A Comparison of Mixed-Integer Programming Models for Nonconvex Piecewise Linear Cost Minimization Problems

Keely L. Croxton • Bernard Gendron • Thomas L. Magnanti
Fisher College of Business, The Ohio State University, Columbus, Ohio 43210
Département d'informatique, et de recherche opérationnelle, and Centre de recherche sur les transports, Université de Montréal, Montréal, Quebec H3C 3J7, Canada
School of Engineering, and Sloan School of Management, Massachusetts Institute of Technology, Cambridge, Massachusetts
croxton@cob.ohio-state.edu • gendron@iro.umontreal.ca • magnanti@mit.edu

We study a generic minimization problem with separable nonconvex piecewise linear costs, showing that the linear programming (LP) relaxation of three textbook mixed-integer programming formulations each approximates the cost function by its lower convex envelope. We also show a relationship between this result and classical Lagrangian duality theory.

(Piecewise Linear; Integer Programming; Linear Relaxation; Lagrangian Relaxation)

1. Introduction
Optimization problems with piecewise linear costs arise in many application domains, including transportation, telecommunications, and production planning. Specific applications include variants of the minimum cost network flow problem with nonconvex piecewise linear costs (Aghezzaf and Wolsey 1994, Balakrishnan and Graves 1989, Chan et al. 1997, Cominetti and Ortega 1997, Croxton 1999, Croxton et al. 2002b), such as the network loading problem (Bienstock and Günlük 1996, Gabrel et al. 1999, Günlük 1999, Magnanti et al. 1995), the facility location problem with staircase costs (Holmberg 1994, Holmberg and Ling 1997), and the merge-in-transit problem (Croxton et al. 2002a). Each of these studies introduces integer variables to model the costs, though the choice of the basic formulation varies and includes three textbook models—the so-called incremental, multiple choice, and convex combination models. The objective of this note is to show that the linear programming (LP) relaxations of these mixed-integer programming (MIP) models are equivalent and that they all approximate the cost function by its lower convex envelope. To the best of our knowledge, although this result might appear to be intuitive, no one has formally established it. We also discuss the relationship between this result and classical Lagrangian duality theory.

The general problem, \( P \), is to minimize the separable sum of piecewise linear functions, subject to linear constraints, which we write as \( \min \{ g(x) : Ax \geq b, 0 \leq x \leq u \} \), with \( g(x) = \sum g_j(x_j) \), \( b \) and \( u \) as vectors, and \( A \) as a matrix. Because the formulations we consider model each function \( g_j(x_j) \) separately, for notational simplicity we will drop the subscript \( j \). We then
let $x$ denote a single variable, called the load, and focus on a single piecewise linear function $g(x)$. This simplification is justified by the fact that the lower convex envelope of a separable sum of functions (defined over a bounded polyhedron) equals the sum of the lower convex envelopes of these functions (Falk 1969).

The function $g(x)$ need not be continuous; it can have positive or negative jumps, though we do assume that the function is lower semicontinuous, that is, $g(x) \leq \liminf_{x \to a} g(x')$. Without loss of generality, we also assume, through a simple translation of the costs if necessary, that $g(0) = 0$. Each piecewise linear segment $s \in \{1, 2, \ldots, S\}$ of the function $g(x)$ has a variable cost, $c^s$ (the slope), a fixed cost, $f^s$ (the cost-intercept), and upper and lower bounds, $b^{s-1}$ and $b^s$ (the breakpoints), on the load corresponding to that segment. We assume $b^0 = 0$. Figure 1 illustrates the notation.

Using this notation, in §2 we present three well-known valid MIP models for the problem. In §3, we show that the LP relaxations of the three formulations are equivalent and that they each approximate the cost function by its lower convex envelope. In §4, we discuss the relationship between this result and classical Lagrangian duality theory.

2. Three Models for Piecewise Linear Costs

Incremental Model

Dantzig (1960) and Vajda (1964) both attribute the incremental model to a paper by Manne and Markowitz (1957). As reported in early textbooks, including those by Dantzig (1963) and Hadley (1964), the incremental model introduces a segment load variable, $z^s$, for each segment, defined as the load on the segment $s$, giving a total load $x = \sum_s z^s$. Feasibility requires that the value on segment $s + 1$ be zero unless segment $s$ is "full," that is, $z^{s+1} > 0$ only if $z^s = b^s - b^{s-1}$. To account for this requirement, the incremental model introduces binary variables, $y^s$, defined by the condition that $y^s = 1$ if $z^s > 0$, and $y^s = 0$ otherwise. Defining $\hat{f}^s = (f^s + c^s b^{s-1}) - (f^{s-1} + c^{s-1} b^{s-1})$ as the gap in the cost at the breakpoint between segment $s - 1$ and segment $s$, we can express problem $P$ as a MIP formulation by writing the objective function as $g(x) = \sum_s c^s z^s + f^s y^s$, with the additional constraints:

$$x = \sum_s z^s,$$

$$y^s = 0.$$
Convex Combination Model
The third formulation we examine is a modification of a formulation described in textbooks by Vajda (1964) and Dantzig (1963), and that appears as early as 1960 (Dantzig 1960). The original formulation was intended for continuous cost functions, so we modify it to handle arbitrary (lower semicontinuous) discontinuous functions. This formulation makes use of the fact that the cost of a load that lies in segment $s$ is a convex combination of the cost of the two endpoints, $b^s_{-1}$ and $b^s$, of segment $s$. By defining multipliers $\mu^s$ and $\lambda^s$ as the weights on these two endpoints, we can write the objective function as $g(x) = \sum_s \mu^s (c^s b^s_{-1} + f^s) + \lambda^s (c^s b^s + f^s)$. The $y$ variables having the same interpretation as in the multiple choice model, the constraints are

$$x = \sum_s (\mu^s b^s_{-1} + \lambda^s b^s),$$

$$\mu^s + \lambda^s = y^s,$$

$$\sum_s y^s \leq 1,$$

$$\mu^s, \lambda^s \geq 0, \quad y^s \in \{0, 1\}. \quad (11)$$

3. Comparing the Three Models
Given that all three of the previous models are valid and that researchers have used each of them in different application contexts, it is natural to ask if one is better than another. An important measure for assessing the quality of any MIP formulation is the strength of its LP relaxation. The following result demonstrates the equivalence of the LP relaxations of these three formulations.

**Proposition 1.** The LP relaxations of the incremental, multiple choice, and convex combination formulations are equivalent in the sense that any feasible solution of one LP relaxation corresponds to a feasible solution to the others with the same cost.

**Proof.** See the appendix.

We can further characterize the LP relaxation of these formulations with the following result.

**Proposition 2.** The LP relaxations of the incremental, multiple choice, and convex combination formulations each approximate the cost function, $g(x)$, with its lower convex envelope.

**Proof.** Because by Proposition 1 the three LP relaxations are equivalent, we need only show that the LP relaxation of one of the formulations approximates the cost function with its lower convex envelope. We will use the convex combination formulation, showing that for any load $\hat{x}$, the objective value of the LP relaxation obtained by optimally choosing the other variables is given by the lower convex envelope of the cost function.

By relaxing the integrality restriction on the $y$ variables, we can combine Constraints (9) and (10) into $\sum_s (\mu^s + \lambda^s) \leq 1$ and we can eliminate the $y$ variables. Therefore, a feasible solution is provided by any representation of $\hat{x}$ as a convex combination of the two $S$ points, $(b^s, c^s b^s + f^s)$, with weight $\lambda^s$, and $(b^s_{-1}, c^s b^s_{-1} + f^s)$, with weight $\mu^s$. As we vary the value of $\hat{x}$, the cost-minimizing convex combination is given by the lower convex envelope of these two $S$ points. Because $g(x)$ is piecewise linear, the lower convex envelope of these two $S$ points is the same as the lower convex envelope of $g(x)$. \qed

Another approach for establishing the convex envelope property would be to characterize the structure, especially for the extreme points, of the underlying LP feasible regions for the three models we have considered. For a development of these results, see Croxton et al. (2002c).

4. Relationship to Lagrangian Duality
By associating a vector, $\gamma \geq 0$, of Lagrangian multipliers with the constraints $Ax \geq b$ and letting $g^\gamma(x) = g(x) - \gamma Ax$, we can write the corresponding Lagrangian subproblem, $LS(\gamma)$, as follows: $Z_{LS(\gamma)} = \min_{x \geq 0} [g^\gamma(x) : 0 \leq x \leq u]$. The resulting Lagrangian dual problem, $LD$, is: $Z_{LD} = \max_{\gamma \geq 0} \gamma b + Z_{LS(\gamma)}$. To establish a relationship between Proposition 2 and classical Lagrangian duality theory, we will use the following theorem, due to Falk (1969).

**Theorem 3.** Let $\gamma^*$ be an optimal solution to $LD$ and $x^*$ an optimal solution to the corresponding Lagrangian
subproblem, \( LS(y^*) \). Then, \( x^* \) minimizes the lower convex envelope of the cost function defined over the bounded polyhedron \( \{ x : Ax \geq b, 0 \leq x \leq u \} \).

To establish the desired relationship, we will show that \( Z_{LD} \) equals the optimal value of the LP relaxation of any of the three formulations, say the multiple choice model (a similar development applies to the other two formulations).

Like \( g(x) \), \( g^*(x) \) is a separable sum of piecewise linear functions, which we write as \( g^*(x) = \sum g_j^*(x_j) \). Consequently, we can formulate the Lagrangian subproblem, \( LS(y) \), using the multiple choice model. Given the constraints of this model, we can assume that the bounding constraints \( 0 \leq x \leq u \) are redundant. The resulting problem decomposes into subproblems of the form: \( \min g_j^*(x_j) \), subject to the constraints of the multiple choice model. If, for notational simplicity, we drop the subscript \( j \), each of these subproblems is defined by the objective function \( \sum c_s(y) z_s + f_s y_s \) and the Constraints (4)-(7). In this expression, \( c_s(y) \) is the slope of the segment \( s \) modified by the introduction of the Lagrangian multipliers. Note that the total load variable, \( x \), does not appear in the objective function. We can derive its value from the values of the segment load variables, \( z^* \). Thus, we can remove Constraint (4).

We could derive the same Lagrangian subproblem as follows: First, reformulate problem \( P \) using the multiple choice model; then, in the resulting MIP formulation, relax constraints \( Ax \geq b \) in a Lagrangian fashion. Clearly, the resulting Lagrangian dual is equivalent to \( LD \), because the Lagrangian subproblems are identical. Classical Lagrangian duality theory in MIP (Geoffrion 1974) implies that the Lagrangian dual and the LP relaxation of the MIP model have equal optimal values if, for any cost function \( \sum c_s(y) z^*_s + f_s y^*_s \), the LP relaxation of the corresponding Lagrangian subproblem has an integral optimal solution. Thus, \( Z_{LD} \) equals the optimal value of the LP relaxation of the multiple choice model if we can show that when we minimize some cost function \( \sum c_s(y) z^*_s + f_s y^*_s \) over the polyhedron \( Q = \{(y, z) : b^{-1} y^* \leq z^*_s \leq b y^*_s, \sum y^*_s \leq 1, y^*_i \geq 0 \} \), the problem has an optimal solution with each \( y^*_i \in [0, 1] \).

This property is easy to establish. Suppose we minimize some cost function \( \sum c_s(y) z^*_s + f_s y^*_s \) over the polyhedron \( Q \). If \( c_s(y) \geq 0 \), then \( z^*_s = b^{-1} y^*_s \) in some optimal solution, while if \( c_s(y) \leq 0 \), then \( z^*_s = b y^*_s \) in some optimal solution. Therefore, we can express each \( z^*_s \) in terms of the \( y^*_s \) variables, and eliminate the \( z^*_s \) variables and the constraints \( b^{-1} y^*_s \leq z^*_s \leq b y^*_s \). The resulting problem has a linear objective function and the single constraint \( \sum y^*_s \leq 1 \) in the nonnegative \( y^*_s \) variables. Because the problem has a single constraint, it has an optimal solution with at most one \( y^*_i = 1 \) and all other \( y \) variables at value zero. Therefore, for some optimal solution the value of each \( y^*_i \) is 0 or 1.

This discussion shows how Lagrangian duality results imply the convex envelope property of the three classical models for optimization problems with nonconvex piecewise linear costs. Conversely, it shows that the convex envelope property of the classical models presages the Lagrangian duality result and further demonstrates the strong relationship between Lagrangian duality and linear programming.

5. Conclusion

We have shown that the LP relaxations of three textbook MIP models for nonconvex piecewise linear minimization problems defined over bounded polyhedra are equivalent, each approximating the cost function with its lower convex envelope. We have also discussed the relationship between these results and classical Lagrangian duality theory.

The equivalence between the three LP relaxations and the fact that they all approximate the lower convex envelope of the cost function has several implications. First, it shows that from the perspective of LP relaxations, choosing among the three models is irrelevant. We might prefer one model to another for other reasons (for example, their use within specific algorithms), but they all provide the same LP relaxations and bounds.

As an algorithmic implication, suppose we use a branch-and-bound algorithm to solve a nonconvex piecewise linear cost minimization problem with a feasible region defined by a bounded polyhedron. There are two obvious relaxations for computing the lower bounds at the nodes of the enumeration
tree: either the lower convex envelope or the LP relaxation of a MIP formulation of the problem. Falk and Soland (1969) studied the first approach, but to the best of our knowledge, no one has ever recognized the fundamental relationship between their method and an LP-based branch-and-bound method: both compute the same lower bounds.

**Appendix. Proof of Proposition 1**

We establish this result by providing translations between feasible solutions of (1) the multiple choice and convex combination formulations and (2) the incremental and multiple choice formulations. We show that these translations give feasible solutions with the same cost. We denote the LP relaxations of the incremental, multiple choice, and convex combination formulations as LP(I), LP(M), and LP(C), respectively.

**Multiple Choice → Convex Combination**

Consider a feasible solution \((x, y, z)\) to LP(M). Because \(b^{-1}y' \leq z' \leq b'y'\), for some value of \(0 \leq \alpha \leq 1\), \(z' = \alpha b^{-1}y' + (1 - \alpha)b'y'\). Let \(\mu' = \alpha y'\) and \(\lambda' = (1 - \alpha)y'\). Then \(\mu' + \lambda' = y'\) and \(z' = \mu b^{-1} + \mu' b\).

Because \(x = \sum z', \ x = \sum (\mu b^{-1} + \mu' b)\). Therefore, \((x, y, \mu, \lambda)\) is feasible for LP(C). The cost of this solution is \(\sum (\mu c b^{-1} + \mu' f) + \lambda' (c b^{-1} + f') = \sum (\mu b^{-1} + \lambda' b) (\mu + \lambda) = \sum c z' + f' y'\), which equals the cost of \((x, y, z)\), the solution to LP(M).

**Convex Combination → Multiple Choice**

Consider a feasible solution \((x, y, \mu, \lambda)\) to LP(C). Define \(z' = \mu b^{-1} + \lambda' b\). The conditions \(b^{-1}y' \leq z' \leq b'y'\) imply that \(b^{-1}y' \leq z' \leq b'y'\). Therefore, \((x, y, z)\) is feasible for LP(M). As shown previously, the cost of this solution is the same as the cost of \((x, y, z)\).

**Incremental → Multiple Choice**

Consider a feasible solution \((x, y, z)\) to LP(I). Let \(w' = z' + b^{-1}y' - b' y'\) and \(v' = y' - y'\). If we add \(b^{-1}y' - b' y'\) to each of the terms in (2), these inequalities become \(b^{-1}y' \leq z' \leq b'y'\). The inequalities (2) imply that \(y' \leq y'\), and thus \(v' \geq 0\). In addition, \(\sum z' = \sum z', \ b'y' - b'y' = \sum z'\). Therefore, \((x, v, w)\) is a feasible solution to LP(M). The cost of this solution is \(\sum (c w' + f' v') = \sum (c(\mu + \lambda)' y' + (f' + c b^{-1}) y' - \sum (f' + c b^{-1}) y') = \sum c z' + \sum (f' + c b^{-1}) y' = \sum c z' + f' y'\), which equals the cost of \((x, y, z)\).

**Multiple Choice → Incremental**

Consider a feasible solution \((x, y, z)\) to LP(M). Let \(w' = z' + (b^{-1} - b^{-1})(\sum y') - b'^{-1}y' + (b' - b'^{-1}) y'\). These definitions imply that \(z' = w' + b^{-1}y' - b'^{-1}y'\) and \(y' = v' - z'\). Also note that \(0 \leq v' \leq 1\). Through substitution, the inequalities (5) imply \(b^{-1}y' \leq z' \leq b'y'\). In addition, \(\sum z' = x\). Therefore, \((x, v, w)\) is a feasible solution to LP(I). Moreover, using the same equations as in the translation from a solution of LP(I) to a solution of LP(M), it is easy to show that the cost of \((x, v, w)\) is the same as the cost of \((x, y, z)\), the solution to LP(M).

**References**


Accepted by Thomas Liebling, former department editor; received July 8, 2002. This paper was with the authors 9 months for 2 revisions.