ARTICLE IN PRESS

Computers & Operations Research ∎ (∎∎∎∎) ∎∎∎–∎∎∎

Contents lists available at ScienceDirect

Computers & Operations Research

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

An exact algorithm for the bilevel mixed integer linear programming problem under three simplifying assumptions

Pan Xu [a], Lizhi Wang [b,*]

[a] University of Maryland, United States
[b] Iowa State University, United States

ARTICLE INFO

Keywords:
Bilevel optimization
Bilevel mixed integer linear programming
Branch-and-bound

ABSTRACT

We present an exact algorithm for the bilevel mixed integer linear programming (BMILP) problem under three simplifying assumptions. Although BMILP has been studied for decades and widely applied to various real world problems, there are only a few BMILP algorithms. Compared to these existing ones, our new algorithm relies on fewer and weaker assumptions, explicitly considers finite optimal, infeasible, and unbounded cases, and is proved to terminate finitely and correctly. We report results of our computational experiments on a small library of BMILP test instances, which we created and made publicly available online.

1. Introduction

We present an exact algorithm for the bilevel mixed integer linear programming (BMILP) problem under three simplifying assumptions:

$$\max_{x,y} \quad \zeta = c^\top x + d_1^\top y \tag{1}$$

$$\text{s.t.} \quad A_1 x + B_1 y \le b_1 \tag{2}$$

$$0 \le x \le X \tag{3}$$

$$x_i \in \mathbb{Z}, \ \forall i \in I \tag{4}$$

$$y \in \underset{\tilde{y}}{\operatorname{argmax}} \{ d_2^\top \tilde{y} : A_2 x + B_2 \tilde{y} \le b_2; \tilde{y} \ge 0; \tilde{y}_j \in \mathbb{Z}, \forall j \in J \}. \tag{5}$$

where $A_1 \in \mathbb{R}^{m_1 \times n_1}$, $A_2 \in \mathbb{Z}^{m_2 \times n_1}$, $B_1 \in \mathbb{R}^{m_1 \times n_2}$, $B_2 \in \mathbb{R}^{m_2 \times n_2}$, $b_1 \in \mathbb{R}^{m_1 \times 1}$, $b_2 \in \mathbb{R}^{m_2 \times 1}$, $c \in \mathbb{R}^{n_1 \times 1}$, $d_1 \in \mathbb{R}^{n_2 \times 1}$, $d_2 \in \mathbb{R}^{n_2 \times 1}$, $I = \{1, \ldots, n_1\}$, $J \subseteq \{1, \ldots, n_2\}$, and $X \in \mathbb{R}^{n_1 \times 1}$. Compared to the general BMILP formulation, our definition of the BMILP problem contains three simplifying assumptions.

Assumption 1. All variables in vector x are required to be integral: $I = \{1, \ldots, n_1\}$.

Assumption 2. All variables in vector x have known bounds: $0 \le x \le X$.

Assumption 3. Matrix A_2 is integral: $A_2 \in \mathbb{Z}^{m_2 \times n_1}$.

The BMILP model we try to solve belongs to the category of bilevel optimization, which has been studied for decades. Pioneers of bilevel optimization models include Bracken and McGill [11–13], Aiyoshi [1,2], Bard [5], and Candler [14], among others. Most early studies [8–10,27,39] focused on the simpler case of bilevel linear program. Since the 1990s, there has been increased attention on more complex models with nonlinear terms [6,4,21,22] or discrete decision variables [7,19,32,38]. Comprehensive reviews of existing bilevel optimization algorithms and applications can be found in [15,37].

Bilevel optimization models have been applied to solve a variety of real world problems, in which the hierarchical structure of decision making widely exists. These applications include revenue management [17], network design [16], national security [36], network interdiction [30,33,41], national agriculture planning [25], and decentralized management of multidivisional firms [3].

Despite its broad applications, BMILP is intrinsically hard to solve, both theoretically and computationally. This can be epitomized by the following two-dimensional example from [28].

Example 1.

$$\sup_{x,y} \quad \zeta = -x + y$$
$$\text{s.t.} \quad 0 \le x \le 1$$
$$y \in \underset{\tilde{y}}{\operatorname{argmax}} \{ -\tilde{y} : 0 \le \tilde{y} \le 1; \tilde{y} \ge x; \tilde{y} \in \mathbb{Z} \}.$$

It can be easily checked that: (i) the bilevel feasible region is nonconvex and disconnected, consisting of the point $(0,0)$ and the line segment between $(0,1)$ and $(1,1)$, including $(1,1)$ but excluding $(0,1)$; (ii) the supremum of ζ is 1, but it is not attainable; and

* Corresponding author. Tel.: +1 5152941757.
E-mail address: lzwang@iastate.edu (L. Wang).

0305-0548/$ - see front matter © 2013 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.cor.2013.07.016

(iii) the optimal objective value of the continuous relaxation, which is 0, fails to provide an upper bound for ζ. These special properties are unique to the hierarchical structure of the bilevel problems, which place significant challenges to designing exact and efficient BMILP algorithms.

The literature on BMILP algorithms is scarce. In their seminal work [32], Moore and Bard presented the first branch-and-bound algorithms for BMILP under several simplifying assumptions. Under the condition that either $I=\{1,\dots,n_1\}$ or $J=\varnothing$, they proved that the algorithms can find an optimal solution if one exists. An example is provided to show that if the upper level contains continuous variables and the lower level discrete ones, then the supremum of the BMILP may not be attainable even if it finitely exists. The authors published another branch-and-bound algorithm in [7], which solves a more restricted model with $m_1=0$, $x\in\{0,1\}^{n_1}$, and $y\in\{0,1\}^{n_2}$. They pointed out that "it cannot be determined whether the algorithm terminates with the optimum." DeNegre and Ralphs [19] invented a BMILP algorithm under the assumptions that $I=\{1,\dots,n_1\}$, $J=\{1,\dots,n_2\}$, and $B_1=0^{m_1\times n_2}$. A Benders decomposition method is proposed in [35]. Genetic algorithms and Tabu search methods for BMILP are presented in [34,40], respectively. Relevant literature also includes [18,28]. In [28], the authors proved the existence of an algorithm that solves BMILP in polynomial time when n_2 is fixed, referring to parametric integer programming approaches [23]. Several bilevel discrete nonlinear programming algorithms [20,24,26,29,31] have also been proposed, which rely on additional simplifying assumptions to achieve convergence or ϵ-convergence.

The algorithm we present here, which will be referred to as $\text{Alg}^{\text{BMILP}}$, differs from the previous ones in the following four aspects. First, we allow the B_1 matrix to be nonzero. As a result, obtaining a bilevel feasible solution is no longer straightforward. This point will be further explained in Section 2.1. Nevertheless, allowing for nonzero B_1 could be practically imperative, because in many real world situations (such as in the energy market) the consequences of the lower level's decisions (such as emissions resulting from energy generation) must be taken into account by the upper level in explicit constraints (such as environmental policy design goals). Second, we allow the lower level to have both continuous and discrete variables. Third, we explicitly consider all possible outcomes of a BMILP, be it infeasible, unbounded, or finite optimal. Fourth, we prove that our algorithm will finitely terminate with the correct output. The three simplifying assumptions are necessary for our algorithm. The first one is to avoid the case of unattainable supremum, the second one is to ensure finite termination, and the third one is to eliminate unavoidable rounding errors from computer representation of irrational numbers.

The rest of the paper is organized as follows. We motivate and present the algorithm in Section 2 and report results from computational experiments in Section 3. Discussion and concluding remarks are made in Section 4.

2. An exact BMILP algorithm

2.1. Definitions and preliminaries

In this section, we introduce some definitions that will be used in the algorithm. We use a vector with a subscript j to refer to the jth element of the vector. We define $\overline{\mathbb{R}}=\mathbb{R}\cup\{+\infty\}\cup\{-\infty\}$ as the set that includes all real numbers as well as positive and negative infinity. For a given parameter x, we define $\mathcal{L}(x)$ as the following parametric MILP (6)–(8), referred to as the lower level problem (LLP):

$$\max_{y} \quad d_2^\top y \tag{6}$$

$$\text{s.t.} \quad A_2x+B_2y\le b_2 \tag{7}$$

$$y\ge 0, \quad y_j\in\mathbb{Z}, \quad \forall j\in J. \tag{8}$$

For a given set of parameters ($l\in\overline{\mathbb{R}}^{m_2\times 1}, u\in\overline{\mathbb{R}}^{m_2\times 1}, w\in\overline{\mathbb{R}}$), we define $\mathcal{B}(l,u,w)$ as the following parametric BMILP problem:

$$\max_{x,y} \quad \zeta=c^\top x+d_1^\top y \tag{9}$$

$$\text{s.t.} \quad A_1x+B_1y\le b_1 \tag{10}$$

$$l\le A_2x\le u \tag{11}$$

$$d_2^\top y\ge w \tag{12}$$

$$0\le x\le X \tag{13}$$

$$x\in\mathbb{Z}^{n_1\times 1} \tag{14}$$

$$y\in\underset{\tilde{y}}{\operatorname{argmax}}\{d_2^\top\tilde{y} : A_2x+B_2\tilde{y}\le b_2; \tilde{y}\ge 0; \tilde{y}_j\in\mathbb{Z}, \forall j\in J\}. \tag{15}$$

For a given set of parameters ($l\in\overline{\mathbb{R}}^{m_2\times 1}, u\in\overline{\mathbb{R}}^{m_2\times 1}, w\in\overline{\mathbb{R}}$), we define $\mathcal{H}(l,u,w)$ as the following parametric MILP (16)–(24), referred to as the high point problem (HPP):

$$\max_{x,y} \quad c^\top x+d_1^\top y \tag{16}$$

$$\text{s.t.} \quad A_1x+B_1y\le b_1 \tag{17}$$

$$A_2x+B_2y\le b_2 \tag{18}$$

$$l\le A_2x\le u \tag{19}$$

$$d_2^\top y\ge w \tag{20}$$

$$0\le x\le X \tag{21}$$

$$x\in\mathbb{Z}^{n_1\times 1} \tag{22}$$

$$y\ge 0 \tag{23}$$

$$y_j\in\mathbb{Z}, \quad \forall j\in J. \tag{24}$$

The term "high point problem" was first used in [10] for bilevel linear program and then in [32] for BMILP, although our definition (16)–(24) is different from theirs.

A solution (x,y) is called LLP optimal if y is an optimal solution to $\mathcal{L}(x)$. For a given HPP $\mathcal{H}(l,u,w)$, a solution (x,y) is called HPP feasible/optimal if (x,y) is a feasible/optimal solution to $\mathcal{H}(l,u,w)$. For a given BMILP $\mathcal{B}(l,u,w)$, a solution (x,y) is called bilevel feasible if it satisfies Constraints (10)–(15). A solution is called bilevel infeasible if it is not bilevel feasible. One can verify from the above definitions that a solution is bilevel feasible if and only if it is both HPP feasible and LLP optimal. A solution (x^*,y^*) is called bilevel optimal if it is bilevel feasible and we have $c^\top x^*+d_1^\top y^*\ge c^\top x^0+d_1^\top y^0$ for any other bilevel feasible solution (x^0,y^0). A BMILP instance is called optimal if it possesses (uniquely or not) a finite optimal solution. A BMILP instance is called infeasible if a bilevel feasible solution does not exist. A BMILP instance is called unbounded if for any real number K there exists a bilevel feasible solution (x^K,y^K) such that $c^\top x^K+d_1^\top y^K\ge K$.

We clarify that in (5), the symbol "argmax{}" refers to the set of optimal solutions of the problem enclosed in the braces, which is the LLP $\mathcal{L}(x)$. If the LLP is infeasible or unbounded, then this set is empty; otherwise this set contains all LLP optimal solutions. The formulation (1)–(5) reads that when the lower level possesses multiple optimal solutions the upper level gets to pick. This is commonly referred to as the optimistic formulation. In contrast, there is also the pessimistic formulation of BMILP, which allows