Convex relaxations for quadrilinear terms

Sonia Cafieri

LIX, École Polytechnique

joint work with:

Jon Lee
IBM T.J. Watson Research Center, N.Y.

Leo Liberti
LIX, École Polytechnique

ARS Workshop, Oct 31st 2008

Outline

Outline

Motivations

- The convex envelopes of graphs of all monomials of degree 2 and 3 on an arbitrary box are explicitly known.
- Such a description is unknown, in general, for degree at least 4.
- Branch-and-Bound based global optimization methods, applied to formulations involving multivariate polynomials, rely on such convex envelopes.

Outline

- Background and Motivations
  - Motivations
  - Existing convex envelopes
- Convex relaxations for $x_1x_2x_3x_4$
  - Investigated convex relaxations
  - Computational assessment
  - Computational results
- Applications to known problems
  - MDGP and HFP problems
  - A bound evaluation algorithm
  - Computational results

Existing convex envelopes

- The bilinear term $x_j x_k$ is replaced by a new variable x_i, and the following linear inequalities are added to the problem relaxation:

$$\begin{array}{rcl} x_i & \geq & x_j^L x_k + x_k^L x_j - x_j^L x_k^L \\ x_i & \geq & x_j^U x_k + x_k^U x_j - x_j^U x_k^U \\ x_i & \leq & x_j^L x_k + x_k^U x_j - x_j^L x_k^U \\ x_i & \leq & x_j^U x_k + x_k^L x_j - x_j^U x_k^L \end{array}$$

(*McCormick's envelope*)

Existing convex envelopes

- The trilinear term $x_j x_k x_h$ is replaced by a variable x_i, and linear inequalities are added to the problem relaxation depending on the signs of the bounds on variables (Meyer & Floudas, 2004).

  e.g. case $x^L \geq 0, y^L \geq 0, z^L \leq 0, z^U \geq 0$: ($x, y, z$ permutation of x_j, x_k, x_h)

$$
\begin{array}{rcl}
x_i & \leq & y^U z^U x + x^U z^U y + x^U y^U z - 2x^U y^U z^U \\
x_i & \leq & y^U z^L x + x^L z^U y + x^L y^U z - x^L y^U z^L - x^L y^U z^U \\
x_i & \leq & y^U z^L x + x^L z^L y + x^L y^L z - x^L y^U z^L - x^L y^L z^L \\
x_i & \leq & y^L z^U x + x^U z^L y + x^U y^L z - x^U y^L z^U - x^U y^L z^L \\
x_i & \leq & y^L z^L x + x^U z^L y + x^L y^L z - x^U y^L z^L - x^L y^L z^L \\
x_i & \leq & y^L z^U x + x^L z^U y + (\theta/(z^U - z^L))z + (-(\theta z^L)/(z^U - z^L) - x^L y^U z^U - x^U y^L z^U + x^U y^U z^L) \\
x_i & \geq & y^U z^L x + x^U z^L y + x^U y^U z - 2x^U y^U z^L \\
x_i & \geq & y^L z^L x + x^U z^U y - x^U y^L z - x^U y^L z^U - x^U y^L z^L \\
x_i & \geq & y^U z^U x + x^L z^U y + x^L y^L z - x^L y^U z^U - x^L y^L z^U \\
x_i & \geq & y^U z^U x + x^L z^L y + x^L y^U z - x^L y^U z^U - x^L y^U z^L \\
x_i & \geq & y^L z^U x + x^U z^U y + x^L y^L z - x^U y^L z^U - x^L y^L z^U \\
x_i & \geq & y^L z^L x + x^L z^L y + (\bar{\theta}/(z^L - z^U))z + (-(\bar{\theta} z^U)/(z^L - z^U) - x^U y^L z^L - x^L y^U z^L + x^U y^U z^U),
\end{array}
$$

$$\theta := x^L y^U z^U - x^U y^U z^L - x^L y^L z^U + x^U y^L z^U, \quad \bar{\theta} := x^U y^L z^L - x^U y^U z^U - x^L y^L z^L + x^L y^U z^L.$$

Existing convex envelopes

- The concave univariate function $f(x_j)$ is replaced by a variable x_i and two inequalities are added to the problem relaxation:

$$
\begin{aligned}
x_i &\leq f(x_j) \\
x_i &\geq f(x_j^L) + \frac{f(x_j^U) - f(x_j^L)}{x_j^U - x_j^L}(x_j - x_j^L).
\end{aligned}
$$

- The convex univariate function $f(x_j)$ is replaced by a variable x_i and two inequalities are added to the problem relaxation:

$$
\begin{aligned}
x_i &\leq f(x_j^L) + \frac{f(x_j^U) - f(x_j^L)}{x_j^U - x_j^L}(x_j - x_j^L) \\
x_i &\geq f(x_j).
\end{aligned}
$$

For concave (convex) $f(x_j)$, the first (second) constraint is a nonlinear over (under)-estimator which is usually replaced by a pre-determined number of tangents to f at various given points.

Existing convex envelopes

- The concave univariate function $f(x_j)$ is replaced by a variable x_i and two inequalities are added to the problem relaxation:

$$\begin{aligned} x_i &\leq f(x_j) \\ x_i &\geq f(x_j^L) + \frac{f(x_j^U) - f(x_j^L)}{x_j^U - x_j^L}(x_j - x_j^L). \end{aligned}$$

- The convex univariate function $f(x_j)$ is replaced by a variable x_i and two inequalities are added to the problem relaxation:

$$\begin{aligned} x_i &\leq f(x_j^L) + \frac{f(x_j^U) - f(x_j^L)}{x_j^U - x_j^L}(x_j - x_j^L) \\ x_i &\geq f(x_j). \end{aligned}$$

For concave (convex) $f(x_j)$, the first (second) constraint is a nonlinear over (under)-estimator which is usually replaced by a pre-determined number of tangents to f at various given points.

Existing convex envelopes

- The term x_j^{2k} for any $k \in \mathbb{N}$ is replaced by a variable x_i and treated as a convex univariate function.
- The term x_j^{2k+1} for any $k \in \mathbb{N}$ is replaced by a variable x_i. If the range of x_j does not include 0, the function is convex or concave. Otherwise, the convex/concave envelope is given in (Liberti & Pantelides, 2003) – a tight linear relaxation is:

$$
\begin{array}{rcl}
(x_j^L)^{2k+1}\left(1+R_k\left(\frac{x_j}{x_j^L}-1\right)\right) & \le x_i \le & (x_j^U)^{2k+1}\left(1+R_k\left(\frac{x_j}{x_j^U}-1\right)\right) \\
(2k+1)(x_j^U)^{2k}x_j - 2k(x_j^U)^{2k+1} & \le x_j \le & (2k+1)(x_j^L)^{2k}x_j - 2k(x_j^L)^{2k+1},
\end{array}
$$

$$R_k = \frac{r_k^{2k+1}-1}{r_k-1}$$

k	r_k	k	r_k
1	-0.5000000000	6	-0.7721416355
2	-0.6058295862	7	-0.7921778546
3	-0.6703320476	8	-0.8086048979
4	-0.7145377272	9	-0.8223534102
5	-0.7470540749	10	-0.8340533676

Existing convex envelopes

- The term x_j^{2k} for any $k \in \mathbb{N}$ is replaced by a variable x_i and treated as a convex univariate function.
- The term x_j^{2k+1} for any $k \in \mathbb{N}$ is replaced by a variable x_i. If the range of x_j does not include 0, the function is convex or concave. Otherwise, the convex/concave envelope is given in (Liberti & Pantelides, 2003) – a tight linear relaxation is:

$$\begin{array}{rcl} (x_j^L)^{2k+1}\left(1+R_k\left(\frac{x_j}{x_j^L}-1\right)\right) & \leq x_i \leq & (x_j^U)^{2k+1}\left(1+R_k\left(\frac{x_j}{x_j^U}-1\right)\right) \\ (2k+1)(x_j^U)^{2k}x_j-2k(x_j^U)^{2k+1} & \leq x_j \leq & (2k+1)(x_j^L)^{2k}x_j-2k(x_j^L)^{2k+1}, \end{array}$$

$$R_k = \frac{r_k^{2k+1}-1}{r_k-1}$$

k	r_k	k	r_k
1	-0.5000000000	6	-0.7721416355
2	-0.6058295862	7	-0.7921778546
3	-0.6703320476	8	-0.8086048979
4	-0.7145377272	9	-0.8223534102
5	-0.7470540749	10	-0.8340533676

Outline

Obtaining convex relaxations

Basic idea: given a sufficiently rich set of “elementary” convex envelopes, compose convex relaxations (albeit not envelopes) of complex functions relatively easily.

Example: given $f(x), g(x)$ with known convex/concave envelopes, in order to obtain a convex relaxation for $f(x)g(x)$:

- apply the bilinear convex envelope to the product w_1w_2,
- replace the necessary “defining constraints”:

$$\begin{aligned} w_1 &= f(x) \\ w_2 &= g(x) \end{aligned}$$

by the convex/concave envelopes of f, g.

Note: this strategy may yields non-unique ways of combining terms (due to the associativity of the product).

Obtaining convex relaxations

Basic idea: given a sufficiently rich set of "elementary" convex envelopes, compose convex relaxations (albeit not envelopes) of complex functions relatively easily.

Example: given $f(x), g(x)$ with known convex/concave envelopes, in order to obtain a convex relaxation for $f(x)g(x)$:

- apply the bilinear convex envelope to the product w_1w_2,
- replace the necessary "defining constraints":

$$\begin{aligned} w_1 &= f(x) \\ w_2 &= g(x) \end{aligned}$$

by the convex/concave envelopes of f, g.

Note: this strategy may yields non-unique ways of combining terms (due to the associativity of the product).

Obtaining convex relaxations

Basic idea: given a sufficiently rich set of "elementary" convex envelopes, compose convex relaxations (albeit not envelopes) of complex functions relatively easily.

Example: given $f(x), g(x)$ with known convex/concave envelopes, in order to obtain a convex relaxation for $f(x)g(x)$:

- apply the bilinear convex envelope to the product w_1w_2,
- replace the necessary "defining constraints":

$$\begin{aligned} w_1 &= f(x) \\ w_2 &= g(x) \end{aligned}$$

by the convex/concave envelopes of f, g.

Note: this strategy may yields non-unique ways of combining terms (due to the associativity of the product).

Quadrilinear term: 3 convex relaxations

Given a quadrilinear term

$$\boxed{x_1x_2x_3x_4}$$

we consider the following three types of term grouping:

$$((x_1x_2)x_3)x_4$$
$$(x_1x_2)(x_3x_4)$$
$$(x_1x_2x_3)x_4$$

and derive three corresponding linear convex relaxations for $x_1x_2x_3x_4$.

Quadrilinear term: 3 convex relaxations

Let us consider:

$$
\begin{aligned}
S_1 &= \{(x,w) \in \mathbb{R}^4 \times \mathbb{R}^3 \,|\, x_i \in [x_i^L, x_i^U], w_1 = x_1x_2, w_2 = w_1x_3, w_3 = w_2x_4\} \\
S_2 &= \{(x,w) \in \mathbb{R}^4 \times \mathbb{R}^3 \,|\, x_i \in [x_i^L, x_i^U],\ w_1 = x_1x_2, w_2 = x_3x_4, w_3 = w_1w_2\} \\
S_3 &= \{(x,w) \in \mathbb{R}^4 \times \mathbb{R}^2 \,|\, x_i \in [x_i^L, x_i^U],\ w_1 = x_1x_2x_3, w_2 = w_1x_4\}
\end{aligned}
$$

S_1 : *bilinear envelope exploited thrice*

S_2 : *bilinear envelope exploited thrice*

S_3 : *bilinear envelope* + *trilinear envelope*

Which one yields the tightest bounds?

Quadrilinear term: 3 convex relaxations

Let us consider:

$$\begin{aligned}
S_1 &= \{(x,w) \in \mathbb{R}^4 \times \mathbb{R}^3 \,|\, x_i \in [x_i^L, x_i^U], w_1 = x_1x_2, w_2 = w_1x_3, w_3 = w_2x_4\} \\
S_2 &= \{(x,w) \in \mathbb{R}^4 \times \mathbb{R}^3 \,|\, x_i \in [x_i^L, x_i^U],\ w_1 = x_1x_2, w_2 = x_3x_4, w_3 = w_1w_2\} \\
S_3 &= \{(x,w) \in \mathbb{R}^4 \times \mathbb{R}^2 \,|\, x_i \in [x_i^L, x_i^U],\ w_1 = x_1x_2x_3, w_2 = w_1x_4\}
\end{aligned}$$

S_1 : *bilinear envelope exploited thrice*

S_2 : *bilinear envelope exploited thrice*

S_3 : *bilinear envelope* + *trilinear envelope*

Which one yields the tightest bounds?

Quadrilinear term: 3 convex relaxations

Let us consider:

$$\begin{aligned}
S_1 &= \{(x,w) \in \mathbb{R}^4 \times \mathbb{R}^3 \,|\, x_i \in [x_i^L, x_i^U], w_1 = x_1x_2, w_2 = w_1x_3, w_3 = w_2x_4\} \\
S_2 &= \{(x,w) \in \mathbb{R}^4 \times \mathbb{R}^3 \,|\, x_i \in [x_i^L, x_i^U],\ w_1 = x_1x_2, w_2 = x_3x_4, w_3 = w_1w_2\} \\
S_3 &= \{(x,w) \in \mathbb{R}^4 \times \mathbb{R}^2 \,|\, x_i \in [x_i^L, x_i^U],\ w_1 = x_1x_2x_3, w_2 = w_1x_4\}
\end{aligned}$$

S_1 : *bilinear envelope exploited thrice*

S_2 : *bilinear envelope exploited thrice*

S_3 : *bilinear envelope* + *trilinear envelope*

Which one yields the tightest bounds?

Outline

- Background and Motivations
  - Motivations
  - Existing convex envelopes
- Convex relaxations for $x_1x_2x_3x_4$
  - Investigated convex relaxations
  - Computational assessment
  - Computational results
- Applications to known problems
  - MDGP and HFP problems
  - A bound evaluation algorithm
  - Computational results

Test instances

- 80 test instances generated varying signs of the bounds/bound interval widths.
- 20 combinations by varying the signs of bounds on the 4 variables (missing cases are equivalent to covered cases by simple symmetry considerations).
- same initial width of the bound intervals for all variables. Then progressively, for $i = 1, 2, 3$, the width of the bound interval of x_i is reduced.

  This simulates the exploration of a single branch of a typical sBB search tree, whose nodes have decreasing range widths.

Comparison of relaxations

The comparison among the considered relaxations is made in terms of the **volume** of the corresponding enveloping polytopes.

Exploiting envelopes for bilinear and trilinear terms leads to an increased number of variables $\Longrightarrow$ the obtained polytopes belong to $\mathbb{R}^7$ and $\mathbb{R}^6$.

$\Longrightarrow$ **Projection** of the polytopes onto the space of $(x, f(x) := x_1x_2x_3x_4) \in \mathbb{R}^5$.

- Computation of the projections: `cdd` software (Fukuda, 2008).
- Computation of the volume of the projected polytopes: `lrs` software (Avis, 2006).

All the results are computed in exact arithmetic.

Comparison of relaxations

The comparison among the considered relaxations is made in terms of the **volume** of the corresponding enveloping polytopes.

Exploiting envelopes for bilinear and trilinear terms leads to an increased number of variables $\Longrightarrow$ the obtained polytopes belong to $\mathbb{R}^7$ and $\mathbb{R}^6$.

$\Longrightarrow$ **Projection** of the polytopes onto the space of $(x, f(x) := x_1x_2x_3x_4) \in \mathbb{R}^5$.

- Computation of the projections: `cdd` software (Fukuda, 2008).
- Computation of the volume of the projected polytopes: `lrs` software (Avis, 2006).

All the results are computed in exact arithmetic.

Outline

Results

Inst.	x_1		x_2		x_3		x_4		$((x_1x_2)x_3)x_4$	$(x_1x_2)(x_3x_4)$	$(x_1x_2x_3)x_4$
	$\|\|\cdot\|\|$	*sign*	$\|\|\cdot\|\|$	*sign*	$\|\|\cdot\|\|$	*sign*	$\|\|\cdot\|\|$	*sign*			
`inst1`	2	$+,+$	2	$+,+$	2	$-,+$	2	$-,+$	8282/45	**1474/15**	1508117/8640
`inst2`	2	$+,+$	2	$+,+$	2	$-,+$	2	$-,-$	10922/45	298793/1215	**1928777/8640**
`inst3`	2	$+,+$	2	$-,+$	2	$-,+$	2	$-,+$	2080/27	2080/27	**3136/45**
`inst4`	2	$+,+$	2	$-,+$	2	$-,+$	2	$-,-$	3424/27	3056/15	**4576/45**
`inst5`	2	$-,+$	2	$-,+$	2	$-,+$	2	$-,+$	*416/15*	*416/15*	*416/15*
`inst6`	2	$-,+$	2	$-,+$	2	$-,+$	2	$-,-$	**736/15**	2080/27	**736/15**
`inst7`	2	$+,+$	2	$-,+$	2	$-,-$	2	$-,+$	1664/9	3056/15	**4736/27**
`inst8`	2	$+,+$	2	$-,+$	2	$-,-$	2	$-,-$	736/3	298793/1215	**6032/27**
`inst9`	2	$-,+$	2	$-,-$	2	$-,-$	2	$-,+$	1664/9	3056/15	**4736/27**
`inst10`	2	$-,+$	2	$-,-$	2	$-,-$	2	$-,-$	736/3	298793/1215	**6032/27**
`inst11`	2	$-,+$	2	$-,+$	2	$-,-$	2	$-,+$	**3136/45**	2080/27	3728/45
`inst12`	2	$-,+$	2	$-,+$	2	$-,-$	2	$-,-$	4576/45	**1474/15**	6608/45
`inst13`	2	$+,+$	2	$+,+$	2	$+,+$	2	$-,+$	40166/195	298793/1215	**38288/195**
`inst14`	2	$+,+$	2	$+,+$	2	$+,+$	2	$-,-$	53686/195	359936/1215	**48688/195**
`inst15`	2	$+,+$	2	$+,+$	2	$-,-$	2	$-,+$	40166/195	298793/1215	**38288/195**
`inst16`	2	$+,+$	2	$+,+$	2	$-,-$	2	$-,-$	53686/195	359936/1215	**48688/195**
`inst17`	2	$+,+$	2	$-,-$	2	$-,-$	2	$-,+$	40166/195	298793/1215	**38288/195**
`inst18`	2	$+,+$	2	$-,-$	2	$-,-$	2	$-,-$	53686/195	359936/1215	**48688/195**
`inst19`	2	$-,-$	2	$-,-$	2	$-,-$	2	$-,+$	40166/195	298793/1215	**4983841/44928**
`inst20`	2	$-,-$	2	$-,-$	2	$-,-$	2	$-,-$	53686/195	359936/1215	**48688/195**

Results

Inst.	x_1		x_2		x_3		x_4		$((x_1 x_2) x_3) x_4$	$(x_1 x_2)(x_3 x_4)$	$(x_1 x_2 x_3) x_4$
	$\lVert \cdot \rVert$	*sign*	$\lVert \cdot \rVert$	*sign*	$\lVert \cdot \rVert$	*sign*	$\lVert \cdot \rVert$	*sign*			
inst21	1	$+,+$	2	$+,+$	2	$-,+$	2	$-,+$	48958/675	**1886/45**	47746613/691200
inst22	1	$+,+$	2	$+,+$	2	$-,+$	2	$-,-$	63358/675	21847/225	**59852213/691200**
inst23	1	$+,+$	2	$-,+$	2	$-,+$	2	$-,+$	11368/375	11368/375	**2128/75**
inst24	1	$+,+$	2	$-,+$	2	$-,+$	2	$-,-$	6056/125	781148/10125	**3128/75**
inst25	1	$-,+$	2	$-,+$	2	$-,+$	2	$-,+$	*104/15*	*104/15*	*104/15*
inst26	1	$-,+$	2	$-,+$	2	$-,+$	2	$-,-$	**184/15**	520/27	**184/15**
inst27	1	$+,+$	2	$-,+$	2	$-,-$	2	$-,+$	81008/1125	781148/10125	**15584/225**
inst28	1	$+,+$	2	$-,+$	2	$-,-$	2	$-,-$	104408/1125	7503097/81000	**19484/225**
inst29	1	$-,+$	2	$-,-$	2	$-,-$	2	$-,+$	416/9	764/15	**1184/27**
inst30	1	$-,+$	2	$-,-$	2	$-,-$	2	$-,-$	184/3	298793/4860	**1508/27**
inst31	1	$-,+$	2	$-,+$	2	$-,-$	2	$-,+$	**784/45**	520/27	932/45
inst32	1	$-,+$	2	$-,+$	2	$-,-$	2	$-,-$	1144/45	**737/30**	1652/45
inst33	1	$+,+$	2	$+,+$	2	$+,+$	2	$-,+$	**8842/105**	21847/225	30404/315
inst34	1	$+,+$	2	$+,+$	2	$+,+$	2	$-,-$	**11362/105**	695674/6075	50144/315
inst35	1	$+,+$	2	$+,+$	2	$-,-$	2	$-,+$	8842/105	21847/225	**25364/315**
inst36	1	$+,+$	2	$+,+$	2	$-,-$	2	$-,-$	11362/105	695674/6075	**31244/315**
inst37	1	$+,+$	2	$-,-$	2	$-,-$	2	$-,+$	8842/105	21847/225	**25364/315**
inst38	1	$+,+$	2	$-,-$	2	$-,-$	2	$-,-$	11362/105	695674/6075	**31244/315**
inst39	1	$-,-$	2	$-,-$	2	$-,-$	2	$-,+$	8842/105	21847/225	**458469/5600**
inst40	1	$-,-$	2	$-,-$	2	$-,-$	2	$-,-$	**11362/105**	695674/6075	50144/315

Results

Inst.	x_1		x_2		x_3		x_4		$((x_1x_2)x_3)x_4$	$(x_1x_2)(x_3x_4)$	$(x_1x_2x_3)x_4$
	$\|\|\cdot\|\|$	*sign*	$\|\|\cdot\|\|$	*sign*	$\|\|\cdot\|\|$	*sign*	$\|\|\cdot\|\|$	*sign*			
inst41	1	$+,+$	1	$+,+$	2	$-,+$	2	$-,+$	335651/12000	**8453/480**	161590649/6144000
inst42	1	$+,+$	1	$+,+$	2	$-,+$	2	$-,-$	421651/12000	874021/24000	**196456649/6144000**
inst43	1	$+,+$	1	$-,+$	2	$-,+$	2	$-,+$	2842/375	2842/375	**532/75**
inst44	1	$+,+$	1	$-,+$	2	$-,+$	2	$-,-$	1514/125	195287/10125	**782/75**
inst45	1	$-,+$	1	$-,+$	2	$-,+$	2	$-,+$	*26/15*	*26/15*	*26/15*
inst46	1	$-,+$	1	$-,+$	2	$-,+$	2	$-,-$	**46/15**	130/27	**46/15**
inst47	1	$+,+$	1	$-,+$	2	$-,-$	2	$-,+$	20252/1125	195287/10125	**3896/225**
inst48	1	$+,+$	1	$-,+$	2	$-,-$	2	$-,-$	26102/1125	7503097/324000	**4871/225**
inst49	1	$-,+$	1	$-,-$	2	$-,-$	2	$-,+$	20252/1125	195287/10125	**3896/225**
inst50	1	$-,+$	1	$-,-$	2	$-,-$	2	$-,-$	26102/1125	7503097/324000	**4871/225**
inst51	1	$-,+$	1	$-,+$	2	$-,-$	2	$-,+$	**196/45**	130/27	233/45
inst52	1	$-,+$	1	$-,+$	2	$-,-$	2	$-,-$	286/45	**737/120**	413/45
inst53	1	$+,+$	1	$+,+$	2	$+,+$	2	$-,+$	**47921/1440**	874021/24000	3961/99
inst54	1	$+,+$	1	$+,+$	2	$+,+$	2	$-,-$	**59201/1440**	56957/1350	6568/99
inst55	1	$+,+$	1	$+,+$	2	$-,-$	2	$-,+$	47921/1440	874021/24000	**15757/495**
inst56	1	$+,+$	1	$+,+$	2	$-,-$	2	$-,-$	59201/1440	56957/1350	**18727/495**
inst57	1	$+,+$	1	$-,-$	2	$-,-$	2	$-,+$	47921/1440	874021/24000	**969001783/35371875**
inst58	1	$+,+$	1	$-,-$	2	$-,-$	2	$-,-$	59201/1440	56957/1350	**368725761/11790625**
inst59	1	$-,-$	1	$-,-$	2	$-,-$	2	$-,+$	47921/1440	874021/24000	**3307195027/243302400**
inst60	1	$-,-$	1	$-,-$	2	$-,-$	2	$-,-$	**59201/1440**	56957/1350	6568/99

Results

Inst.	x_1		x_2		x_3		x_4		$((x_1x_2)x_3)x_4$	$(x_1x_2)(x_3x_4)$	$(x_1x_2x_3)x_4$
	$\|\|\cdot\|\|$	*sign*	$\|\|\cdot\|\|$	*sign*	$\|\|\cdot\|\|$	*sign*	$\|\|\cdot\|\|$	*sign*			
`inst61`	1	$+,+$	1	$+,+$	1	$-,+$	2	$-,+$	335651/48000	**8453/1920**	161590649/24576000
`inst62`	1	$+,+$	1	$+,+$	1	$-,+$	2	$-,-$	421651/48000	874021/96000	**196456649/24576000**
`inst63`	1	$+,+$	1	$-,+$	1	$-,+$	2	$-,+$	1421/750	1421/750	**133/75**
`inst64`	1	$+,+$	1	$-,+$	1	$-,+$	2	$-,-$	757/250	195287/40500	**391/150**
`inst65`	1	$-,+$	1	$-,+$	1	$-,+$	2	$-,+$	*13/30*	*13/30*	*13/30*
`inst66`	1	$-,+$	1	$-,+$	1	$-,+$	2	$-,-$	**23/30**	65/54	**23/30**
`inst67`	1	$+,+$	1	$-,+$	1	$-,-$	2	$-,+$	12851/1875	1733/240	**20/3**
`inst68`	1	$+,+$	1	$-,+$	1	$-,-$	2	$-,-$	10609/1250	3203327/360000	**97/12**
`inst69`	1	$-,+$	1	$-,-$	1	$-,-$	2	$-,+$	12851/1875	1733/240	**20/3**
`inst70`	1	$-,+$	1	$-,-$	1	$-,-$	2	$-,-$	10609/1250	3203327/360000	**97/12**
`inst71`	1	$-,+$	1	$-,+$	1	$-,-$	2	$-,+$	**133/75**	1421/750	641/300
`inst72`	1	$-,+$	1	$-,+$	1	$-,-$	2	$-,-$	**391/150**	943/360	1141/300
`inst73`	1	$+,+$	1	$+,+$	1	$+,+$	2	$-,+$	1162283/94080	1673383477/129600000	**34879/2940**
`inst74`	1	$+,+$	1	$+,+$	1	$+,+$	2	$-,-$	1377883/94080	24832097/1620000	**39779/2940**
`inst75`	1	$+,+$	1	$+,+$	1	$-,-$	2	$-,+$	1162283/94080	1673383477/129600000	**34879/2940**
`inst76`	1	$+,+$	1	$+,+$	1	$-,-$	2	$-,-$	1377883/94080	24832097/1620000	**39779/2940**
`inst77`	1	$+,+$	1	$-,-$	1	$-,-$	2	$-,+$	1162283/94080	1673383477/129600000	**34879/2940**
`inst78`	1	$+,+$	1	$-,-$	1	$-,-$	2	$-,-$	1377883/94080	24832097/1620000	**39779/2940**
`inst79`	1	$-,-$	1	$-,-$	1	$-,-$	2	$-,+$	1162283/94080	1673383477/129600000	**45232093/516096000**
`inst80`	1	$-,-$	1	$-,-$	1	$-,-$	2	$-,-$	1377883/94080	24832097/1620000	**39779/2940**

Outline

Molecular Distance Geometry Problem

- The MDGP is the problem of finding an embedding in $\mathbb{R}^3$ of a weighted graph $G = (V, E)$ such that all Euclidean distances between points in the embedding are the same as the corresponding edge weights in the graph.

  The *main application* is to find the 3-dimensional structure of a molecule given a subset of the atomic distances.

- Given a set V of n *atoms*, a set E of *inter-atomic distances* $d_{ij} = d(\{i,j\})$ for $\{i,j\} \in E$, a NLP formulation of the MDGP is:

$$\min_x f(x)$$

$$f(x) = \sum_{\{i,j\}\in E} (||x_i - x_j||^2 - d_{ij}^2)^2.$$

Molecular Distance Geometry Problem

- The MDGP is the problem of finding an embedding in $\mathbb{R}^3$ of a weighted graph $G = (V, E)$ such that all Euclidean distances between points in the embedding are the same as the corresponding edge weights in the graph.

  The *main application* is to find the 3-dimensional structure of a molecule given a subset of the atomic distances.

- Given a set V of n *atoms*, a set E of *inter-atomic distances* $d_{ij} = d(\{i,j\})$ for $\{i,j\} \in E$, a NLP formulation of the MDGP is:

$$\min_x f(x)$$

$$f(x) = \sum_{\{i,j\}\in E} (||x_i - x_j||^2 - d_{ij}^2)^2.$$

Molecular Distance Geometry Problem

A typical term $(||x_i - x_j||^2 - d_{ij}^2)^2$ expanded is:

(each 3-vector x_i has components (x_{i1}, x_{i2}, x_{i3}))

$$
\begin{aligned}
(||x_i - x_j||^2 - d_{ij}^2)^2 = \quad & - \quad 4\, x_{i2}\, x_{j2}\, x_{i3}^2 - 4\, x_{i1}\, x_{j1}\, x_{i2}^2 - 4\, x_{i1}\, x_{j1}\, x_{j2}^2 - 4\, x_{i1}\, x_{j1}\, x_{i3}^2 + 4\, x_{i1}\, x_{j1}\, d_{ij}^2 \\
& - \quad 4\, x_{i1}^2\, x_{i2}\, x_{j2} - 4\, x_{j1}^2\, x_{i2}\, x_{j2} - 4\, x_{i2}^2\, x_{i3}\, x_{j3} - 4\, x_{i1}\, x_{j1}\, x_{j3}^2 - 4\, x_{j1}^2\, x_{i3}\, x_{j3} \\
& + \quad 4\, x_{i2}\, x_{j2}\, d_{ij}^2 - 4\, x_{i2}\, x_{j2}\, x_{j3}^2 + 4\, x_{i3}\, x_{j3}\, d_{ij}^2 - 4\, x_{j2}^2\, x_{i3}\, x_{j3} - 4\, x_{i1}^2\, x_{i3}\, x_{j3} \\
& + \quad 8\, x_{i1}\, x_{j1}\, x_{i2}\, x_{j2} + 8\, x_{i1}\, x_{j1}\, x_{i3}\, x_{j3} + 8\, x_{i2}\, x_{j2}\, x_{i3}\, x_{j3} + x_{i1}^4 + x_{j1}^4 + x_{i2}^4 \\
& + \quad x_{j2}^4 + x_{i3}^4 + x_{j3}^4 - 4\, x_{i1}^3\, x_{j1} + 6\, x_{i1}^2\, x_{j1}^2 + 2\, x_{i1}^2\, x_{i2}^2 + 2\, x_{i1}^2\, x_{j2}^2 \\
& + \quad 2\, x_{i1}^2\, x_{i3}^2 + 2\, x_{i1}^2\, x_{j3}^2 - 2\, x_{i1}^2\, d_{ij}^2 - 4\, x_{i1}\, x_{j1}^3 + 2\, x_{j1}^2\, x_{i2}^2 + 2\, x_{j1}^2\, x_{j2}^2 \\
& + \quad 2\, x_{j1}^2\, x_{i3}^2 + 2\, x_{j1}^2\, x_{j3}^2 - 2\, x_{j1}^2\, d_{ij}^2 - 4\, x_{i2}^3\, x_{j2} + 6\, x_{i2}^2\, x_{j2}^2 + 2\, x_{i2}^2\, x_{i3}^2 \\
& + \quad 2\, x_{i2}^2\, x_{j3}^2 - 2\, x_{i2}^2\, d_{ij}^2 - 4\, x_{i2}\, x_{j2}^3 + 2\, x_{j2}^2\, x_{i3}^2 + 2\, x_{j2}^2\, x_{j3}^2 - 2\, x_{j2}^2\, d_{ij}^2 \\
& - \quad 4\, x_{i3}^3\, x_{j3} + 6\, x_{i3}^2\, x_{j3}^2 - 2\, x_{i3}^2\, d_{ij}^2 - 4\, x_{i3}\, x_{j3}^3 - 2\, x_{j3}^2\, d_{ij}^2 + d_{ij}^4.
\end{aligned}
$$

many quartic terms!

Hartree-Fock Problem

It is a known problem in quantum chemistry: finding spatial orbitals of electrons in a closed-shell atomic system.

- Non-relativistic time-independent Schrödinger equation: $H\Psi_n = E\Psi_n$ (H = Hamiltonian operator of the system, representing the total energy).
- HF model: the electrons in atoms and molecules move independently of each other, the motion of each one of the electrons being determined by the attractive electrostatic potential of the nuclei and by a repulsive average field due to all the other electrons of the system.
- The approximate solutions Φ_n of the Schrödinger equation are products of *orbitals* $\{\varphi_i\}$, which are solutions of the HF equations.
- Orbitals approximated by suitable bases $\{\chi_s \mid s \leq b\}$:

$$\bar{\varphi}_i := \sum_{s \leq b} c_{si} \chi_s$$

$\bar{\varphi}_i$ approximations of φ_i.

Hartree-Fock Problem

It is a known problem in quantum chemistry: finding spatial orbitals of electrons in a closed-shell atomic system.

- Non-relativistic time-independent Schrödinger equation: $H\Psi_n = E\Psi_n$ (H = Hamiltonian operator of the system, representing the total energy).
- HF model: the electrons in atoms and molecules move independently of each other, the motion of each one of the electrons being determined by the attractive electrostatic potential of the nuclei and by a repulsive average field due to all the other electrons of the system.
- The approximate solutions Φ_n of the Schrödinger equation are products of *orbitals* $\{\varphi_i\}$, which are solutions of the HF equations.
- Orbitals approximated by suitable bases $\{\chi_s \mid s \leq b\}$:

$$\bar{\varphi}_i := \sum_{s \leq b} c_{si} \chi_s$$

$\bar{\varphi}_i$ approximations of φ_i.

Hartree-Fock Problem

It is a known problem in quantum chemistry: finding spatial orbitals of electrons in a closed-shell atomic system.

- Non-relativistic time-independent Schrödinger equation: $H\Psi_n = E\Psi_n$ (H = Hamiltonian operator of the system, representing the total energy).
- HF model: the electrons in atoms and molecules move independently of each other, the motion of each one of the electrons being determined by the attractive electrostatic potential of the nuclei and by a repulsive average field due to all the other electrons of the system.
- The approximate solutions Φ_n of the Schrödinger equation are products of *orbitals* $\{\varphi_i\}$, which are solutions of the HF equations.
- Orbitals approximated by suitable bases $\{\chi_s \mid s \leq b\}$:

$$\bar{\varphi}_i := \sum_{s \leq b} c_{si} \chi_s$$

$\bar{\varphi}_i$ approximations of φ_i.

Hartree-Fock Problem

- HFP: finding a set of coefficients c_{si} such that the $\bar{\varphi}_i$ are the best possible approximations of the spatial orbitals.
- NLP problem: minimize a suitable energy function (quality of the approximation) s.t. $\{\bar{\varphi}_i\}$ is an orthonormal set:

$$
\begin{array}{ll}
\min & E(c) \\
\text{s.t.} & \left\langle \sum_{s \le b} c_{si}\chi_s \, , \, \sum_{s \le b} c_{sj}\chi_s \right\rangle = \delta_{ij} \quad \forall i \le j \le n \\
& c^L \le c \le c^U \, .
\end{array}
$$

orthonormality constraints are quadratic in the decision variables c .

Hartree-Fock Problem

- HFP: finding a set of coefficients c_{si} such that the $\bar{\varphi}_i$ are the best possible approximations of the spatial orbitals.
- NLP problem: minimize a suitable energy function (quality of the approximation) s.t. $\{\bar{\varphi}_i\}$ is an orthonormal set:

$$
\begin{aligned}
\min \quad & E(c) \\
\text{s.t.} \quad & \left\langle \sum_{s \le b} c_{si}\chi_s \, , \, \sum_{s \le b} c_{sj}\chi_s \right\rangle = \delta_{ij} \quad \forall i \le j \le n \\
& c^L \le c \le c^U \, .
\end{aligned}
$$

orthonormality constraints are quadratic in the decision variables c .

Outline

Bound evaluation algorithm

- The natural application of tight lower bounds computed through a convex relaxation is within the sBB algorithm.
- Our alternative: a simplified *partial sBB* algorithm.
  At each branching step, the algorithm only records the most promising node and discards the other, thus *exploring a single branch up to a leaf*.
- A very simple branching strategy (the variable index i maximizing $|x_i^* - \bar{x}_i|$); termination: either on iteration limit or on reaching a node that is infeasible or that contains the global optimum.

Bound evaluation algorithm

- The natural application of tight lower bounds computed through a convex relaxation is within the sBB algorithm.
- Our alternative: a simplified *partial sBB* algorithm.
  At each branching step, the algorithm only records the most promising node and discards the other, thus *exploring a single branch up to a leaf*.
- A very simple branching strategy (the variable index i maximizing $|x_i^* - \bar{x}_i|$); termination: either on iteration limit or on reaching a node that is infeasible or that contains the global optimum.

Constructing the convex relaxation

`solver_RQuarticConvex` within ROSE

- **Step1:** replace each nonlinear term by an added variable
- **Step2:** add a defining constraint “added variable = nonlinear term” to the problem
- **Step 3:** replace each defining constraint by a convex relaxation.

Note: The 3 different convex relaxations yielded by the different defining constraints due to the different associativity precedences in

$$((x_1x_2)x_3)x_4 \quad (x_1x_2)(x_3x_4) \quad (x_1x_2x_3)x_4$$

are implemented.

Outline

Computational results

Instance	$((x_1x_2)x_3)x_4$	$(x_1x_2)(x_3x_4)$	$(x_1x_2x_3)x_4$
`lavor5`	-1580.81	-1758.4	**-683.82**
`lavor6`	-2652.05	-2746.92	**-1117.05**
`lavor7`	-3427.96	-3411.96	**-2378.53**
`beryllium`	**-22.6887**	-21.8208	-17.988
`neon`	-1292.64	-1306.4	**-1342.71**

Conclusions

- We proposed a computational approach to determine the best among three convex relaxations for quadrilinear terms.
- The results suggest that a strategy taking into account widths and signs of the input bounds could be used in a sBB solver in order to automatically select the best relaxation procedure.

Future work

- Computational experiences on large scale instances.
- Implementation of a full sBB.
- Theoretical study on the relaxation strenght of quadrilinear terms.

Conclusions

- We proposed a computational approach to determine the best among three convex relaxations for quadrilinear terms.
- The results suggest that a strategy taking into account widths and signs of the input bounds could be used in a sBB solver in order to automatically select the best relaxation procedure.

Future work

- Computational experiences on large scale instances.
- Implementation of a full sBB.
- Theoretical study on the relaxation strenght of quadrilinear terms.

Conclusions

- We proposed a computational approach to determine the best among three convex relaxations for quadrilinear terms.
- The results suggest that a strategy taking into account widths and signs of the input bounds could be used in a sBB solver in order to automatically select the best relaxation procedure.

Future work

- Computational experiences on large scale instances.
- Implementation of a full sBB.
- Theoretical study on the relaxation strenght of quadrilinear terms.