Comparison of Convex Relaxations for Monomials of Odd Degree

LEO LIBERTI

DEI, Politecnico di Milano, P.zza L. Da Vinci, 20133 Milano, Italy

(liberti@elet.polimi.it)

13 June 2002

Abstract

Branch-and-Bound is one of the most widely used technique for performing global optimization of Nonlinear Programs (NLPs) and Mixed-Integer Nonlinear Programs (MINLPs). In most implementations of this technique the lower bound is obtained by solving a convex relaxation of the original problem. As shown in [9, 10, 11] the convex relaxation can be formed by replacing the nonconvex terms with their convex relaxations. Tight convex underestimators are already available for many types of nonconvex terms, including bilinear and trilinear products, linear fractional terms, and concave and convex univariate functions. However, terms which are piecewise concave and convex are not explicitly catered for. A frequently occurring example of such a term is the monomial of odd degree. We propose a novel convex nonlinear envelope for odd power terms of the form x^{2k+1} ($k \in \mathbb{N}$), where $x \in [a, b]$ and $a < 0 < b$; the derived relaxation is continuous and differentiable everywhere in $[a, b]$. We then make a comparison with an existing convex relaxation for odd power monomials where we show that the novel convex relaxation gains better results in a Branch-and-Bound implementation.

1 Introduction

In this paper, we address NLPs in the following formulation:

$$\left.\begin{array}{ccc} \min_x & f(x) & \\ \alpha & \leq g(x) \leq & \beta \\ a & \leq x \leq & b \end{array}\right\} \tag{1}$$

where x are the problem variables, α, β are the lower and upper bounds of the constraints, a, b are the lower and upper bounds of the variables. f is the (possibly nonlinear) objective function and g is a set of (possibly nonlinear) constraints.

The most effective deterministic methods for the global solution of MINLPs in the general formulation (1) are spatial Branch-and-Bound algorithms. The whole search space is iteratively partitioned into progressively smaller regions. For each region, lower and upper bounds to the objective functions are computed, and the lowest upper bound is stored as the current best solution to the problem. A region is fathomed and discarded when the bounds are close enough or when the objective function lower bound for the region is higher than the current best solution. The algorithm terminates when there are no more regions to examine, at which point the current best solution is the solution of the problem.

One of the most critical steps in the algorithm is the generation of the lower bound: a good lower bound means rapid fathoming of regions, and as a consequence a lot less regions to examine during the algorithm run. There are many methods to provide a lower bound to the objective function. Lagrangian relaxation, either applied directly [4, 1] or as part of an altogether different algorithm (e.g. Generalized Benders Decomposition), relies on a reformulation of the problem so that it becomes box-constrained:

$$\left.\begin{array}{ll} \max_{u,v} & \inf\{f(x) + u(g-\beta) + v(\alpha - g) \mid a \leq x \leq b\} \\ & u, v \geq 0 \end{array}\right\} \tag{2}$$

Another method to provide lower bounds (esp. with respect to bilinear terms) is the RLT (Reformulation-Linearization Technique), proposed in [8]. By this method, new constraints are obtained by multiplying existing linear constraints by problem variables and other problem constraints and substituting the resulting bilinearities so that the new constraint becomes linear. On solving this linear problem to optimality, we find a lower bound to the original problem. One of the most common methods for providing lower bounds, however, is that of forming and locally solving a convex relaxation of the original problem [7, 2, 11]. The fact that the relaxation is convex implies that the local solution is also the global one, and hence a guaranteed underestimator of the objective function value.

The convex relaxation of any MINLP can be formed by isolating the nonconvex terms and replacing them with the respective convex relaxations. Tight convex underestimators are already available for many types of nonconvex term, including bilinear and trilinear products, linear fractional terms, and concave and convex univariate functions. However, terms which are piecewise concave and convex are not explicitly catered for. A frequently occurring example of such a term is x^{2k+1}, where $k \in \mathbb{N}$ and the range of x includes zero. In this article, we propose two convex relaxations for such terms: one nonlinear and one linear. We then briefly compare these relaxations to another possible relaxation of the monomial of odd degree, based on reformulating it to a bilinear product, and we show that our proposed relaxation is tighter. Finally, we present some numerical results that highlight this point.

2 Novel Convex Relaxation

Consider the monomial x^{2k+1} in the range $x \in [a, b]$ where $a < 0 < b$. Let c, d be the x-coordinates of the points C, D where the tangents from points A and B respectively meet the curve (see fig. 1 below). The shape of the convex underestimator of x^{2k+1} depends on the relative magnitude of b and c. In particular, if $c < b$ (as in fig. 1), a convex underestimator can be formed from the tangent from $x = a$ to $x = c$ followed by the curve x^{2k+1} from $x = c$ to $x = b$. On the other hand, if $c > b$ (cf. fig. 2), a convex underestimator is the straight line passing through A and B (similarly for the concave overestimator of x^{2k+1} in the range $x \in [a, b]$). If $d > a$, the overestimator is given by the upper tangent from B to D followed by the curve x^{2k+1} from D to A, as in fig. 1. If $d < a$ the overestimator is the straight line from A and B. Note that onditions $c > b$ and $d < a$ cannot both hold simultaneously.

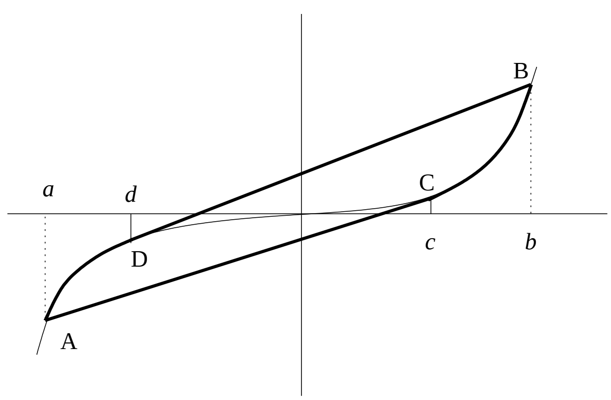

Figure 1: Tightest (nonlinear) convex envelope of x^{2k+1}.

In order to build the convex envelope of x^{2k+1} we therefore require the equations of the tangents that pass through points A, C and B, D. We do this by equating the slope of the line $\overline{AC}$ to the gradient of x^{2k+1} at $x = c$:

$$\frac{c^{2k+1} - a^{2k+1}}{c - a} = (2k+1)c^{2k} \tag{3}$$

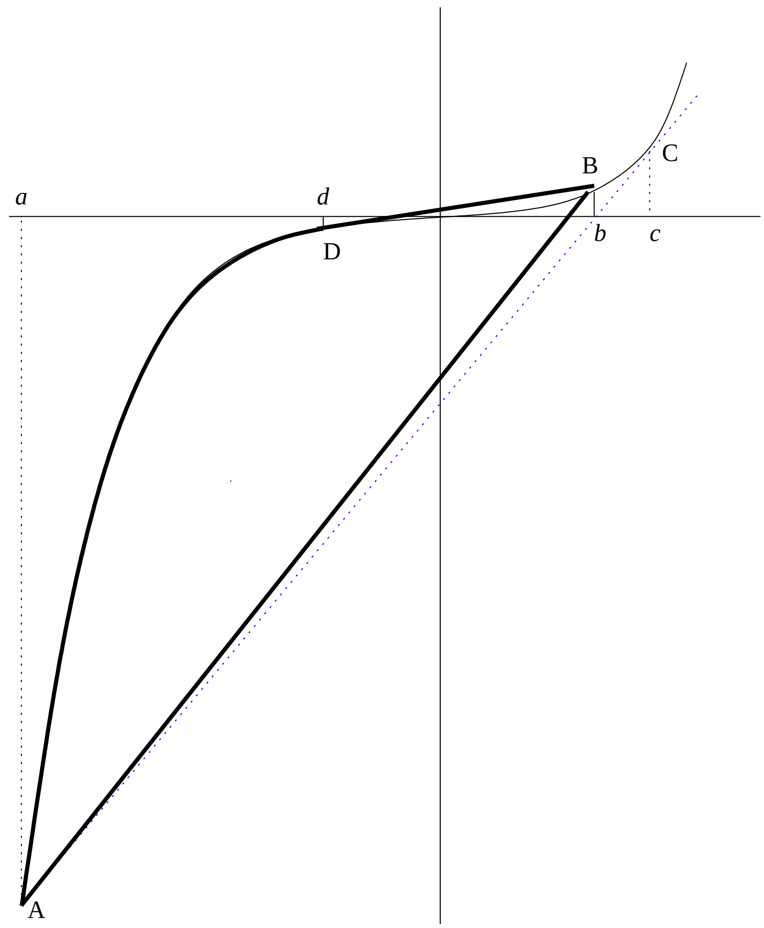

Figure 2: The case when $c > b$.

Hence c is a root of the polynomial:

$$P^k(x,a) \equiv (2k)x^{2k+1} - a(2k+1)x^{2k} + a^{2k+1} \tag{4}$$

It can be shown by induction on k that:

$$P^k(x,a) = a^{2k-1}(x-a)^2 Q^k(\frac{x}{a}) \tag{5}$$

where the polynomial $Q^k(x)$ is defined as:

$$Q^k(x) \equiv 1 + \sum_{i=2}^{2k} i x^{i-1}. \tag{6}$$

Thus, the roots of $P^k(x,a)$ can be obtained from the roots[1] of $Q^k(x)$. However, generally, polynomials of degree higher than 4 cannot be solved by radicals. In fact the Galois group of $Q^3(x) \equiv 6x^5 + 5x^4 + 4x^3 + 3x^2 + 2x + 1$ is isomorphic to S_5 which is not solvable. For details on Galois theory and the solvability of polynomials, see [12]. Notice further that $Q^k(x)$ does not depend on the range of x; moreover, $Q^k(x)$ has exactly one real root, r_k, for any $k \geq 1$, and this lies in $[-1 + 1/2k, -0.5]$ [5]. Hence, the roots of $Q^k(x)$ for different k can be computed *a priori* to arbitrary precision using simple numerical schemes (e.g. bisection). A table of these roots is presented in table 1 for $k \leq 10$.

[1] Although $P^k(x,a)$ also has $x = a$ as a root, this is not of practical interest.

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

Table 1: Numerical values of the roots of $Q^k(x)$ for $k = 1, .., 10$ (to 10 significant digits).

If the roots shown in the second column of table 1 are denoted by r_k, then the tangent points c and d in fig. 1 are simply $c = r_k a$ and $d = r_k b$. The lower and upper tangent lines are given respectively by:

$$a^{2k+1} + \frac{c^{2k+1} - a^{2k+1}}{c - a}(x - a) \tag{7}$$

$$b^{2k+1} + \frac{d^{2k+1} - b^{2k+1}}{d - b}(x - b) \tag{8}$$

Hence, the convex envelope for $z = x^{2k+1}$ when $x \in [a, b]$ and $a < 0 < b$:

$$l_k(x) \leq \quad z \quad \leq u_k(x) \tag{9}$$

is as follows:

- If $c < b$, then:

$$l_k(x) = \begin{cases} a^{2k+1}\left(1 + R_k\left(\frac{x}{a} - 1\right)\right) & \text{if } x < c \\ x^{2k+1} & \text{if } x \geq c \end{cases} \tag{10}$$

  otherwise:

$$l_k(x) = a^{2k+1} + \frac{b^{2k+1} - a^{2k+1}}{b - a}(x - a) \tag{11}$$

- If $d > a$, then:

$$u_k(x) = \begin{cases} x^{2k+1} & \text{if } x \leq d \\ b^{2k+1}\left(1 + R_k\left(\frac{x}{b} - 1\right)\right) & \text{if } x > d \end{cases} \tag{12}$$

  otherwise:

$$u_k(x) = a^{2k+1} + \frac{b^{2k+1} - a^{2k+1}}{b - a}(x - a) \tag{13}$$

where we have used the constant $R_k \equiv \frac{r_k^{2k+1} - 1}{r_k - 1}$.

We can relax the nonlinear convex envelope to a linear one by dropping the "follow the curve" requirements on either side of the tangency points, and using the lower and upper tangent as convex underestimator and concave overestimator respectively, as follows:

$$a^{2k+1}\left(1 + R_k\left(\frac{x}{a} - 1\right)\right) \quad \leq z \leq \quad b^{2k+1}\left(1 + R_k\left(\frac{x}{b} - 1\right)\right) \tag{14}$$

We can tighten the envelope further by also considering the tangents to the curve at the endpoints A, B:

$$(2k+1)b^{2k}x - 2kb^{2k+1} \quad \leq z \leq \quad (2k+1)a^{2k}x - 2ka^{2k+1}. \tag{15}$$

A similar idea for producing the convex envelope of x^3 was discussed in [7] (p. 159), but the treatment was limited to one demonstrative example.

3 Alternative Convex Relaxation

One common way to derive a convex relaxation for the monomial of odd degree x^{2k+1} is to reformulate it to wx and add a constraint $w = x^{2k}$. The bilinear term $z = wx$ can then be relaxed via the McCormick's convex under- and over-estimators [7]

$$\begin{array}{rcl} z & \geq & w^L x + x^L w - w^L x^L \\ z & \geq & w^U x + x^U w - w^U x^U \\ z & \leq & w^U x + x^L w - w^U x^L \\ z & \leq & w^L x + x^U w - w^L x^U. \end{array}$$

where $w^L \leq w \leq w^U, x^L \leq x \leq x^U$; whereas the term x^{2k} is convex and hence an envelope is given by the function itself and the secant. Thus the reformulation is:

$$\begin{array}{rcl} z & = & wx \\ w & = & x^{2k} \\ a \leq & x & \leq b \\ 0 \leq & w & \leq w^U = \max\{a^{2k}, b^{2k}\} \end{array}$$

and the convex relaxation is:

$$\begin{array}{rcl} aw \leq & z & \leq bw \\ w^U x + bw - w^U b \leq & z & \leq w^U x + aw - w^U a \\ x^{2k} \leq & w & \leq a^{2k} + \dfrac{b^{2k} - a^{2k}}{b - a}(x - a) \\ a \leq & x & \leq b \\ 0 \leq & w & \leq w^U \end{array}$$

After some algebraic manipulation, we can eliminate w to obtain the following nonlinear convex envelope for z:

$$\frac{w^U a}{a - b}(x - b) \leq z \leq \frac{w^U b}{b - a}(x - a) \tag{16}$$

$$bx^{2k} + w^U(x - b) \leq z \leq ax^{2k} + w^U(x - a) \tag{17}$$

$$a\left(a^{2k} + \frac{b^{2k} - a^{2k}}{b - a}(x - a)\right) \leq z \leq b\left(a^{2k} + \frac{b^{2k} - a^{2k}}{b - a}(x - a)\right) \tag{18}$$

4 Numerical Results

We have generated some numerical results that illustrate the difference between the two convex envelopes described above for x^{2k+1}. We solved the problem

$$\left.\begin{array}{rl} \min_{x,y} & x - y \\ y & = x^{2k+1} \\ -2 & \leq x, y \leq 2 \end{array}\right\} \tag{19}$$

to global optimality using the spatial Branch-and-Bound algorithm described in [11] within the *ooOPS* implementation [6], both with the novel convex (linear) relaxation and with the alternative convex relaxation. Table 2 lists the number of iterations taken by the algorithm when k varies. The first column lists

k	Iterations (novel rel.)	Iterations (alternative rel.)	k	Iterations (novel rel.)	Iterations (alternative rel.)
1	11	43	10	15	59
2	7	63	9	11	59
3	19	55	8	19	59
4	15	59	11	19	59
5	11	63	12	19	9
6	11	55	13	15	9
7	19	55	14	15	59

Table 2: Numerical results from the simple test problem.

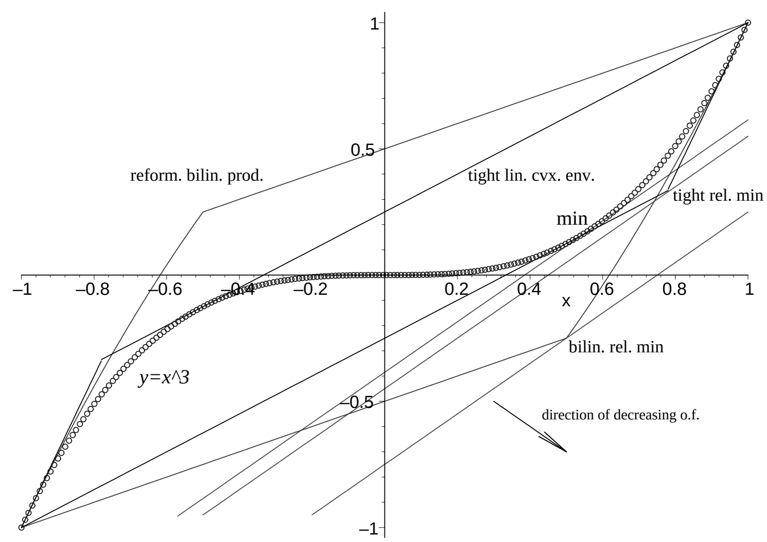

Figure 3: Graphical description of simple test problem (in 2D).

the results relative to the novel convex relaxation (section 2), the second those relative to the alternative convex relaxation (section 3).

The results clearly substantiate the theory: the novel convex relaxation gives better performance in comparison to the alternative relaxation based on reformulation to bilinear product. In the case of x^3 we can see why this happens in fig. 3, 4. The direction of minimization of the objective function $x - y$ is such that the minimum over the original feasible region is very near the minimum over the novel convex relaxation. However, the minimum over the alternative convex relaxation is further away. Hence the performance gain.

5 Conclusion

We have proposed a tight novel convex envelope for monomials of odd degree when the range of the defining variable includes zero, i.e. when they are piecewise convex and concave. We have then compared it with another possible relaxation (based on reformulation to bilinear product) and shown that the

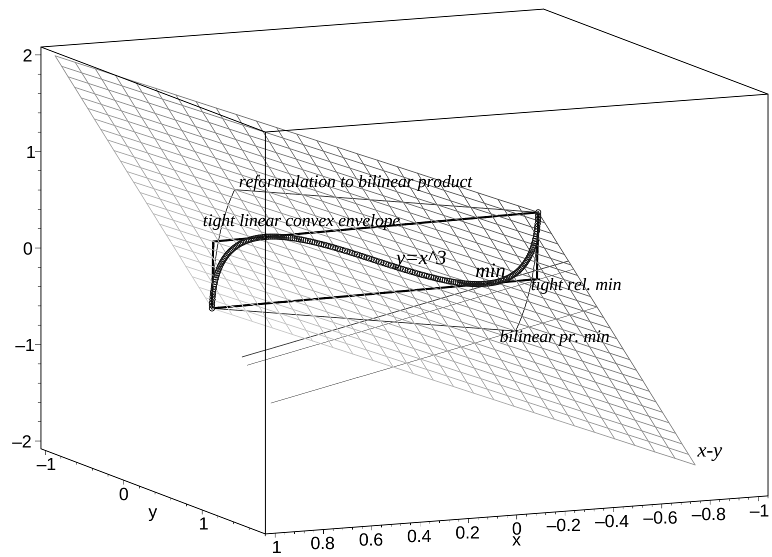

Figure 4: Graphical description of simple test problem (in 3D).

former performs better than the latter when tested in a Branch-and-Bound algorithm.

References

[1] N. Adhya, M. Tawarmalani, and N.V. Sahinidis. A lagrangian approach to the pooling problem. *Industrial and Engineering Chemistry Research*, 38:1956–1972, 1999.

[2] 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 & Chemical Engineering*, 22(9):1137–1158, 1998.

[3] M.S. Bazaraa and C.M. Shetty. *Nonlinear Programming*. Wiley, Chichester, 1979.

[4] A.Ben-Tal, G.Eiger, and V.Gershovitz. Global minimization by reducing the duality gap. *Mathematical Programming*, 63:193–212, 1994.

[5] L. Liberti and C.C. Pantelides. Tightest convex envelopes of monomials of odd degree. *Journal of Global Optimization*, *To appear*.

[6] L. Liberti, P. Tsiakis, B. Keeping, and C.C. Pantelides. *oo$\mathcal{OPS}$*. Centre for Process Systems Engineering Technical Report, Chemical Engineering Department, Imperial College, London, UK, 2002.

[7] G.P. McCormick. Computability of global solutions to factorable nonconvex programs: Part I — Convex underestimating problems. *Mathematical Programming*, 10:146–175, 1976.

[8] H.D. Sherali. Global optimization of nonconvex polynomial programming problems having rational exponents. *Journal of Global Optimization*, 12:267–283, 1998.

[9] E.M. Smith. *On the Optimal Design of Continuous Processes*. PhD thesis, Imperial College of Science, Technology and Medicine, University of London, October 1996.

[10] E.M. Smith and C.C. Pantelides. Global optimisation of nonconvex MINLPs. *Computers and Chemical Engineering*, 21:S791–S796, 1997.

[11] E.M. Smith and C.C. Pantelides. A symbolic reformulation/spatial branch-and-bound algorithm for the global optimisation of nonconvex MINLPs. *Computers and Chemical Engineering*, 23:457–478, 1999.

[12] I. Stewart. *Galois Theory*. Chapman & Hall, London, 2nd edition, 1989.