Mixed-Integer Nonlinear Optimisation: Piecewise Linear Underestimators

Ruth Misener
r.misener@imperial.ac.uk

Computational Optimisation Group
Centre for Process Systems Engineering
Department of Computing

Imperial College London

computational optimisation group

Centre for Process Systems Engineering

2017/05/17

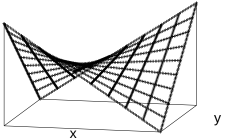
Piecewise Linear Underestimators

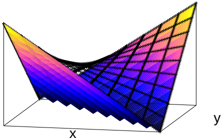
Key Idea

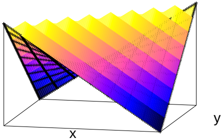
Nonconvex bilinear optimisation problems are really difficult. Maybe we could solve a MILP to get an approximate solution or a bound?

- Earliest work (to Ruth's knowledge!): Glover, *Mgt Science*, 1975.
- Process systems engineering: Meyer & Floudas, *AIChE J*, 2006, Karuppiah & Grossmann, *Comput Chem Eng*, 2006.

There are many other examples in the years 2007 - 2017! See the work of Bienstock, Castro, Dey, Floudas, Grossmann, Martin, etc.







This presentation: Misener, Thompson, Floudas, *Comput Chem Eng*, 2010.

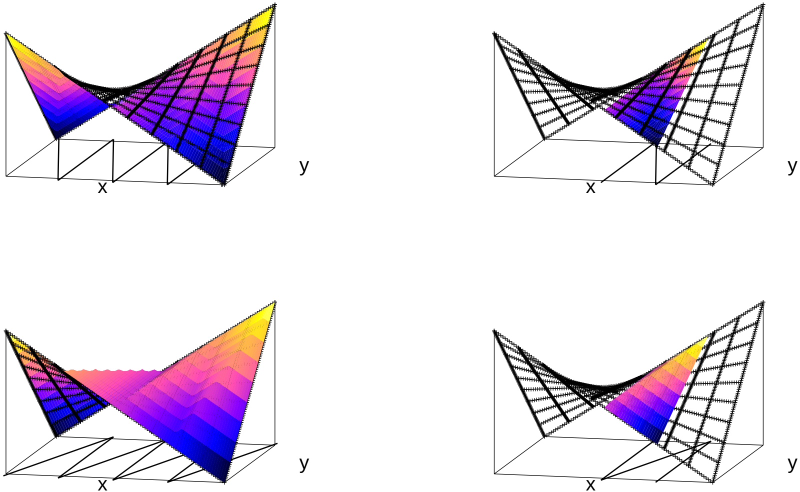
Disjunctive Program

Consider bilinear function $z = x \cdot y$ defined on a box $x,\, y \in \mathbb{R},\, x^L \le x \le x^U,\, y^L \le y \le y^U$. We partition one of the variables (x) into N_P segments and come up with the disjunctive program:

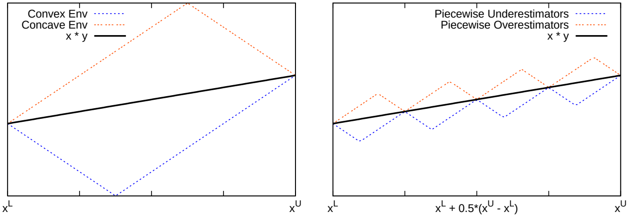
$$\bigvee n_P \in \{1, \ldots, N_P\} \left[\begin{array}{l} W(n_P) \\ z \ge x \cdot y^L + (x^L + a \cdot (n_P - 1)) \cdot (y - y^L) \\ z \ge x \cdot y^U + (x^L + a \cdot n_P) \cdot (y - y^U) \\ z \le x \cdot y^L + (x^L + a \cdot n_P) \cdot (y - y^L) \\ z \le x \cdot y^U + (x^L + a \cdot (n_P - 1)) \cdot (y - y^U) \\ x^L + a \cdot (n_P - 1) \le x \le x^L + a \cdot n_P \\ y^L \le y \le y^U \end{array} \right]$$

There are N_P segments on range $\left[x^L,\, x^U\right]$ and each segment is bounded by $\left[x^L + a \cdot (n_P - 1),\, x^L + a \cdot n_P\right]$ $\forall\, n_P \in \{1, \ldots, N_P\}$ where $a = \frac{x^U - x^L}{N_P}$.

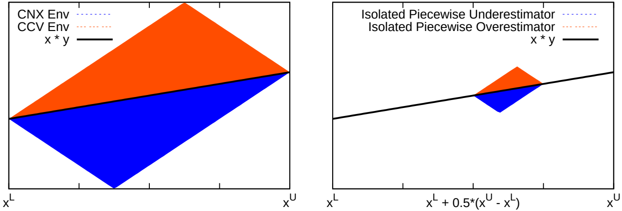
Piecewise Under- & Overestimators when ($N_P = 4$)



Projecting the Under- & Overestimators



Under- & Overestimators of $(x \cdot y)$ Projected on $\left(y = y^L + \frac{5}{8} \cdot (y^U - y^L)\right)$



Feasible Set Corresponding to the Relaxation of $(x \cdot y)$ Projected on $\left(y = y^L + \frac{5}{8} \cdot (y^U - y^L)\right)$

Number of Additional Vars & Constraints

Number of Additional Vars & Constraints for the McCormick, Linear, & Logarithmic Relaxation of a Single Bilinear Term

	Continuous Vars	Binary Vars	Constraints
McC Hull	1	–	4
PW Linear	$N_P + 1$	N_P	$N_P + 8$
PW Log	$2 \cdot N_P + 1$	$\lceil \log_2 N_P \rceil$	$N_P + 2 \cdot \lceil \log_2 N_P \rceil + 8$

Underestimators with a Linear Number of Binary Vars [1/4]

Introduce two additional variable sets:

- Binary switch: $\lambda \in \{0,\, 1\}^{N_P}$
- Continuous switch: $\Delta y \in \left[0,\, y^U - y^L\right]^{N_P}$

The binary switch λ is active, i.e. $\lambda(n_P) = 1$, for the segment where $(x^L + a \cdot (n_P - 1) \leq x \leq x^L + a \cdot n_P)$ and is otherwise inactive:

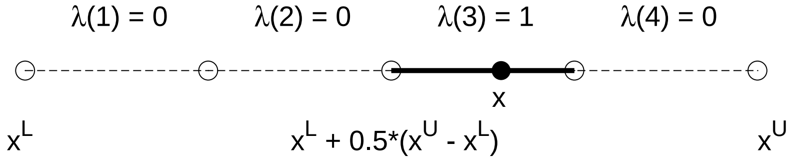
$$\sum_{n_P=1}^{N_P} \lambda(n_P) = 1$$

$$x^L + \sum_{n_P=1}^{N_P} a \cdot (n_P - 1) \cdot \lambda(n_P) \leq x \leq x^L + \sum_{n_P=1}^{N_P} a \cdot n_P \cdot \lambda(n_P)$$

Alternatively, we could declare λ to be a special ordered set of type 1 (SOS1) because exactly one member of the set is non-zero.

Underestimators with a Linear Number of Binary Vars [2/4]

Activation of a Single $\lambda(n_P)$, i.e. $n_P = 3 \implies \lambda(3) = 1$, according to the Value of x ($N_P = 4$ and $x = x^L + 2.6 \cdot a$)

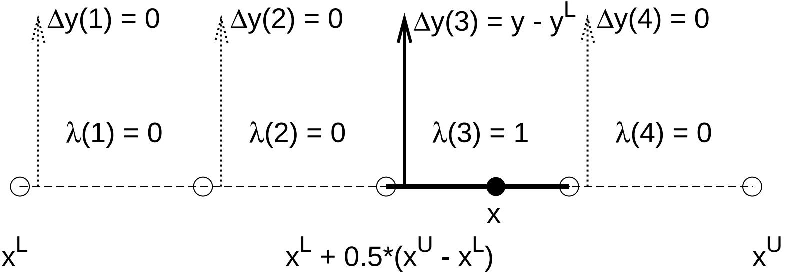


Underestimators with a Linear Number of Binary Vars [3/4]

Continuous switch variable $\Delta y(n_P)$ is $(y - y^L)$ when $\lambda(n_P) = 1$ and 0 else:

$$y = y^L + \sum_{n_P=1}^{N_P} \Delta y(n_P)$$

$$0 \leq \Delta y(n_P) \leq (y^U - y^L) \cdot \lambda(n_P) \quad \forall\, n_P \in \{1, \ldots, N_P\}$$



Activation of a Single $\Delta y(n_P)$ According to the Value of x ($N_P = 4$ and $x = x^L + 2.6 \cdot a$)

Underestimators with a Linear Number of Binary Vars [4/4]

$$z \geq x \cdot y^L + \sum_{n_P=1}^{N_P} \left[x^L + a \cdot (n_P - 1) \right] \cdot \ \Delta y(n_P)$$

$$z \geq x \cdot y^U + \sum_{n_P=1}^{N_P} \left[x^L + a \cdot \ \ n_P \qquad \right] \cdot \left[\Delta y(n_P) - (y^U - y^L) \cdot \lambda(n_P) \right]$$

$$z \leq x \cdot y^L + \sum_{n_P=1}^{N_P} \left[x^L + a \cdot \ \ n_P \qquad \right] \cdot \ \Delta y(n_P)$$

$$z \leq x \cdot y^U + \sum_{n_P=1}^{N_P} \left[x^L + a \cdot (n_P - 1) \right] \cdot \left[\Delta y(n_P) - (y^U - y^L) \cdot \lambda(n_P) \right]$$

$$x^L \leq x \leq x^U; \quad y^L \leq y \leq y^U$$

Logarithmic Number of Binary Variables [1/4]

Define a parameter N_L and 3 additional variable sets so that the number of continuous variables, binary variables, and constraint equations all scale logarithmically:

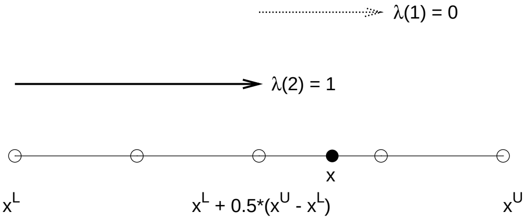
- Number of Logarithmic Binary Variables: $N_L = \lceil \log_2 N_P \rceil$
- Binary switch: $\lambda \in \{0,\, 1\}^{N_L}$
- Continuous switch: $\hat{\lambda} \in \{0,\, 1\}^{N_P}$
- Continuous switch: $\Delta y \in \left[0,\, y^U - y^L\right]^{N_P}$

Although we model the mapping from the partition containing x to λ using a base-2 representation, note that any injective function $B : \{1,\, \ldots,\, N_P\} \mapsto \{0,\, 1\}^{\lceil \log_2 N_P \rceil}$ could formulate the SOS1-like constraints for the activation of exactly one of the N_P segments:

$$x^L + \sum_{n_L=1}^{N_L} 2^{n_L-1} \cdot a \cdot \lambda(n_L) \le x \le x^L + a + \sum_{n_L=1}^{N_L} 2^{n_L-1} \cdot a \cdot \lambda(n_L)$$

Logarithmic Number of Binary Variables [2/4]

Activation of $\lambda(n_L)$ according when $N_P = 4$ and $x = x^L + 2.6 \cdot a$:



The figure illustrates the example case: $N_P = 4$ and $x = x^L + 2.6 \cdot a$.
Setting $\lambda(n_L = 1) = 1$ and $\lambda(n_L = 2) = 0$ results in the appropriate:

$$x \geq x^L + \sum_{n_L=1}^{N_L} 2^{n_L-1} a\lambda(n_L) = x^L + 2a\lambda(1) + a\lambda(2) = x^L + 2a$$

$$x \leq x^L + a + \sum_{n_L=1}^{N_L} 2^{n_L-1} a\lambda(n_L) = x^L + a + 2a\lambda(1) + a\lambda(2) = x^L + 3a$$

Logarithmic Number of Binary Variables [3/4]

The $\Delta y(n_P)$ variables are equal to $(y - y^L)$ for each active subdomain:

$$\sum_{n_P=1}^{N_P} \hat{\lambda}(n_P) = 1$$

$$\sum_{n_P : \lfloor \frac{n_P - 1}{2^{N_L - n_L}} \rfloor \bmod 2 = 0} \hat{\lambda}(n_P) \leq (1 - \lambda(n_L)) \quad \forall\, n_L \in \{1, \ldots, N_L\}$$

$$\sum_{n_P : \lfloor \frac{n_P - 1}{2^{N_L - n_L}} \rfloor \bmod 2 = 1} \hat{\lambda}(n_P) \leq \lambda(n_L) \quad \forall\, n_L \in \{1, \ldots, N_L\}$$

$$\Delta y(n_P) \leq (y^U - y^L) \cdot \hat{\lambda}(n_P) \quad \forall\, n_P \in \{1, \ldots, N_P\}$$

$$y = y^L + \sum_{n_P=1}^{N_P} \Delta y(n_P)$$

Logarithmic Number of Binary Variables [4/4]

$$z \geq x \cdot y^L + \sum_{n_P=1}^{N_P} \left[x^L + a \cdot (n_P - 1)\right] \cdot \ \Delta y(n_P)$$

$$z \geq x \cdot y^U + \sum_{n_P=1}^{N_P} \left[x^L + a \cdot \ \ n_P \qquad \right] \cdot \left[\Delta y(n_P) - (y^U - y^L) \cdot \hat{\lambda}(n_P)\right]$$

$$z \leq x \cdot y^L + \sum_{n_P=1}^{N_P} \left[x^L + a \cdot \ \ n_P \qquad \right] \cdot \ \Delta y(n_P)$$

$$z \leq x \cdot y^U + \sum_{n_P=1}^{N_P} \left[x^L + a \cdot (n_P - 1)\right] \cdot \left[\Delta y(n_P) - (y^U - y^L) \cdot \hat{\lambda}(n_P)\right]$$

$$x^L \leq x \leq x^U; \quad y^L \leq y \leq y^U$$

LP Relaxation of Logarithmic Partitioning Scheme nondominated by the convex hull? [1/4]

Relax $\lambda \in \{0,\, 1\}^{N_L}$ to $\lambda \in [0,\, 1]^{N_L}$:

$$
\begin{aligned}
z \quad & \geq x \cdot y^L + \sum_{n_P=1}^{N_P} \left[x^L + a \cdot (n_P - 1) \right] \cdot \Delta y(n_P) \\
& \overset{(1)}{=} x \cdot y^L + x^L \cdot y - x^L \cdot y^L + \sum_{n_P=1}^{N_P} a \cdot (n_P - 1) \cdot \Delta y(n_P) \\
& \overset{(2)}{\geq} x \cdot y^L + x^L \cdot y - x^L \cdot y^L
\end{aligned}
$$

Equality (1) holds because $\sum_{n_P=1}^{N_P} x^L \cdot \Delta y(n_P) = x^L \cdot y - x^L \cdot y^L$. Inequality (2) follows from $a \cdot (n_P - 1) \cdot \Delta y(n_P) \geq 0 \;\forall\; n_P \in \{1,\, \dots,\, N_P\}$.

LP Relaxation of Logarithmic Partitioning Scheme nondominated by the convex hull? [2/4]

Relax $\lambda \in \{0,\, 1\}^{N_L}$ to $\lambda \in [0,\, 1]^{N_L}$:

$$
\begin{aligned}
z &\geq xy^U + \sum_{n_P=1}^{N_P} \left[x^L + a\ \ n_P \qquad \right] \left[\Delta y(n_P) - (y^U - y^L)\hat{\lambda}(n_P)\right] \\
&\overset{(1)}{=} xy^U + x^U y - x^U y^U + \sum_{n_P=1}^{N_P} a(n_P - N_P) \left[\Delta y(n_P) - (y^U - y^L)\hat{\lambda}(n_P)\right] \\
&\overset{(2)}{\geq} xy^U + x^U y - x^U y^U
\end{aligned}
$$

Equality (1) follows from $x^L + an_P = x^U + a(n_P - N_P)$ and $\sum_{n_P=1}^{N_P} x^U \left[\Delta y(n_P) - (y^U - y^L)\hat{\lambda}(n_P)\right] = x^U y - x^U y^U$. Inequality (2) from: $a(n_P - N_P) \left[\Delta y(n_P) - (y^U - y^L)\hat{\lambda}(n_P)\right] \geq 0 \ \forall\ n_P \in \{1,\, \ldots,\, N_P\}$.

LP Relaxation of Logarithmic Partitioning Scheme nondominated by the convex hull? [3/4]

Relax $\lambda \in \{0,\, 1\}^{N_L}$ to $\lambda \in [0,\, 1]^{N_L}$:

$$
\begin{aligned}
z \;\; &\leq x \cdot y^L + \sum_{n_P=1}^{N_P} \left[x^L + a \cdot \;\; n_P \qquad \right] \cdot \Delta y(n_P) \\
&\overset{(1)}{=} x \cdot y^L + x^U \cdot y - x^U \cdot y^L + \sum_{n_P=1}^{N_P} a \cdot (n_P - N_P) \cdot \Delta y(n_P) \\
&\overset{(2)}{\leq} x \cdot y^L + x^U \cdot y - x^U \cdot y^L
\end{aligned}
$$

Equality (1) holds by the definition of a and $y - y^L = \sum_{n_P=1}^{N_P} \Delta y(n_P)$.

Inequality (2) from $a \cdot (n_P - N_P) \cdot \Delta y(n_P) \leq 0 \; \forall \; n_P \in \{1,\, \ldots,\, N_P\}$.

LP Relaxation of Logarithmic Partitioning Scheme nondominated by the convex hull? [4/4]

Relax $\lambda \in \{0, 1\}^{N_L}$ to $\lambda \in [0, 1]^{N_L}$:

$$\begin{aligned} z \;&\leq xy^U + \sum_{n_P=1}^{N_P} \left[x^L + a(n_P - 1)\right] \left[\Delta y(n_P) - (y^U - y^L)\hat{\lambda}(n_P)\right] \\ &\overset{(1)}{=} xy^U + x^L y - x^L y^U + \sum_{n_P=1}^{N_P} a(n_P - 1) \left[\Delta y(n_P) - (y^U - y^L)\hat{\lambda}(n_P)\right] \\ &\overset{(2)}{\leq} xy^U + x^L y - x^L y^U \end{aligned}$$

Equality (1) is a result of $y - y^L = \sum_{n_P=1}^{N_P} \Delta y(n_P)$ and $\sum_{n_P=1}^{N_P} \hat{\lambda}(n_P) = 1$.

Inequality (2) holds by $\Delta y(n_P) \leq (y^U - y^L) \cdot \hat{\lambda}(n_P)$.