Comparing Formulations for Piecewise Convex Problems

Renan Spencer Trindade[1] [*], Claudia D'Ambrosio[1], Antonio Frangioni[2], Claudio Gentile[3]

[1] LIX, CNRS, École Polytechnique, Institut Polytechnique de Paris, Palaiseau, France
{rst,dambrosio}@lix.polytechnique.fr

[2] Dipartimento di Informatica, Università di Pisa, Pisa, Italy
frangio@di.unipi.it

[3] Istituto di Analisi dei Sistemi ed Informatica "Antonio Ruberti", Consiglio Nazionale delle Ricerche, Rome, Italy
claudio.gentile@iasi.cnr.it

Abstract

In this paper we address non-convex Mixed-Integer Non-Linear Programs where the non-convexity is manifested as the sum of non-convex univariate functions. Motivated by the Sequential Convex Mixed Integer Non Linear Programming technique, we compare the three classical different formulations for piecewise problems: the *incremental model*, the *multiple choice model*, and the *convex combination model*. For piecewise-*linear* functions, these models are known to be equivalent. We show that this is not the case for piecewise-*convex* functions, where one of the three formulations is weaker than the other two. Computational results on a target application illustrate the practical impact of this property.

Keywords : *Global optimization, Non-convex separable functions, Sequential Convex MINLP technique.*

1 Introduction

Mixed-Integer Nonlinear Programs (MINLPs) have been more in the focus of researchers in recent years, due to their ability to model an extremely wide variety of real-world applications. However, solving practical MINLPs to global optimality, especially non-convex ones, remains very challenging. It is therefore of paramount importance to exploit all structural properties of the MINLPs at hand. In [3], the Sequential Convex Mixed Integer Non Linear Programming (SC-MINLP) technique has been defined for MINLPs where the non-convexity is manifested as the sum of non-convex univariate functions. This work, motivated by applications in air traffic management, focuses on the same class problems, i.e.,

$$\min \textstyle\sum_{j\in N} c_j x_j \tag{1}$$

$$f_i(x) + \textstyle\sum_{j\in H(i)} g_{ij}(x_j) \le 0 \qquad i \in M \tag{2}$$

$$l_j \le x_j \le u_j \qquad j \in N \tag{3}$$

$$x_j \in \mathbb{Z} \qquad j \in I. \tag{4}$$

The sets M, N, $I \subseteq N$, and $H(i) \subseteq N$ are finite. The functions $f_i : \mathbb{R}^n \to \mathbb{R}$ are convex and multivariate. The functions $g_{ij} : \mathbb{R} \to \mathbb{R}$ are non-convex univariate. We assume that l_j and

[*]This research has been supported by the Chair "Integrated Urban Mobility", sponsored by Uber and backed by L'X- École Polytechnique and La Fondation de l'École Polytechnique

u_j are finite bounds for x_j that appear in g_{ij} functions. Even in the case where there are no integer variables, i.e., $I = \emptyset$, the problem remains $\mathcal{NP}$-hard.

The SC-MINLP technique deals with this problem by computing a piecewise-convex relaxation of each $g_{ij}(x_j)$, which can be strengthened with the Perspective Reformulation technique [2]. That approach uses the well-known *Incremental Model* (IM) to formulate the piecewise-convex functions. In this work, we focus on alternative models for these, in particular the well-known *Convex-Combination* (CCM) and *Multiple Choice* (MCM) ones. These are typically used interchangeably for Mixed-Integer *Linear* Programs, since they have similar size and provide the same lower bound [1]. However, this is *not* the case for *nonlinear* piecewise-convex programs.

2 Piecewise models for Sequential Convex MINLP technique

The SC-MINLP framework computes the breakpoints $l_j = l_{ij}^1 < l_{ij}^2 < \ldots < l_{ij}^{s(ij)} < l_{ij}^{s(ij)+1} = u_j$ where the non-convex functions g_{ij} change convexity/concavity. In practice, this is done by computing by the zeros of the second derivative of g_{ij} using some algebraic package, such as MATLAB. Then, for fixed i and $j \in H(i)$, we denote by $S_{ij} = \{\, s \,:\, g_{ij}$ is *concave* in the sub-interval $[l_{ij}^s, l_{ij}^{s+1}]\,\}$, and $\check{S}(ij) = \{\, s \,:\, g_{ij}$ is *convex* in the sub-interval $[l_{ij}^s, l_{ij}^{s+1}]\,\}$. On S_{ij} the function is substituted with its the best possible convex relaxation (a linear function), while the convex parts are kept as they are. This defines a *convex* MINLP, whose continuous relaxation therefore provides valid lower bounds. However, this reformulation step can actually be done in different ways. In [3, 2], the following Incremental Model is used.

2.1 Incremental Model (IM)

The IM introduces a segment load variable, x_{ij}^s, for each sub-interval $[l_{ij}^s, l_{ij}^{s+1}]$. Feasibility requires that the value on x_{ij}^{s+1} be zero unless x_{ij}^s is "full," that is, $x_{ij}^{s+1} > 0$ only if $x_{ij}^s = l_{ij}^{s+1} - l_{ij}^s$. For this purpose the IM introduces binary variables, y_{ij}^s, defined by the condition that $y_{ij}^s = 1$ if $x_{ij}^s > 0$, and $y_{ij}^s = 0$ otherwise, finally yielding

$$\min \textstyle\sum_{j \in N} c_j x_j \tag{5}$$

$$\bar{f}_i(x) + \textstyle\sum_{j \in H(i)} \sum_{s \in \check{S}(ij)} z_{ij}^s \leq 0 \qquad i \in M \tag{6}$$

$$z_{ij}^s \geq g_{ij}(l_{ij}^s + x_{ij}^s) - g_{ij}(l_{ij}^s) \qquad s \in \check{S}(ij),\ j \in H(i),\ i \in M \tag{7}$$

$$x_j = l_j + \textstyle\sum_{s \in S(ij)} x_{ij}^s \qquad j \in H(i), i \in M \tag{8}$$

$$(l_{ij}^{s+1} - l_{ij}^s) y_{ij}^{s+1} \leq x_{ij}^s \leq (l_{ij}^{s+1} - l_{ij}^s) y_{ij}^s \qquad s \in S(ij),\ j \in H(i),\ i \in M \tag{9}$$

$$y_{ij}^s \in \{0, 1\} \qquad s \in S(ij),\ j \in H(i),\ i \in M \tag{10}$$

$$x_j \in \mathbb{Z} \qquad j \in I \tag{11}$$

where $\bar{f}_i = f_i(x) + \sum_{j \in H(i)} g_{ij}(l_{ij}^1) + \sum_{s \in \hat{S}(ij)} \alpha_{ij}^s x_{ij}^s$. Clearly, $\bar{f}_i$ is convex since f_i was.

2.2 Multiple Choice Model (MCM)

The MCM is an alternative definition of the segment variables. The load variable x_{ij}^s, for each segment s, defines the total load $x_{ij}^s = x_j$ and $y_{ij}^{s+1} = 1$, if x_j lies on the sub-interval $[l_{ij}^s, l_{ij}^{s+1}]$. Otherwise, $x_{ij}^s = y_{ij}^{s+1} = 0$. In this formulation, at most one y_{ij}^{s+1} will equal one.

$$\min \sum_{j\in N} c_j x_j \tag{12}$$

$$\bar{f}_i(x) + \sum_{j\in H(i)} \sum_{s\in \check{S}(ij)} z_{ij}^s \le 0 \qquad i \in M \tag{13}$$

$$z_{ij}^s \ge g_{ij}(x_{ij}^s) - g_{ij}(0) \qquad s \in \check{S}(ij),\ j \in H(i),\ i \in M \tag{14}$$

$$x_j = \sum_{s\in S(ij)} x_{ij}^s \qquad j \in H(i), i \in M \tag{15}$$

$$l_{ij}^s y_{ij}^s \le x_{ij}^s \le l_{ij}^{s+1} y_{ij}^s \qquad s \in S(ij),\ j \in H(i),\ i \in M \tag{16}$$

$$\sum_{s\in S(ij)} y_{ij}^s = 1 \qquad i \in M,\ j \in H(i) \tag{17}$$

$$y_{ij}^s \in \{0,1\} \qquad s \in S(ij),\ j \in H(i),\ i \in M \tag{18}$$

$$x_j \in \mathbb{Z} \qquad j \in I \tag{19}$$

Again, $\bar{f}_i = f_i(x) + \sum_{j\in H(i)} g_{ij}(0) \sum_{s\in \check{S}(ij)} y_{ij}^s + \sum_{s\in \hat{S}(ij)} (\alpha_{ij}^s x_{ij}^s + (g_{ij}(l_{ij}^s) - \alpha_{ij}^s l_{ij}^s) y_{ij}^s)$ is clearly convex since f_i was.

2.3 Convex Combination Model (CCM)

As in the MCM, this formulation considers the variable that x_{ij}^s defines the total load, i.e., $x_{ij}^s = x_j$ if x_j lies on the sub-interval $[l_{ij}^s, l_{ij}^{s+1}]$. However, the load and its cost are computed as a convex combination of the load/cost of the two endpoints of the segment. By defining multipliers μ_{ij}^s and λ_{ij}^s as the weights of these two endpoints, this yields

$$\min \sum_{j\in N} c_j x_j \tag{20}$$

$$\bar{f}_i(x) + \sum_{j\in H(i)} \sum_{s\in \check{S}(ij)} z_{ij}^s \le 0 \qquad i \in M \tag{21}$$

$$z_{ij}^s \ge g_{ij}(l_{ij}^s \mu_{ij}^s + l_{ij}^{s+1} \lambda_{ij}^s) - g_{ij}(0) \qquad s \in \check{S}(ij),\ j \in H(i),\ i \in M \tag{22}$$

$$x_j = \sum_{s\in S(ij)} (l_{ij}^s \mu_{ij}^s + l_{ij}^{s+1} \lambda_{ij}^s) \qquad j \in H(i), i \in M \tag{23}$$

$$\mu_{ij}^s + \lambda_{ij}^s = y_{ij}^s \qquad s \in S(ij),\ j \in H(i),\ i \in M \tag{24}$$

$$\sum_{s\in S(ij)} y_{ij}^s = 1 \qquad i \in M,\ j \in H(i) \tag{25}$$

$$y_{ij}^s \in \{0,1\} \qquad s \in S(ij),\ j \in H(i),\ i \in M \tag{26}$$

$$x_j \in \mathbb{Z} \qquad j \in I \tag{27}$$

where $\bar{f}_i = f_i(x) + \sum_{j\in H(i)} g_{ij}(0) \sum_{s\in \check{S}(ij)} (\mu_{ij}^s + \lambda_{ij}^s) + \sum_{s\in \hat{S}(ij)} (g_{ij}(l_{ij}^s)\mu_{ij}^s + g_{ij}(l_{ij}^{s+1})\lambda_{ij}^s)$. Note that writing "$g_{ij}(0) \sum_{s\in \check{S}(ij)} (\mu_{ij}^s + \lambda_{ij}^s)$" instead of "$g_{ij}(0) \sum_{s\in \check{S}(ij)} y_{ij}^s$" is a bit of a gimmick that changes nothing, except making the notation little bit more symmetric.

Computional Results

Several real-world applications can be modeled as a problem in the class of (1)-(4), among which aircraft traffic control problems, uncapacitated facility location problem, unit commitment and scheduling problem, and, more in general, the non linear knapsack problem. For lack of space, we focus on the latter and compare the three formulations on a set of 40 instances. All the instances were solved with BARON 19.7.13.

The non linear knapsack problem is the same considered in [2], i.e.,

$$
\begin{aligned}
&\max \textstyle\sum_{j\in N} p_j \\
&\quad p_j \le \frac{c_j}{1 + b_j \exp(-a_j(x_j + d_j))} && j \in N \\
&\quad \textstyle\sum_{j\in N} x_j \le C \\
&\quad 0 \le x_j \le U_j && j \in N
\end{aligned}
$$

For each value of $|N| \in \{10, 20, 50, 100\}$ we randomly generated 10 instances, where $a_j \in [0.1, 0.2]$, $b_j \in [0, 100]$, $c_j \in [0, 100]$, and $d_j \in [-100, 0]$ were uniformly drawn in the corresponding intervals. We fixed $U_j = 100$ for all $j \in N$ and $C = 100|N|/2$.

In the following table, we present, for each value of $|N|$, the average results for the non-convex MINLP formulation ("Original Solution" in the table) and the three convex MINLP relaxations presented in the previous sections. For each of the convex MINLP formulations, we present the objective function of the best solution found within the time limit of 120 seconds both for the convex MINLP ("Integer" in the table) and its continuous relaxation ("Relax" in the table). Clearly, the best solution value of the non-convex MINLP problem is smaller than the solution of the convex MINLP relaxations (which is smaller than the solution of its continuous relaxation). Note also that the "Integer" values of the three convex MINLP relaxation should be the same, as they are equivalent formulations. It might not be true when the time limit is hit, see, for example, for $|N| = 50$ and $|N| = 100$.

The most interesting thing to notice is the continuous relaxation value for the three convex MINLP formulations ("Relax"). In general, the continuous relaxations of MCM and CCM are equivalent while IM is worse (unless the time limit is reach). However, this does not always imply better CPU times. We plan to study this aspect more extensively both from a computational and a theoretical viewpoint.

TAB. 1: Computational results for non linear knapsack problem.

Instance	Original	Incremental		Multiple Choice		Convex Combination	
Items	Solution	Integer	Relax	Integer	Relax	Integer	Relax
10	193.372	196.034	205.313	196.035	201.773	196.035	201.773
20	250.601	251.077	274.085	251.079	259.449	251.079	259.449
50	337.435	337.997	394.567	337.652	344.921	338.001	345.098
100	422.279	422.597	504.163	422.162	429.720	414.255	418.741
Average	299.985	301.001	343.249	300.811	308.044	298.864	305.269

References

[1] Keely L. Croxton, Bernard Gendron, and Thomas L. Magnanti. A Comparison of Mixed-Integer Programming Models for Nonconvex Piecewise Linear Cost Minimization Problems. *Management Science*, 49(9):1268–1273, 2003.

[2] Claudia D'Ambrosio, Antonio Frangioni, and Claudio Gentile. Strengthening the sequential convex MINLP technique by perspective reformulations. *Optimization Letters*, 13(4):673–684, jun 2019.

[3] Claudia D'Ambrosio, Jon Lee, and Andreas Wächter. An algorithmic framework for minlp with separable non-convexity. In Jon Lee and Sven Leyffer, editors, *Mixed Integer Nonlinear Programming*, pages 315–347, New York, NY, 2012. Springer New York.