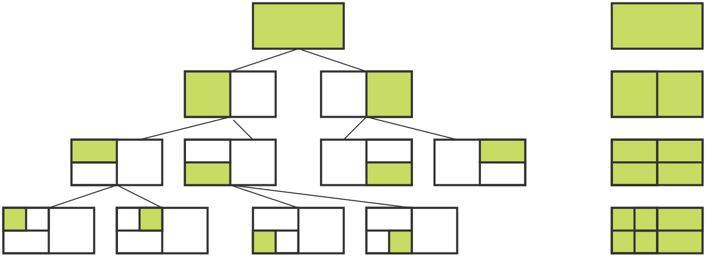
Type	Iteration	Obj. funct.	x_1	x_2	$F(x_1, x_2)$
LP1	1	−40.0000	20.0000	20.0000	30 359
LP1	2	−28.4720	8.47199	20.0000	14.9321
LP1	3	−21.6378	9.19722	12.4406	0.957382
LP2	4	−21.1639	8.56022	12.6037	0.229455
MILP	5	−20.9065	8.90647	12	0.00442134
MILP	6	−20.9036	8.90362	12	$4.22619 \cdot 10^{-6}$

- Solution times compared to some other MINLP solvers:

Solver	Iterations	Time (s)	Implementation
New algorithm	6	0.7	Prototype in Mathematica + CBC
ECP	21	1.5	GAMS 24.2 + CPLEX
DICOPT	11	1.5	GAMS 24.2 + CONOPT + CPLEX

4. Aspects on frameworks for nonconvex MINLP problems

Convex relaxation: branching vs reformulation



- **Branching:** n convex subproblems (the subproblems with the green domains are solved using a branching strategy)
- **Reformulation:** the entire nonconvex MINLP problem is reformulated to a convex relaxed MINLP problem solved sequentially.

Convex envelopes of functions or sets for tight convex relaxations

- Does a convex envelope $c(x) = \operatorname{conv} g(x)$ of a nonconvex function g in an inequality constraint $g(x) \leq 0$ give the tightest convex relaxation of $g(x) \leq 0$ when replacing it with $c(x) \leq 0$?

Convex relaxations and envelopes in literature

Tuy 1998

"A nonconvex inequality constraint $g(\mathbf{x}) \leq 0,\ \mathbf{x} \in X$, where X is a convex set in $\mathbb{R}^n$, can often be handled by replacing it with a convex inequality constraint $c(\mathbf{x}) \leq 0$ where $c(\mathbf{x})$ is a convex minorant of $g(\mathbf{x})$ on X. The latter inequality is then called a convex relaxation of the former. **Of course, the tightest relaxation is obtained when $c(\mathbf{x}) = \operatorname{conv} g(\mathbf{x})$, the convex envelope, *i.e.,* the largest convex minorant, of $g(\mathbf{x})$.**"

Let's see...

- Could it be possible to find some function q, other than $c(x) = \text{conv}\, g(x)$, with the property:

$$N \subset C_q \subset C_c,$$

where

$$\begin{aligned} N &= \{x | g(x) \leq 0\} \\ C_q &= \{x | q(x) \leq 0\} \\ C_c &= \{x | c(x) \leq 0\} \end{aligned}$$

for all $x \in X$ such that C_q would still be a convex set?

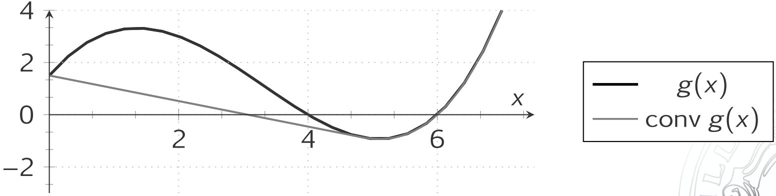
The convex envelope of a function

Consider the function

$$g(x) = 0.00506x^4 + 0.09553x^3 - 1.2774x^2 + 2.8821x + 1.5.$$

The convex envelope of the nonconvex function $g(x)$ on the interval $[0,7]$ is given by

$$\text{conv } g(x) = \begin{cases} -0.488764x + 1.5 & \text{if } 0 \leq x \leq 4.8312, \\ g(x) & \text{if } 4.8312 < x \leq 7. \end{cases}$$



The αBB underestimator, Floudas (2000)

Convex underestimator for twice-differentiable functions

A function $g(\mathbf{x}) \in C^2$ has the convex underestimator

$$\boxed{\hat{g}(\mathbf{x}) = g(\mathbf{x}) + \sum_i \alpha(\underline{x}_i - x_i)(\overline{x}_i - x_i)}$$

for $x_i \in [\underline{x}_i, \overline{x}_i]\ \forall i$ if and only if the parameter α fulfills

$$\alpha \geq \max\left\{0, -\tfrac{1}{2}\min_i \lambda_i\right\}$$

where the λ_i's are the eigenvalues of the Hessian of $g(\mathbf{x})$ on the interval $[\underline{x}_i, \overline{x}_i]$.
Different methods for calculating the α-values are available, *e.g.,* the scaled Gerschgorin method.

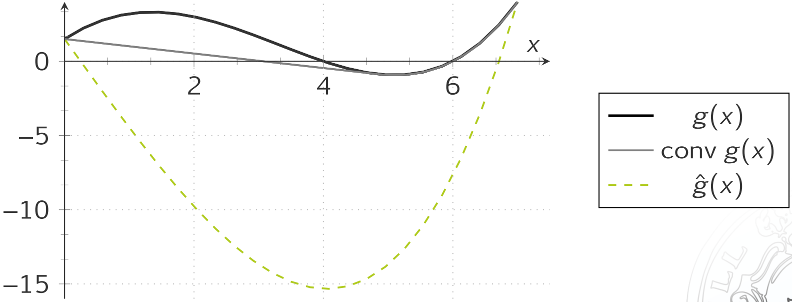
The αBB underestimator, illustration

- For example for the function

$$g(x) = 0.00506x^4 + 0.09553x^3 - 1.2774x^2 + 2.8821x + 1.5,$$

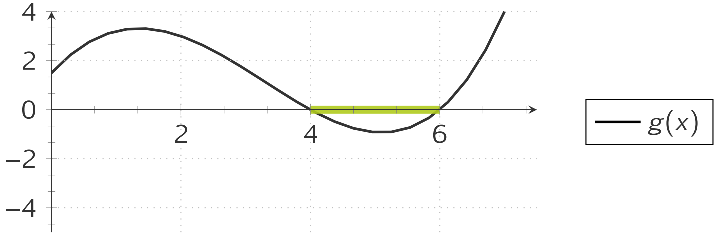
where $0 \leq x \leq 7$, the αBB underestimator becomes

$$\hat{g}(x) = g(x) + 1.2774(0 - x)(7 - x).$$



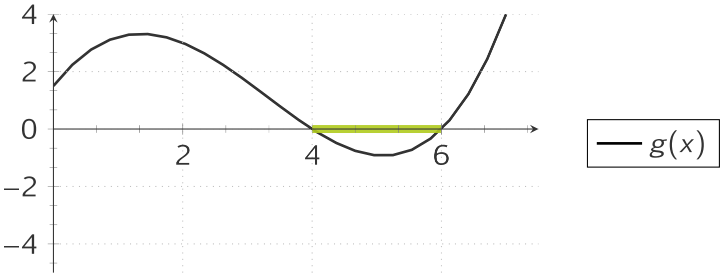
Convex envelope of the level set

- Observe that the convex envelope of a function $g(x)$ is the tightest convex relaxation of the function in question, but does not generally give the tightest convex relaxation of a level set $L = \{x \mid g(x) \leq \alpha\}$ (in this case $\alpha = 0$).



- The tightest convex relaxation of L is conv L, *i.e.,* the convex hull of L.
- The convex envelope of the set L is given by the border of its convex hull.

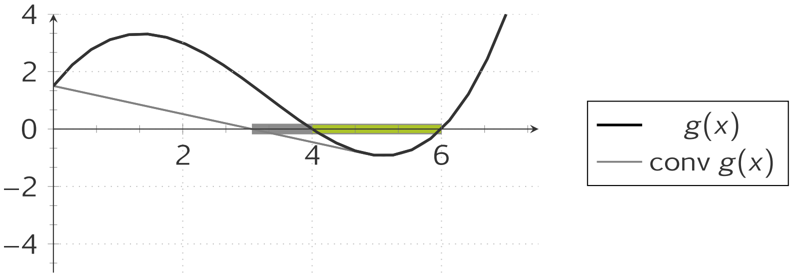
Convex relaxations of the level set $L = \{x | g(x) \leq 0\}$



- The level sets $L_{\alpha}^{g} = \{x | g(x) \leq \alpha\}$ are:

$$L_{\alpha=0}^{g} = [4, 6]$$

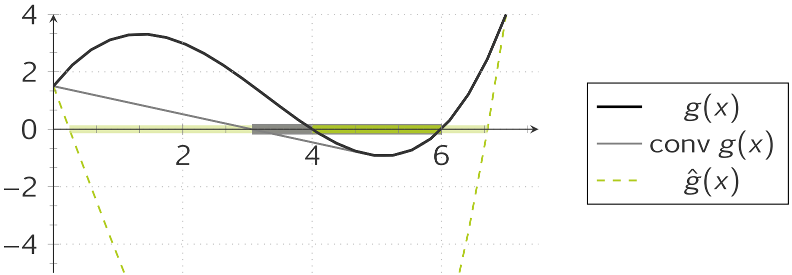
Convex relaxations of the level set $L = \{x | g(x) \leq 0\}$



- The level sets $L_\alpha^g = \{x | g(x) \leq \alpha\}$ are:

$$L_{\alpha=0}^{g} = [4, 6] \qquad L_{\alpha=0}^{\text{conv } g} = [3.069, 6]$$

Convex relaxations of the level set $L = \{x | g(x) \leq 0\}$

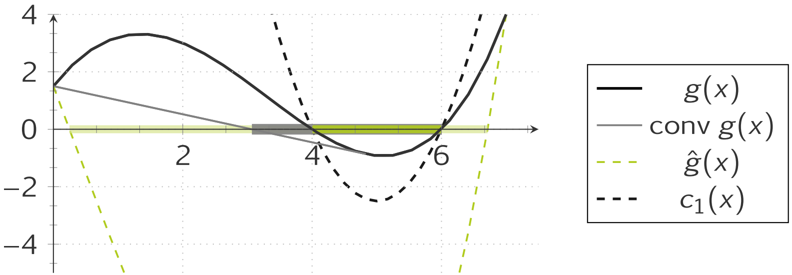


- The level sets $L_{\alpha}^{g} = \{x | g(x) \leq \alpha\}$ are:

$$L_{\alpha=0}^{g} = [4,6] \qquad L_{\alpha=0}^{\text{conv } g} = [3.069, 6]$$

$$L_{\alpha=0}^{\hat{g}} = [0.248, 6.713]$$

Convex relaxations of the level set $L = \{x | g(x) \leq 0\}$

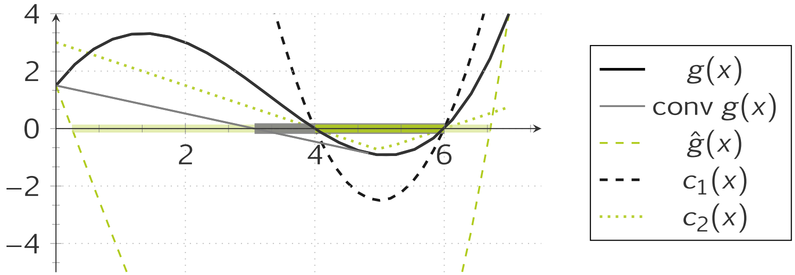


- The level sets $L_\alpha^g = \{x | g(x) \leq \alpha\}$ are:

$$L_{\alpha=0}^{g} = [4, 6] \qquad L_{\alpha=0}^{\text{conv } g} = [3.069, 6]$$
$$L_{\alpha=0}^{\hat{g}} = [0.248, 6.713] \qquad L_{\alpha=0}^{c_1} = [4, 6]$$

- A possible tight convex relaxation: $c_1(x) = \frac{5}{2}(x-4)(x-6)$.

Convex relaxations of the level set $L = \{x | g(x) \leq 0\}$



- The level sets $L_\alpha^g = \{x | g(x) \leq \alpha\}$ are:

$$L_{\alpha=0}^{g} = [4, 6] \qquad L_{\alpha=0}^{\text{conv } g} = [3.069, 6]$$

$$L_{\alpha=0}^{\hat{g}} = [0.248, 6.713] \qquad L_{\alpha=0}^{c_1} = L_{\alpha=0}^{c_2} = [4, 6]$$

- Another tight convex relaxation:
  $c_2(x) = \max\left\{-\frac{3}{4}(x-4),\ \frac{3}{4}(x-6)\right\}.$

A nonconvex size constraint in two dimensions

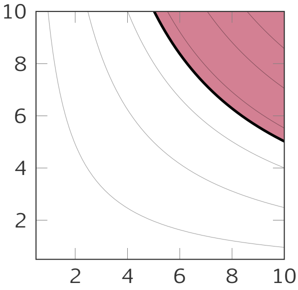
- Consider the inequality constraint

$$g(\mathbf{x}) \leq 0,$$

where

$$g(\mathbf{x}) = 50 - x_1 \cdot x_2, \quad 0.5 \leq x_1,\ x_2 \leq 10.$$

- The contour plot of the constraint function g is



McCormick convex relaxation

- The convex envelope of the negative bilinear term $-x_1 x_2$ is

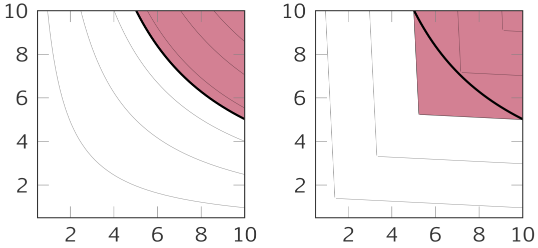
$$\max\{-\overline{x}_1 x_2 - \underline{x}_2 x_1 + \overline{x}_1 \underline{x}_2,\ -\underline{x}_1 x_2 - \overline{x}_2 x_1 + \underline{x}_1 \overline{x}_2\}$$

  where the bounds of the variables are $\underline{x}_i \leq x_i \leq \overline{x}_i$.

- If $0.5 \leq x_1,\ x_2 \leq 10$, we then obtain

$$\begin{aligned}\operatorname{conv} g(\mathbf{x}) = 50 - \max\{&-10 \cdot x_1 - 0.5 \cdot x_2 + 5,\\ &-0.5 \cdot x_1 - 10 \cdot x_2 + 5\}\end{aligned}$$

The level sets for the McCormick relaxation



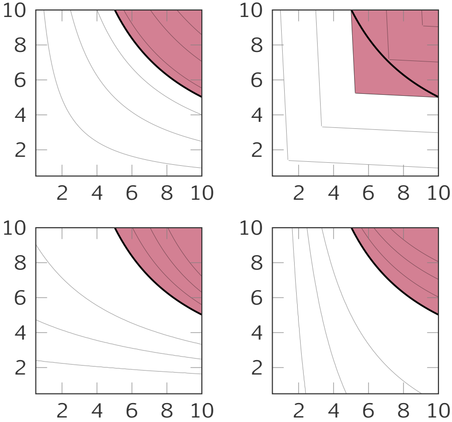
Left: The level set $L^{g}_{\alpha=0}$. *Right:* The level set $L^{\text{conv } g}_{\alpha=0}$.

- Observe that, although $L^{g}_{\alpha=0}$ is a convex set, replacing $g(\mathbf{x}) \leq 0$ with $\text{conv } g(\mathbf{x}) \leq 0$ does not give the tightest convex relaxation of $L^{g}_{\alpha=0}$.

A convex reformulation

- By reformulating
$$g(\mathbf{x}) = 50 - x_1 \cdot x_2$$
at $g(\mathbf{x}) = 0$ we can, in this case, obtain the following convex constraints exactly defining the border of the level set $L^g_{\alpha=0}$:
$$c_1(\mathbf{x}) = \frac{50}{x_2} - x_1 \qquad \text{and} \qquad c_2(\mathbf{x}) = \frac{50}{x_1} - x_2.$$
- Since $c_1(\mathbf{x})$ and $c_2(\mathbf{x})$ exactly define the border of $L^g_{\alpha=0}$, it follows that
$$L^{c_1}_{\alpha=0} \equiv L^{c_2}_{\alpha=0} \equiv L^g_{\alpha=0}.$$

The level sets for the convex reformulation



Upper left: The level set $L^{g}_{\alpha=0}$. *Upper right:* The level set $L^{\mathrm{conv}\,g}_{\alpha=0}$.
Lower left: The level set $L^{c_1}_{\alpha=0}$. *Lower right:* The level set $L^{c_2}_{\alpha=0}$.

3D illustration of the relaxations

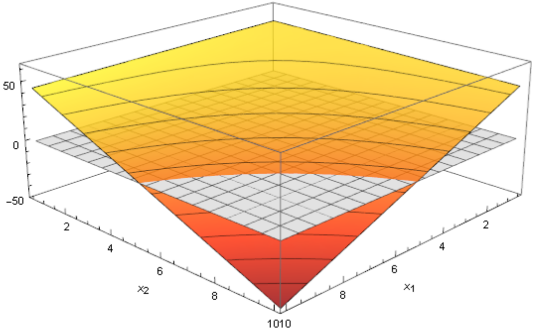


Illustration of $g(x)$

3D illustration of the relaxations

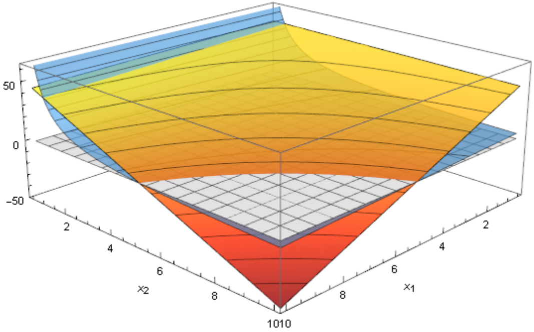


Illustration of $g(\mathbf{x})$ and $c_1(\mathbf{x})$

3D illustration of the relaxations

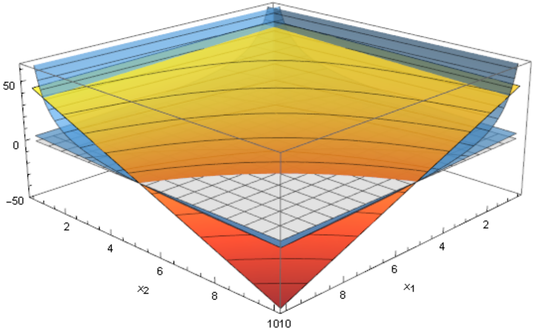


Illustration of $g(\mathbf{x})$, $c_1(\mathbf{x})$ and $c_2(\mathbf{x})$

5. A reformulation algorithm for solving C^2 MINLP problems

Introduction

- A framework for reformulating nonconvex (twice-differentiable – C^2) mixed integer nonlinear programming (MINLP) problems to convex form is presented.
  - The framework is an extension to a previously introduced reformulation technique for signomial problems.
  - For C^2-constraints, convex reformulations are made in an extended variable-space using variants of the αBB quadratic convex underestimator.
  - With the framework, a nonconvex problem can be reformulated to a larger convex MINLP problem solved in one step or to a sequence of smaller relaxed MINLP problems solved iteratively.

The considered problem-type

Nonconvex problem

$$\begin{aligned} \text{min.} \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & \mathbf{q}(\mathbf{x}) + \mathbf{h}(\mathbf{x}) \leq 0 \\ & \underline{\mathbf{x}} \leq \mathbf{x} \leq \overline{\mathbf{x}} \end{aligned}$$

- $f(\mathbf{x})$ is a convex function
- $\mathbf{q}(\mathbf{x})$ are convex functions
- $\mathbf{h}(\mathbf{x})$ are nonconvex twice-differentiable (C^2) functions
- the variables in $\mathbf{x}$ are reals, binaries or integers

- Nonconvex twice-differentiable functions (incl. signomials) can be convexified using an αBB-type reformulation.

Convex underestimation of C^2-functions

- A convex underestimator for twice-differentiable functions in a box-domain from, *e.g.,* Floudas (2000).

Convex underestimation of C^2-functions

- A convex underestimator for twice-differentiable functions in a box-domain from, *e.g.,* Floudas (2000).

Theorem

A function $g(\mathbf{x}) \in C^2$ has the convex underestimator

$$\boxed{\hat{g}(\mathbf{x}) = g(\mathbf{x}) + \sum_i \alpha(\underline{x}_i - x_i)(\overline{x}_i - x_i)}$$

for $x_i \in [\underline{x}_i, \overline{x}_i]\ \forall i$ if and only if the parameter α fulfills

$$\alpha \geq \max\left\{0, -\frac{1}{2}\min_i \lambda_i\right\}$$

where the λ_i's are the eigenvalues of the Hessian matrix of $g(\mathbf{x})$ on the interval $[\underline{x}_i, \overline{x}_i]$.

- Several methods for calculating the α-values are available.

Gerschgorin's circle theorem

Theorem

Let $A \in \mathbb{C}^{n\times n}$ with entries a_{ij} and define $R_i = \sum_{j\neq i} |a_{ij}|$. Every eigenvalue of A lies within at least one of the Gerschgorin disks

$$D(a_{ii}, R_i) = \{x : |x - a_{ii}| \leq R_i\}.$$

Gerschgorin's circle theorem

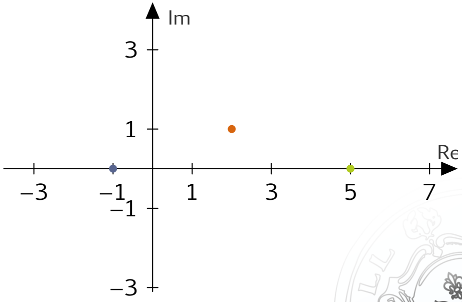
Theorem

Let $A \in \mathbb{C}^{n\times n}$ with entries a_{ij} and define $R_i = \sum_{j\neq i} |a_{ij}|$. Every eigenvalue of A lies within at least one of the Gerschgorin disks

$$D(a_{ii}, R_i) = \{x : |x - a_{ii}| \leq R_i\}.$$

Example

$$A = \begin{bmatrix} 2+i & 2 & -1 \\ 1 & 5 & i \\ 1 & -1 & -1 \end{bmatrix}$$



Gerschgorin's circle theorem

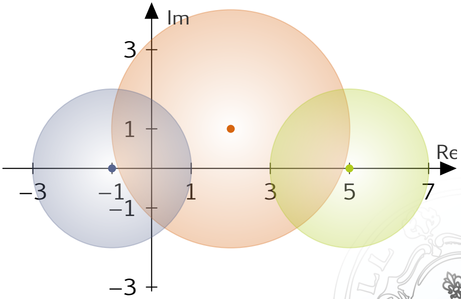
Theorem

Let $A \in \mathbb{C}^{n\times n}$ with entries a_{ij} and define $R_i = \sum_{j\neq i} |a_{ij}|$. Every eigenvalue of A lies within at least one of the Gerschgorin disks

$$D(a_{ii}, R_i) = \{x : |x - a_{ii}| \leq R_i\}.$$

Example

$$A = \begin{bmatrix} 2+i & 2 & -1 \\ 1 & 5 & i \\ 1 & -1 & -1 \end{bmatrix}$$



Gerschgorin's circle theorem

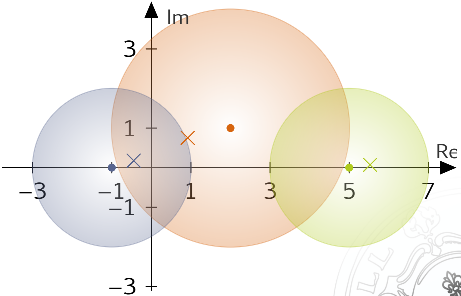
Theorem

Let $A \in \mathbb{C}^{n\times n}$ with entries a_{ij} and define $R_i = \sum_{j\neq i} |a_{ij}|$. Every eigenvalue of A lies within at least one of the Gerschgorin disks

$$D(a_{ii}, R_i) = \{x : |x - a_{ii}| \leq R_i\}.$$

Example

$$A = \begin{bmatrix} 2+i & 2 & -1 \\ 1 & 5 & i \\ 1 & -1 & -1 \end{bmatrix}$$

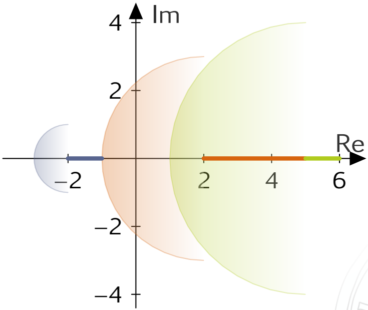


Extending Gerschgorin's circle theorem to interval matrices

- The circle theorem can be extended to interval matrices by considering the worst case.
- Positive-semidefiniteness is wanted, therefore "worst case" should be interpreted as lowest eigenvalue.

Example

$$H = \begin{bmatrix} [2,5] & [-1,3] & 0 \\ [-1,3] & [5,6] & [-1,0] \\ 0 & [-1,0] & [-2,-1] \end{bmatrix}$$



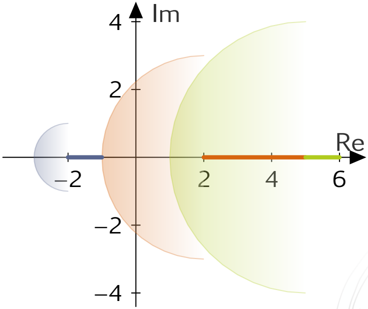
Diagonal αBB using the Gerschgorin Method

- The function is underestimated by adding the perturbation $-\sum_i \alpha_i(\overline{x}_i - x_i)(x_i - \underline{x}_i)$.
- To guarantee positive-semidefiniteness we set the constraints $\underline{h_{ii}} - R_i + 2\alpha_i \geq 0,\ i = 1, 2, \ldots, n$.

$$\begin{bmatrix} [2,5] & [-1,3] & 0 \\ [-1,3] & [5,6] & [-1,0] \\ 0 & [-1,0] & [-2,-1] \end{bmatrix}$$

$$+\begin{bmatrix} 0 & & \\ & 0 & \\ & & 0 \end{bmatrix}$$

$$=\begin{bmatrix} [2,5] & [-1,3] & 0 \\ [-1,3] & [5,6] & [-1,0] \\ 0 & [-1,0] & [-2,-1] \end{bmatrix}$$



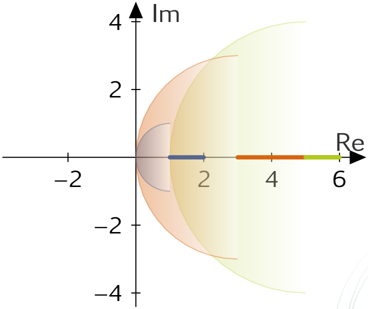
Diagonal αBB using the Gerschgorin Method

- The function is underestimated by adding the perturbation $-\sum_i \alpha_i(\overline{x}_i - x_i)(x_i - \underline{x}_i)$.
- To guarantee positive-semidefiniteness we set the constraints $\underline{h_{ii}} - R_i + 2\alpha_i \geq 0,\ i = 1, 2, \ldots, n$.

$$\begin{bmatrix} [2,5] & [-1,3] & 0 \\ [-1,3] & [5,6] & [-1,0] \\ 0 & [-1,0] & [-2,-1] \end{bmatrix}$$

$$+\begin{bmatrix} 1 & & \\ & 0 & \\ & & 3 \end{bmatrix}$$

$$= \begin{bmatrix} [3,6] & [-1,3] & 0 \\ [-1,3] & [5,6] & [-1,0] \\ 0 & [-1,0] & [1,4] \end{bmatrix}$$



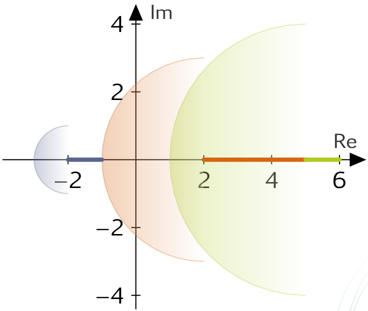
Diagonal and off-diagonal αBB

- The function can also be underestimated by adding $-\sum_i \alpha_i(\overline{x}_i - x_i)(x_i - \underline{x}_i) + \sum_i \sum_{j>i} \beta_{ij} x_i x_j$ as in Skjäl et al. (2012).
- To guarantee positive-semidefiniteness we can then manipulate the diagonal and off-diagonal elements of the resulting Hessian matrix: the radius and midpoint of each Gerschgorin circle will be altered in the constraints $\underline{h_{ii}} + 2\alpha_i - \sum_{j\neq i} \left| h_{ij}^{'} + \beta_{ij} \right| \geq 0 \; \forall i,\, h_{ij}^{'} \in [\underline{h_{ij}}, \overline{h_{ij}}]$.

$$\begin{bmatrix} [2,5] & [-1,3] & 0 \\ [-1,3] & [5,6] & [-1,0] \\ 0 & [-1,0] & [-2,-1] \end{bmatrix}$$

$$+\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

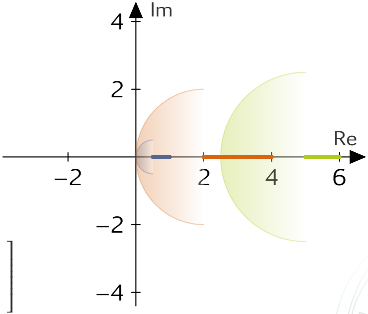
$$=\begin{bmatrix} [2,5] & [-1,3] & 0 \\ [-1,3] & [5,6] & [-1,0] \\ 0 & [-1,0] & [-2,-1] \end{bmatrix}$$



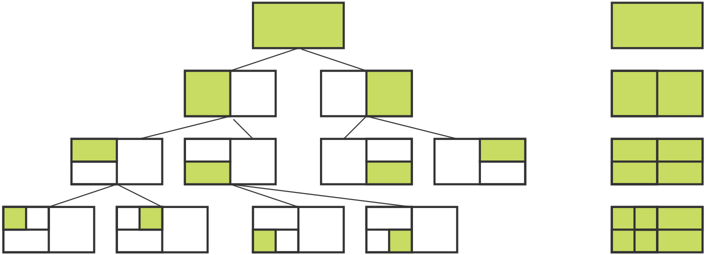
Diagonal and off-diagonal αBB

- The function can also be underestimated by adding $-\sum_i \alpha_i(\overline{x}_i - x_i)(x_i - \underline{x}_i) + \sum_i \sum_{j>i} \beta_{ij} x_i x_j$ as in Skjäl et al. (2012).
- To guarantee positive-semidefiniteness we can then manipulate the diagonal and off-diagonal elements of the resulting Hessian matrix: the radius and midpoint of each Gerschgorin circle will be altered in the constraints $\underline{h_{ii}} + 2\alpha_i - \sum_{j\neq i} \left| h'_{ij} + \beta_{ij} \right| \geq 0 \; \forall i, \, h'_{ij} \in [\underline{h_{ij}}, \overline{h_{ij}}]$.

$$
\begin{bmatrix} [2,5] & [-1,3] & 0 \\ [-1,3] & [5,6] & [-1,0] \\ 0 & [-1,0] & [-2,-1] \end{bmatrix}
+ \begin{bmatrix} 0 & -1 & 0 \\ -1 & 0 & 1/2 \\ 0 & 1/2 & 5/2 \end{bmatrix}
= \begin{bmatrix} [2,5] & [-2,2] & 0 \\ [-2,2] & [5,6] & [-1/2,1/2] \\ 0 & [-1/2,1/2] & [1/2,3/2] \end{bmatrix}
$$



Branching vs reformulation



- **Branching:** n convex subproblems (the subproblems with the green domains are solved using a branching strategy)
- **Reformulation:** a sequence of convex MINLP problems are solved (the whole domain is considered in each iteration)

Including αBB in the reformulation framework

- To be able to reformulate the problem in subdomains without branching, a convex quadratic function αx^2 is added to and a variable $\widehat{W}$ subtracted from the nonconvex C^2 constraint, *i.e.*,

$$\underbrace{h(x) + \alpha x^2 - \widehat{W}}_{\text{convex}} \leq 0.$$

Including αBB in the reformulation framework

- To be able to reformulate the problem in subdomains without branching, a convex quadratic function αx^2 is added to and a variable $\widehat{W}$ subtracted from the nonconvex C^2 constraint, *i.e.*,

$$\underbrace{h(x) + \alpha x^2 - \widehat{W}}_{\text{convex}} \leq 0.$$

- If α is large enough, then the reformulated constraint will be convex.

Including αBB in the reformulation framework

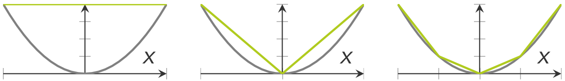
- To be able to reformulate the problem in subdomains without branching, a convex quadratic function αx^2 is added to and a variable $\widehat{W}$ subtracted from the nonconvex C^2 constraint, *i.e.,*

$$\underbrace{h(x) + \alpha x^2 - \widehat{W}}_{\text{convex}} \leq 0.$$

- If α is large enough, then the reformulated constraint will be convex.
- If $\alpha x^2 - \widehat{W} \leq 0$, then the reformulated constraint underestimates the original one.

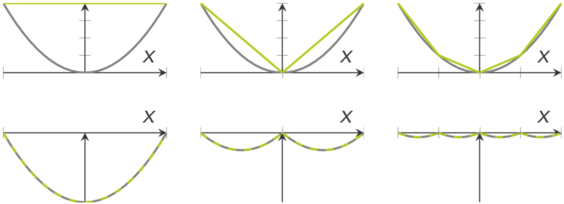
The convex reformulation in subdomains

$$\boxed{h(x) + \alpha x^2 - \widehat{W} \leq 0}$$



The convex reformulation in subdomains

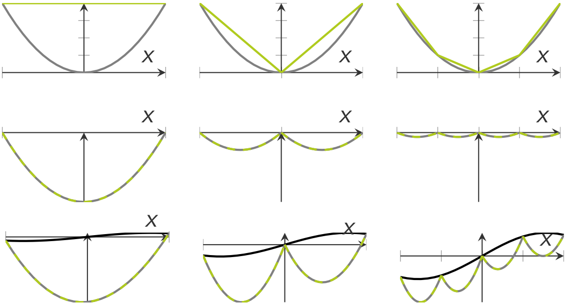
$$\boxed{h(x) + \alpha x^2 - \widehat{W} \leq 0}$$



- If α in αx^2 is large enough then $h(x) + \alpha x^2 - \widehat{W}$ will be convex.
- If $\widehat{W}$ is given by a PLF of αx^2 then $h(x)$ is also underestimated in each subdomain since $\alpha x^2 - \widehat{W} \leq 0$.

The convex reformulation in subdomains

$$\boxed{h(x) + \alpha x^2 - \widehat{W} \leq 0}$$

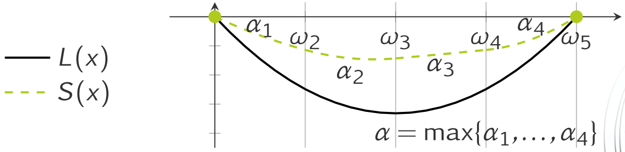


The spline αBB underestimator

- The spline αBB-underestimator is a smooth convex piecewise polynomial expression

$$S(x) = \begin{cases} \alpha_1 x^2 + \beta_1 x + \gamma_1 & \text{if } x \in [\omega_1, \omega_2] \\ \alpha_2 x^2 + \beta_2 x + \gamma_2 & \text{if } x \in [\omega_2, \omega_3] \\ \vdots & \vdots \\ \alpha_{K-1} x^2 + \beta_{K-1} x + \gamma_{K-1} & \text{if } x \in [\omega_{K-1}, \omega_K], \end{cases}$$

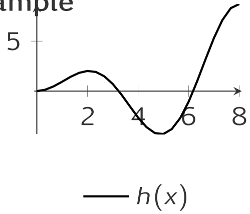
- The α_k's ensure convexity. The β_k and γ_k for $k \in \{2, \ldots, K-1\}$ ensure smoothness and continuity, and β_1, γ_1 gives $S(\omega_1) = S(\omega_K) = 0$.



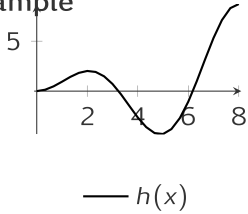
An illustrative example

- Consider the function

$$h(x) = x \cdot \sin x + x/10.$$



An illustrative example



- Consider the function

$$h(x) = x \cdot \sin x + x/10.$$

- The convex underestimators are then

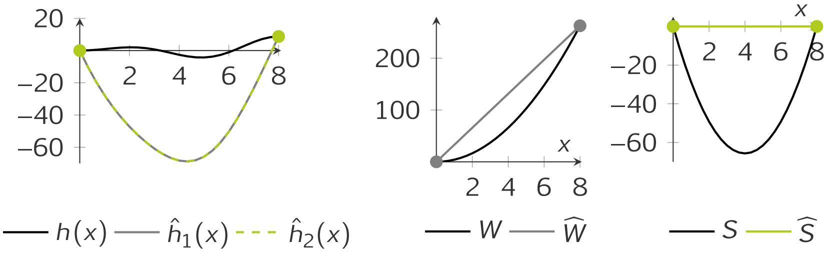
$$\hat{h}_1(x) = x \cdot \sin x + x/10 + \alpha x^2 - \widehat{W}$$

for the reformulated αBB understimator using constant α and

$$\hat{h}_2(x) = x \cdot \sin x + x/10 + S(x) - \widehat{S}$$

for the reformulated spline αBB underestimator, where $\widehat{W}$ is the PLF of $W = \alpha x^2$ and $\widehat{S}$ is the PLF of the spline function $S(x)$.

An illustrative example

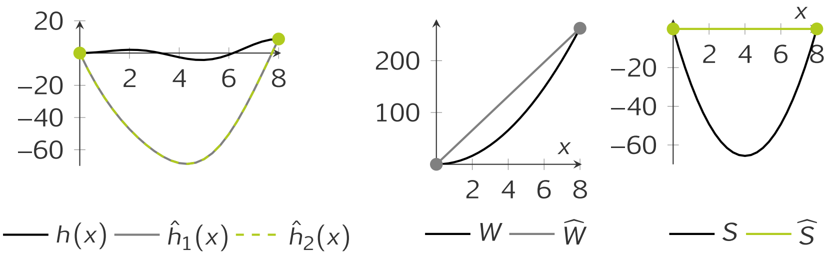


$$\hat{h}_1(x) = x \cdot \sin x + x/10 + \boxed{\alpha x^2 - \widehat{W}} \qquad \alpha = 4.10293$$

$$\hat{h}_2(x) = x \cdot \sin x + x/10 + \boxed{S(x) - \widehat{S}}$$

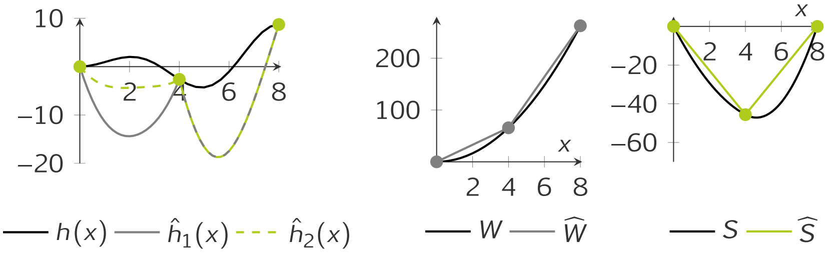
$\widehat{W}$ is the PLF of $W = \alpha x^2$ and $\widehat{S}$ is the PLF of $S(x)$

An illustrative example



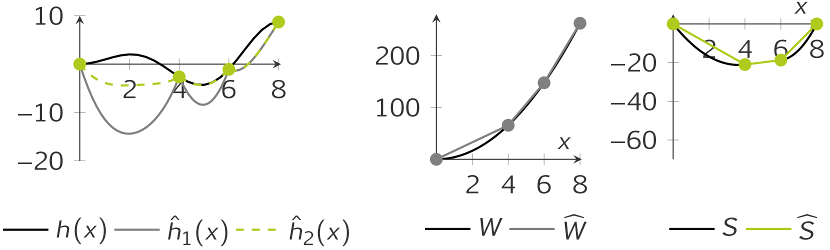
$$W(x) = 4.1x^2 \qquad S(x) = 4.1x^2 - 32.8x, 0 \leq x \leq 8$$

An illustrative example



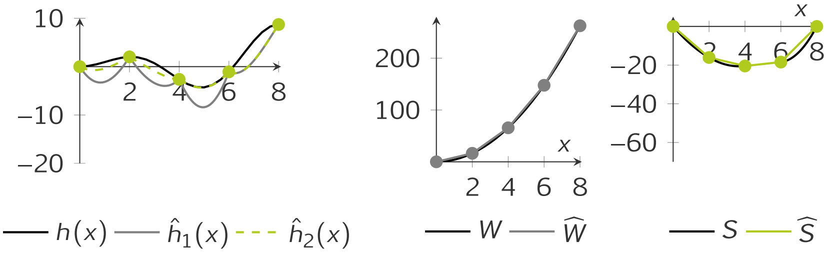
$$W(x) = 4.1x^2 \qquad S(x) = \begin{cases} 1.6x^2 - 17.8x & 0 \le x \le 4 \\ 4.1x^2 - 37.8x + 40.0 & 4 \le x \le 8 \end{cases}$$

An illustrative example



$$W(x) = 4.1x^2 \qquad S(x) = \begin{cases} 1.6x^2 - 11.7x & 0 \leq x \leq 4 \\ 1.1x - 25.6 & 4 \leq x \leq 6 \\ 4.1x^2 - 48.1x + 122.1 & 6 \leq x \leq 8 \end{cases}$$

An illustrative example



$$W(x) = 4.1x^2 \qquad S(x) = \begin{cases} 1.3x^2 - 10.7x & 0 \le x \le 2 \\ 1.6x^2 - 11.8x + 1.1 & 2 \le x \le 4 \\ 1.1x - 24.5 & 4 \le x \le 6 \\ 4.1x^2 - 48.2x + 123.2 & 6 \le x \le 8 \end{cases}$$

Generalization to N dimensions

- The formulation can easily be extended from one to N dimensions by using the underestimators

$$h(\mathbf{x}) + \sum_{i=1}^{N} \left(\alpha_i x_i^2 - \widehat{W}_i \right) \leq 0, \quad \mathbf{x} = (x_1, x_2, \ldots, x_N), \quad \text{or}$$

$$h(\mathbf{x}) + \sum_{i=1}^{N} \left(S_i(x_i) - \widehat{S}_i \right) \leq 0, \quad \mathbf{x} = (x_1, x_2, \ldots, x_N).$$

when using the reformulated versions of the original αBB and spline αBB underestimators respectively.

Generalization to N dimensions

- The formulation can easily be extended from one to N dimensions by using the underestimators

$$h(\mathbf{x}) + \sum_{i=1}^{N} \left(\alpha_i x_i^2 - \widehat{W}_i \right) \leq 0, \quad \mathbf{x} = (x_1, x_2, \ldots, x_N), \quad \text{or}$$

$$h(\mathbf{x}) + \sum_{i=1}^{N} \left(S_i(x_i) - \widehat{S}_i \right) \leq 0, \quad \mathbf{x} = (x_1, x_2, \ldots, x_N).$$

  when using the reformulated versions of the original αBB and spline αBB underestimators respectively.

- Here $\widehat{W}_i$ is the PLF of $W_i = \alpha_i x_i^2$ and $\widehat{S}_i$ is the PLF of S_i.

Reformulation or implementation in a global optimization algorithm

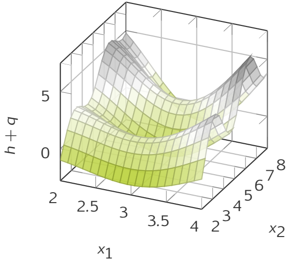
- The underestimator can be used for reformulation or directly implemented in a global optimization algorithm, *e.g.*, αGO, for solving nonconvex MINLP problems with C^2-constraints, *c.f.*, Lundell et al. (2013).
- A sequence of overestimated convex MINLP problems is solved (see Eronen et al. (2012) for convex MINLP methods) until the solution fulfills the constraints in the original nonconvex problem.
- The feasible region of the overestimated convexified problem is reduced in each iteration by improving the PLFs of $W = \alpha_i x_i^2$ or $S(x)$.

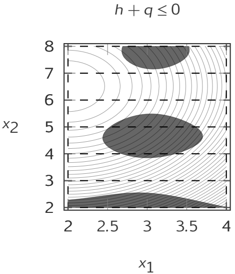
The original nonconvex MINLP problem

$$\begin{aligned}
\text{minimize} \quad & f(x_1,x_2) = (2x_1-4)^2 + (x_2-13/2)^2 \\
\text{subject to} \quad & \underbrace{x_1\cos^2 x_2 + x_2\sin^2 x_1 - 3/x_2}_{h(x_1,x_2)} + \underbrace{x_1/2 - 5/2}_{q(x_1)} \le 0, \\
& 2 \le x_1 \le 4, \quad 2 \le x_2 \le 8, \quad x_1 \in \mathbb{R}, \quad x_2 \in \mathbb{Z}.
\end{aligned}$$

The original nonconvex MINLP problem

$$\begin{aligned}
\text{minimize} \quad & f(x_1,x_2) = (2x_1-4)^2 + (x_2-13/2)^2 \\
\text{subject to} \quad & \underbrace{x_1\cos^2 x_2 + x_2\sin^2 x_1 - 3/x_2}_{h(x_1,x_2)} + \underbrace{x_1/2 - 5/2}_{q(x_1)} \le 0, \\
& 2 \le x_1 \le 4, \quad 2 \le x_2 \le 8, \quad x_1 \in \mathbb{R}, \quad x_2 \in \mathbb{Z}.
\end{aligned}$$



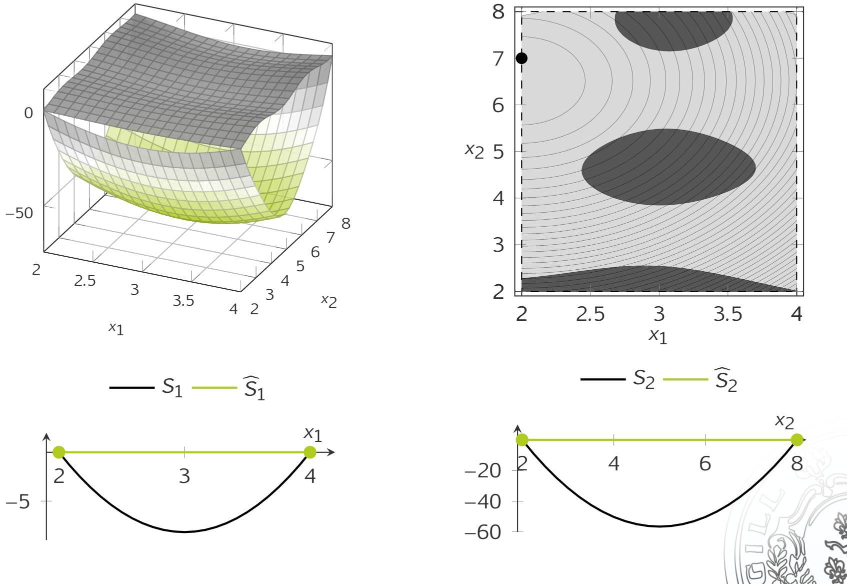


The reformulated MINLP problem

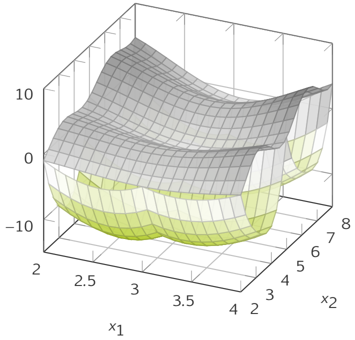
$$
\begin{aligned}
\text{minimize} \quad & f(x_1, x_2) = (2x_1 - 4)^2 + (x_2 - 13/2)^2 \\
\text{subject to} \quad & x_1 \cos^2 x_2 + x_2 \sin^2 x_1 - 3/x_2 + x_1/2 - 5/2 \\
& \qquad\qquad + S_1(x_1) + S_2(x_2) - \widehat{S}_1 - \widehat{S}_2 \leq 0, \\
& \widehat{S}_1 = \mathrm{PLF}(S_1(x_2), V_1 ; \Omega_1),\ \widehat{S}_2 = \mathrm{PLF}(S_2(x_2), V_2 ; \Omega_2), \\
& 2 \leq x_1 \leq 4, \quad 2 \leq x_2 \leq 8, \quad x_1 \in \mathbb{R}, \quad x_2 \in \mathbb{Z}, \\
& V_i \text{ and } \Omega_i \text{ are sets including the variables} \\
& \text{and breakpoints in } \mathrm{PLF}_i \text{ of } S_i(x_1)
\end{aligned}
$$

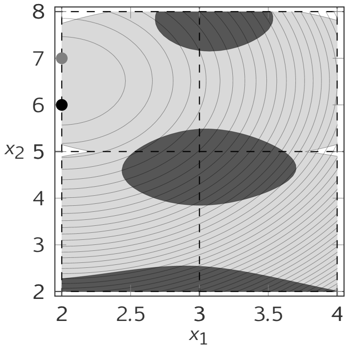
- This reformulated problem is convex in the extended variable space consisting of the original variables x_1 and x_2, as well as, those needed for the PLFs in V_1 and V_2.

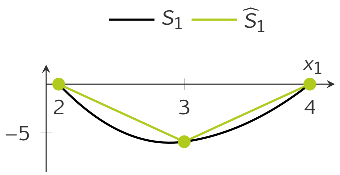
αGO iteration 1

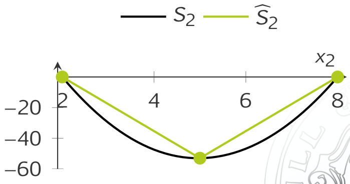


αGO iteration 2

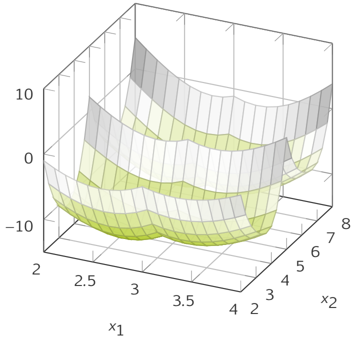


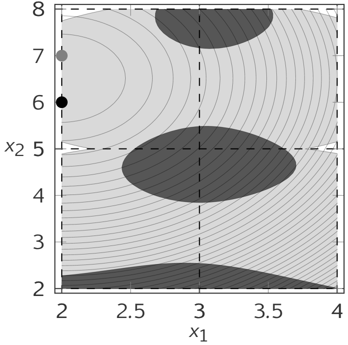


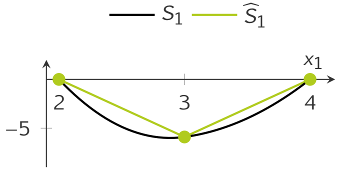


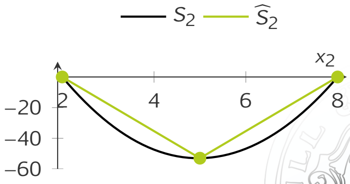


αGO iteration 2

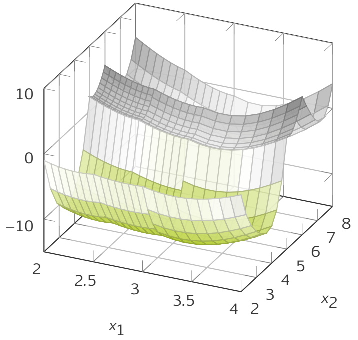


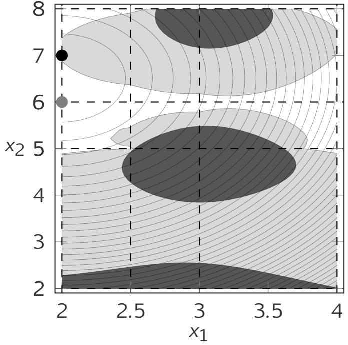


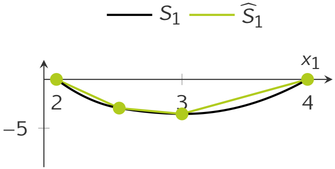


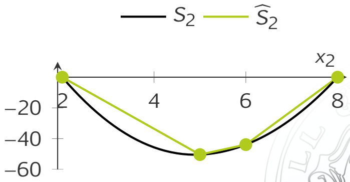


αGO iteration 3

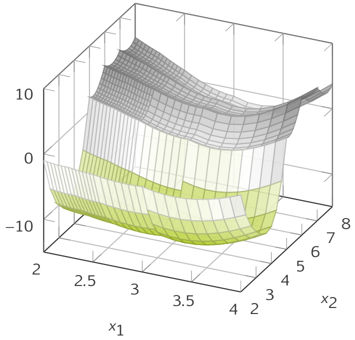


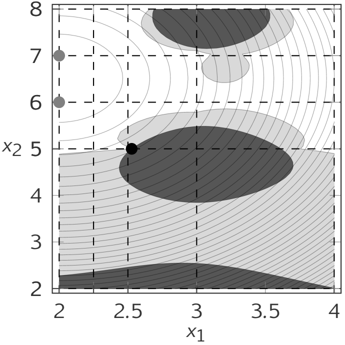


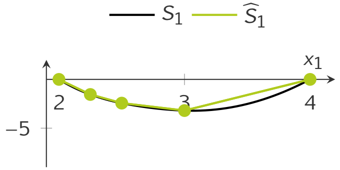


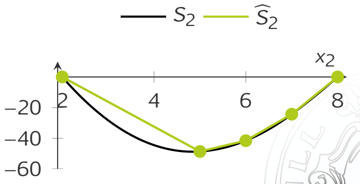


αGO iteration 4

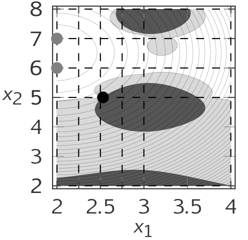








αGO iteration 5 and 6



Iter.	Regions	$f(x_1, x_2)$	x_1	x_2	$h(x_1, x_2) + q(x_1)$
1	1	0.2500	2.0	7	4.9959
2	4	0.2500	2.0	6	4.9959
3	9	0.2500	2.0	7	4.9959
4	16	3.3630	2.52749	5	0.0273
5	20	3.3767	2.53074	5	0.0139
6	24	3.3848	2.53263	5	0.0061

Summary

1. Introduction – a short background to MINLP

2. Some aspects on convex MINLP algorithms
   - Convex functions and convex sets
   - Smooth and nonsmooth functions

3. A new algorithm for solving convex MINLP problems

4. Aspects on solving nonconvex MINLP problems
   - Convex relaxations in BB and relaxation frameworks
   - Convex envelopes of functions or level sets

5. A reformulation algorithm for solving C^2 MINLP problems

Some references

C. S. Adjiman, S. Dallwig, C. A. Floudas, and A. Neumaier.
A global optimization method, αBB, for general twice-differentiable constrained NLPs – I. Theoretical advances.
Computers and Chemical Engineering, 22(9):1137–1158, 1998.

M. A. Duran and I. E. Grossmann.
An outer-approximation algorithm for a class of mixed-integer nonlinear programs.
Mathematical Programming, 36(3):307–339, 1986.

V.-P. Eronen, M. M. Mäkelä, and T. Westerlund.
Extended cutting plane method for a class of nonsmooth nonconvex MINLP problems.
Optimization, (available online):1–21, 2013.

A. Lundell, A. Skjäl, and T. Westerlund.
A reformulation framework for global optimization.
Journal of Global Optimization, 57(1):115–141, 2013.

Y. Nesterov.
Introductory lectures on convex optimization: A basic course.
Kluwer Academic Publishers, 2004.

T. Westerlund and F. Pettersson.
An extended cutting plane method for solving convex MINLP problems.
Computers & Chemical Engineering, 19:131–136, 1995.

The end of the presentation

Thank you for listening!

The presentation including relevant references will be available at **www.abo.fi/ose**