Operations Research Letters 41 (2013) 627–632

Contents lists available at ScienceDirect

Operations Research Letters

journal homepage: www.elsevier.com/locate/orl

Locally ideal formulations for piecewise linear functions with indicator variables

Srikrishna Sridhar [a], Jeff Linderoth [b], James Luedtke [b,*]

[a] School of Computer Sciences, University of Wisconsin-Madison, 1209 Dayton Street, Madison, WI 53706, United States

[b] Department of Industrial and Systems Engineering, University of Wisconsin-Madison, 1513 University Avenue, Madison, WI 53706, United States

ARTICLE INFO

Article history:
Received 12 April 2013
Received in revised form
19 August 2013
Accepted 20 August 2013
Available online 29 August 2013

Keywords:
Piecewise linear function
Mixed-integer programming
Indicator variables

ABSTRACT

In this paper, we consider mixed integer linear programming (MIP) formulations for piecewise linear functions (PLFs) that are evaluated when an indicator variable is turned on. We describe modifications to standard MIP formulations for PLFs with desirable theoretical properties and superior computational performance in this context.

1. Introduction

Optimization problems involving piecewise linear functions (PLFs) appear in a wide range of applications. PLFs are frequently used to approximate nonlinear functions and to model cost functions involving economies of scale and fixed charges. Problems involving non-convex PLFs are commonly formulated as mixed integer programming (MIP) problems [4,17,8,2,23].

Consider a univariate PLF $f : [B_0, B_n] \to \mathbb{R}$ with its domain $[B_0, B_n]$ divided into an increasing sequence of breakpoints $\{B_0, B_1, \dots, B_n\}$. For simplicity, we assume that $f(\cdot)$ is continuous, $B_0 = 0$ and $f(0) = 0$. Our results can be extended to the case when $f(\cdot)$ is lower semi-continuous, $B_0 \neq 0$, and $f(B_0) \neq 0$. The function $f(\cdot)$ can be written as

$$f(x) := m_i x + c_i, \quad x \in [B_{i-1}, B_i] \ \forall i \in \{1, \dots, n\} \tag{1}$$

where $m_i \in \mathbb{R}$, $c_i \in \mathbb{R}$ and $B_0 < B_1 < \cdots < B_n$.

In this paper, we present MIP formulations for PLFs where setting a binary indicator variable to zero forces the argument of the function of $f(\cdot)$ to zero which in turn forces the function to take a zero value. In other words,

$$z = 0 \Rightarrow x = 0, f(x) = 0. \tag{2}$$

The goal of this work is to present a theoretical and computational comparison of MIP formulations that enforce the logical conditions in (2). Specifically, we examine properties of different formulations of the three variable set

$$X := \bigcup_{i=1}^{n} \{(x, y, z) : x \in [B_{i-1}, B_i],\ y = m_i x + c_i, z = 1\} \bigcup \{(0, 0, 0)\}. \tag{3}$$

In some applications, notably those where the PLF appears in a minimization objective, the relevant set to study has the variable y constrained to lie in the epigraph of a function. We denote $X^{\geq}$ as the set where the equality relationship $y = m_i x + c_i$ in (3) is replaced with $y \geq m_i x + c_i$.

Methods for modeling PLFs include specially ordered sets of type II (SOS2) [4], the incremental model, or delta method (Delta) [17], the multiple choice model (MCM) [13], the convex combination (CC) model [8], the disaggregated convex combination model (DCC) [19], and approaches that require only logarithmically many binary variables [24]. Table 1 lists several applications in the literature that have modeled PLFs using these well-known methods in conjunction with variable upper bound constraints of the form

$$x \leq B_n z \tag{4}$$

to enforce the logical on–off condition (2).

In this work, we propose a simple modeling artifice for PLFs that also enforces the logical condition (2), and we demonstrate its desirable theoretical and computational properties. We start by describing the idea using SOS2 to model a PLF as

$$x = \sum_{i=0}^{n} \lambda_i B_i, \qquad y = \sum_{i=0}^{n} \lambda_i F_i \qquad 1 = \sum_{i=0}^{n} \lambda_i \tag{5}$$

$$\boldsymbol{\lambda} := \left\{\lambda_i \in \mathbb{R}_+ : \forall i \in \{0, \dots, n\}\right\} \text{ is SOS2.}$$

* Corresponding author.
E-mail addresses: srikris@cs.wisc.edu (S. Sridhar), linderot@cae.wisc.edu (J. Linderoth), jrluedt1@wisc.edu (J. Luedtke).

0167-6377/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.orl.2013.08.010

Table 1
Applications using PLFs with indicator variables.

Ref.	Application	Model
[18]	Gas network optimization	SOS2
[1]	Transmissions expansion planning	Delta
[12]	Oil field development	CC
[6]	Thermal unit commitment	Delta
[16]	Sales resource allocation	MCM

In this formulation, the function $f(\cdot)$ and its argument x are expressed as convex combinations of breakpoints $\boldsymbol{B} := \{B_0, \ldots, B_n\}$ and their corresponding function evaluations $\{F_0, \ldots, F_n\}$ where $F_i := f(B_i) = m_iB_i + c_i$. The formulation introduces a non-negative set of variables $\boldsymbol{\lambda} \in \mathbb{R}^{n+1}$ that satisfy the SOS2 property—at most two of the variables can be positive, and if two variables are positive then they must be consecutive in the ordered set. Most modern general purpose MIP solvers enforce the SOS2 condition algorithmically by branching [4].

Using variable upper bound constraints (4) to enforce the logical condition (2) has two problems. First, the use of "bigM" constraints may considerably weaken the LP relaxation of the MIP formulation. Second, the model introduces an additional constraint $x \le B_nz$.

We propose the following simple strengthening that replaces $x \le B_nz$ and $\sum_{i=0}^{n} \lambda_i = 1$ with

$$\sum_{i=0}^{n} \lambda_i = z. \tag{6}$$

Setting the binary variable $z = 0$ in (6) forces $\lambda_i = 0\ \forall i \in \{0, \ldots, n\}$, which in turn forces the function to take a zero value. If the binary variable $z = 1$, then $\sum_{i=0}^{n} \lambda_i = 1$, which reduces to (5). We show in Section 2.1 that a formulation using (6) has the desirable property of being *locally ideal*, while one that uses $x \le B_nz$ does not.

In Section 2, we also show how to strengthen MIP formulations of X that use the incremental model, the multiple choice model, the convex combination model, the disaggregated convex combination model, and logarithmic models to model the PLF. Therefore, this formulation strengthening technique could be directly applied to all of the applications listed in Table 1. In all cases, we show that our model retains the desirable theoretical property of the underlying PLF modeling method, either idealness or sharpness, but using a variable upper bound constraint $x \le B_nz$ destroys the property. Borghetti et al. [5] created a formulation of X that employed the strengthening techniques we describe. They used the convex combination method to model the PLFs which does not have the locally ideal property [23]. In the case that the PLFs are convex, we describe a connection between the formulation strengthening techniques we describe and the *perspective reformulation* [11]. The Delta, MCM, CC, and DCC MIP formulations can all be extended to model multivariate piecewise linear functions [23]. When the multivariate piecewise linear functions are combined with the structure (2), our strengthened formulation can be similarly applied. We omit the details to simplify exposition. We conclude with a computational study on a practical application to illustrate the benefits of the new formulations. In our experiments, we observed that our formulation computes optimal solutions on average 40 times faster.

2. Properties of MIP formulations

Padberg and Rijal [21] define a *locally ideal* MIP formulation as one where the vertices of its corresponding LP relaxation satisfy all required integrality conditions. Extending this definition, Croxton et al. [7] and Keha et al. [14] define a locally ideal SOS2 formulation as one whose LP relaxation has extreme points that all satisfy the SOS2 property. As shown by Vielma et al. [23], all commonly used MIP formulations of PLFs, except for the original convex combination (CC) model, are known to be locally ideal. In this section, we demonstrate the theoretical strength of the proposed formulations for X that include the logical condition (2).

2.1. SOS2 model

We consider the following two SOS2-based formulations for X:

$$S_1 := \Big\{(x, y, \boldsymbol{\lambda}, z) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}_+^{n+1} \times \{0, 1\} : x = \sum_{i=0}^{n} B_i\lambda_i,\ y = \sum_{i=0}^{n} F_i\lambda_i,\ 1 = \sum_{i=0}^{n} \lambda_i,\ x \le B_nz,\ \boldsymbol{\lambda} \text{ is SOS2}\Big\}$$

$$S_2 := \Big\{(x, y, \boldsymbol{\lambda}, z) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}_+^{n+1} \times \{0, 1\} : x = \sum_{i=0}^{n} B_i\lambda_i,\ y = \sum_{i=0}^{n} F_i\lambda_i,\ z = \sum_{i=0}^{n} \lambda_i,\ \boldsymbol{\lambda} \text{ is SOS2}\Big\}$$

where S_1 is a standard SOS2 model for PLFs that uses the constraint (4), while formulation S_2 uses the constraint (6) to model the logical condition (2). One can easily show that both S_1 and S_2 are valid formulations of X. In other words, for either $T = S_1$ or $T = S_2$,

$$X = \Big\{(x, y, z) : \exists \boldsymbol{\lambda} \in \mathbb{R}^{n+1} \text{ s.t } (x, y, z, \boldsymbol{\lambda}) \in T\Big\}.$$

We use the standard definition of the *linear programming* (LP) relaxation of a model as the relaxation obtained by replacing integrality restrictions on variables with simple bound restrictions and by removing adjacency requirements for SOS2 variables. We now prove that the formulation S_2 is locally ideal while S_1 is not.

Theorem 1. *Formulation S_2 is locally ideal.*

Proof. The LP relaxation of S_2 has $n + 4$ variables, three equality constraints

$$x = \sum_{i=0}^{n} B_i\lambda_i, \qquad y = \sum_{i=0}^{n} F_i\lambda_i, \qquad z = \sum_{i=0}^{n} \lambda_i,$$

and $n+2$ inequality constraints, $z \le 1$ and $\lambda_i \ge 0\ \forall i = 0, 1, \ldots, n$. Extreme points of the LP relaxation of S_2 have $n + 4$ binding constraints, which forces at least n variables from $\boldsymbol{\lambda} \in \mathbb{R}_+^{n+1}$ to be exactly equal to zero. Thus, the extreme points of the LP relaxation of S_2 are

$$\{(x = B_i,\ y = F_i,\ \boldsymbol{\lambda} = B_i\vec{e}_i,\ z = 1)\ \forall i \in \{1, \ldots, n\}\} \bigcup (x = 0, y = 0, \boldsymbol{\lambda} = \vec{0}, z = 0) \tag{7}$$

where $\vec{e}_i$ are the n dimensional unit vectors. All points in (7) have $z \in \{0, 1\}$ and satisfy the SOS2 properties for the $\boldsymbol{\lambda}$ variables. Hence, S_2 is locally ideal. □

A point $(x, y, \boldsymbol{\lambda}, z)$ can be an extreme point of the set

$$P_2^{\ge} := \Big\{(x, y, \boldsymbol{\lambda}, z) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R}_+^{n+1} \times [0, 1] : x = \sum_{i=0}^{n} B_i\lambda_i,\ y \ge \sum_{i=0}^{n} F_i\lambda_i,\ z = \sum_{i=0}^{n} \lambda_i\Big\}$$

only if $y = \sum_{i=0}^{n} F_i\lambda_i$. Therefore, the proof of Theorem 1 also establishes that expressing the logical condition (2) using (6) also