J Comb Optim (2011) 22:594–608
DOI 10.1007/s10878-010-9307-9

Optimality conditions for a bilevel matroid problem

Diana Fanghänel

Published online: 4 March 2010

Abstract In bilevel programming there are two decision makers, the leader and the follower, who act in a hierarchy. In this paper we deal with a weighted matroid problem where each of the decision makers has a different set of weights. The independent set of the matroid that is chosen by the follower determines the payoff to both the leader and the follower according to their different weights. The leader can increase his payoff by changing the weights of the follower, thus influencing the follower's decision, but he has to pay a penalty for this. We want to find an optimum strategy for the leader. This is a bilevel programming problem with continuous variables in the upper level and a parametric weighted matroid problem in the lower level. We analyze the structure of the lower level problem. We use this structure to develop local optimality criteria for the bilevel problem that can be verified in polynomial time.

Keywords Bilevel programming · Combinatorial optimization · Matroids

1 Introduction

Bilevel programming problems are used to model decision processes with two decision makers that are called leader and follower, respectively. At this, leader and follower solve linked optimization problems. The leader modifies with his decision the problem of the follower. In return, the value of this decision is also influenced by the decision of the follower. The task in bilevel programming is to find a good strategy for the leader to optimize his objective function value.

Due to the many applications there exists a large number of papers on bilevel programming. An overview on this is given in the books and bibliographies of Bard (1998) and Dempe (2002, 2003).

D. Fanghänel (✉)
FB Elektrotechnik/Informatik, Universität Kassel, Wilhelmshöher Allee 73, 34121 Kassel, Germany
e-mail: d.fanghaenel@uni-kassel.de

However, there are only a small number of papers (Dempe 2002; Dempe et al. 2005; Fanghänel 2006a, 2006b; Fanghänel and Dempe 2009; Kalashnikov and Ríos-Mercado 2001; Vicente et al. 1996) that investigate problems where the decision variables of the follower are discrete and those of the leader are continuous. More established is the related theory of inverse combinatorial optimization. An overview on inverse combinatorial problems can be found in Heuberger (2004).

To make progress in the theory of bilevel programming problems with discrete lower and continuous upper level, it seems reasonable to start with the easiest possible cases, such as problems that involve in polynomial time solvable combinatorial problems.

The easiest problems in combinatorial optimization are the weighted matroid problems. They can be solved by using the well-known Greedy Algorithm and are well investigated. The theory of these problems is common knowledge and can be found, e.g., in Cook et al. (1998), Schrijver (2003). To the best of our knowledge there are no papers that investigate bilevel problems with a parametric weighted matroid problem in the lower level and a continuous problem in the upper level.

Thus, in this paper we consider a problem where the second problem is a parametric weighted matroid problem. We start the investigations with the mathematical formulation of the problem and some important definitions. In Sects. 3 and 4 we explore the structure of the solution sets and the so-called regions of stability.

In 1994 it was proved by Vicente et al. (1994), that it is NP-hard to check local optimality for a given feasible point of a linear bilevel problem. In Sect. 5 we show that checking local optimality is less difficult for the bilevel matroid problem. Actually, we provide local optimality conditions that can be tested in polynomial time $\mathcal{O}(\gamma n^2)$.

2 Mathematical formulation of the problem

In combinatorial optimization matroids are an important concept that has been developed during the last century. It was introduced to find a common language for similar problems in, e.g., linear algebra, graph theory, and projective geometry. Many notations in matroid theory have their origin in these areas.

Matroids can be introduced by different equivalent definitions. We will use the independent sets to define them.

Definition 1 Let E with $|E| = n > 1$ be a finite set. Further, let $\mathcal{I} \subseteq 2^E$ be a subset of the partition set of E. We denote the pair $(E, \mathcal{I})$ as a matroid iff the following conditions are satisfied:

(I0) $\emptyset \in \mathcal{I}$,
(I1) $I_1 \subseteq I_2$ with $I_2 \in \mathcal{I}$ implies $I_1 \in \mathcal{I}$,
(I2) for all sets $I_1, I_2 \in \mathcal{I}$ with $|I_1| < |I_2|$ there exists some element $e \in I_2 \backslash I_1$ with $I_1 \cup \{e\} \in \mathcal{I}$.

The elements $I \in \mathcal{I}$ are called independent sets. All other subsets $A \subseteq E$ with $A \notin \mathcal{I}$ are called dependent. A base is an independent set that is not a subset of a larger independent set. All bases of the matroid have the same cardinality that is called rank of the matroid. The set of all bases we denote with $\mathcal{B}$.

Typical examples for matroids are, e.g., the partition matroid and the graphical matroid. For the partition matroid the set E is partitioned into disjoint subsets $E = E_1 \cup \cdots \cup E_k$, and the independent sets are defined by

$$\mathcal{I} = \{I \subseteq E : |I \cap E_i| \leq n_i \text{ for all } i = 1, \ldots, k\}$$

with given natural numbers $n_i \in \mathbb{N}$, $i = 1, \ldots, k$. The graphical matroid is defined on a graph $G = (V, E)$ with the independent sets

$$\mathcal{I} = \{I \subseteq E : G(I) = (V, I) \text{ is a forest}\}.$$

A programming problem that is often solved for a matroid $(E, \mathcal{I})$ is the so-called weighted matroid problem

$$\max\{w(I) : I \in \mathcal{I}\} \tag{WMP}$$

where $w(I) = \sum_{e \in I} w_e$ is the sum of all weights w_e with $e \in I$ for a given vector $w \in \mathbb{R}^E$. The weighted matroid problem (WMP) can be solved easily with the Greedy Algorithm.

Assume two players with different weight vectors want to solve weighted matroid problems in the following hierarchical way. The so-called follower solves the weighted matroid problem with respect to his weight vector $w^F \in \mathbb{R}^E_+$ with $w^F_e > 0$ for all elements $e \in E$. His optimum choice $I \in \mathcal{I}$ also determines the result $w^L(I)$ of the second player (so-called leader). We assume that the leader knows the weights of the follower. Therefore he can predict what the follower will do. To obtain a better result, he can influence the choice of the follower by increasing the weights w^F by a vector $x \in \mathbb{R}^E_+$. For this increase the leader has to pay a penalty $c^\top x$ where $c \in \mathbb{R}^E_+$ is a given cost vector with $c_e > 0$ for all elements $e \in E$. We ask for the best decision $x \in \mathbb{R}^E_+$ of the leader.

This is a bilevel programming problem. In formulas we can write it as

$$\begin{aligned} &\max_x \{w^L(I) - c^\top x : I \in \Psi(x),\ x \geq \mathbf{0}\}, \\ &\text{s.t.} \quad \Psi(x) = \operatorname*{Argmax}_I \{(w^F + x)(I) : I \in \mathcal{I}\}. \end{aligned} \tag{LP}$$

The problem (LP) is the problem of the leader. At this, the set $\Psi(x)$ is the set of all optimal solutions of the problem of the follower

$$\max_I \{(w^F + x)(I) : I \in \mathcal{I}\} \tag{FP}$$

depending on the parameter $x \geq \mathbf{0}$. Obviously, it holds $\Psi(x) \neq \emptyset$ for all parameters x. But for some $x \geq \mathbf{0}$ the set $\Psi(x)$ is not a singleton, i.e., $|\Psi(x)| > 1$. This causes some uncertainty in the definition of problem (LP). In such a case the leader can not predict the decision of the follower. In bilevel programming usually two different approaches are proposed to handle such a situation. The optimistic approach is used if there is an agreement, that the follower chooses that of his optimum solutions, that is best for the leader. In this case the leader has to maximize the so-called optimistic solution function

$$\phi_o(x) = \max_{I \in \Psi(x)} w^L(I) - c^\top x$$

with respect to $x \geq \mathbf{0}$, i.e., he has to solve the problem

$$\max_x \{\phi_o(x) : x \geq \mathbf{0}\}. \tag{Opt}$$

The local/global optimal solutions of this problem (Opt) are called local/global optimistic solutions.

If there is no agreement, it is possible that the follower chooses that of his optimum solutions, that is worst for the leader. In this case the leader maximizes the so-called pessimistic solution function

$$\phi_p(x) = \min_{I \in \Psi(x)} w^L(I) - c^\top x,$$

i.e., he solves the problem

$$\max_x \{\phi_p(x) : x \geq \mathbf{0}\}. \tag{Pess}$$

The local/global optimal solutions of this problem (Pess) are called local/global pessimistic solutions. In this paper, we want to determine both local optimistic and local pessimistic solutions.

3 The solution sets

In this section we investigate the properties of the solution sets $\Psi(x)$. At this we use, that the problem of the follower (FP) is a weighted matroid problem with the weight vector $w^F + x$. Therefore, it can be solved with the Greedy Algorithm.

With the Greedy Algorithm (cf. Cook et al. 1998; Schrijver 2003) both maximum independent sets and maximum bases of a matroid can be computed. Let $w : \mathbb{R}^E \to \mathbb{R}$ be the weight vector. Then we can compute a maximum independent set for the problem $\max\{w(I) : I \in \mathcal{I}\}$ with the following algorithm.

Algorithm 1 (Greedy Algorithm 1—maximum independent sets)

1. *Order the element of E as $e_1, \ldots, e_n$ such that $w_{e_1} \geq w_{e_2} \geq \cdots \geq w_{e_n}$.*
2. *Set $I = \emptyset$.*
3. *For $i = 1, \ldots, n$, if $I \cup \{e_i\} \in \mathcal{I}$ and $w_{e_i} \geq 0$, set $I = I \cup \{e_i\}$.*

However, sometimes we are searching for a maximum base, i.e., we want to solve the problem $\max\{w(I) : I \in \mathcal{B}\}$. Then the following version of the Greedy Algorithm can be used.

Algorithm 2 (Greedy Algorithm 2—maximum bases)

1. *Order the element of* E *as* $e_1, \dots, e_n$ *such that* $w_{e_1} \geq w_{e_2} \geq \dots \geq w_{e_n}$.
2. *Set* $I = \emptyset$.
3. *For* $i = 1, \dots, n$, *if* $I \cup \{e_i\} \in \mathcal{I}$, *set* $I = I \cup \{e_i\}$.

Clearly, if the weights w_e are nonnegative for all $e \in E$, the Algorithm 1 coincides with Algorithm 2. Due to the assumptions, this holds also for the weight vector $w = w^F + x$. Thus, we can determine some element of the set $\Psi(x)$ with Algorithm 2.

If some independent set $I \in \mathcal{I}$ is given, we can verify $I \in \Psi(x)$ with the following well-known theorem.

Theorem 1 (Cook et al. 1998) *Let* $\mathcal{M} = (E, \mathcal{I})$ *be a matroid. Let* $I \in \mathcal{I}$ *and* $x \in \mathbb{R}^E_+$. *Then it holds* $I \in \Psi(x)$ *if and only if the following conditions are satisfied*:

1. *The independent set* $I \in \mathcal{I}$ *is a base of the matroid, i.e.,* $I \in \mathcal{B}$.
2. *If* $q \notin I$, $e \in I$, *and* $(I \cup \{q\}) \backslash \{e\} \in \mathcal{I}$, *then* $w^F_q + x_q \leq w^F_e + x_e$ *holds.*

However, we need a description of the whole set $\Psi(x)$. This can be answered by specifying a theorem on polymatroids (Fujishige 2005, Theorem 3.15).

Theorem 2 *Let* $x \in \mathbb{R}^E_+$, *and assume that the distinct values of* $w^F_e + x_e$ $(e \in E)$ *are given by*

$$w_1 > w_2 > \dots > w_p > 0.$$

Then the sets

$$A_i = \{e \in E : w^F_e + x_e \geq w_i\}, \quad i = 1, \dots, p$$

form a chain $A_1 \subset A_2 \subset \dots \subset A_p = E$. *With these sets matroids* $\mathcal{M}_{A_i/A_{i-1}} = (A_i \backslash A_{i-1}, \mathcal{I}_{A_i/A_{i-1}})$ *are defined with* $A_0 = \emptyset$ *and*

$$\mathcal{I}_{A_i/A_{i-1}} = \{I \subseteq A_i \backslash A_{i-1} : \exists I' \subseteq A_{i-1} \text{ with } I \cup I' \in \mathcal{I}\},$$

$i = 1, \dots, p$. *The set of all bases of* $\mathcal{M}_{A_i/A_{i-1}}$ *we denote with* $\mathcal{B}_{A_i/A_{i-1}}$. *Then it is*

$$\Psi(x) = \mathcal{B}_{A_1/A_0} \oplus \mathcal{B}_{A_2/A_1} \oplus \dots \oplus \mathcal{B}_{A_p/A_{p-1}},$$

i.e., $I \in \Psi(x)$ *iff there exist bases* $I_i \in \mathcal{B}_{A_i/A_{i-1}}$ *with* $I = I_1 \cup I_2 \cup \dots \cup I_p$.

Remarks

1. The matroid $\mathcal{M}_{A_i/A_{i-1}}$ is generated by repeatedly applying matroid deletion and contraction.
2. As an example, if $\mathcal{M}$ is the graphical matroid for a graph G, $\mathcal{M}_{A_i/A_{i-1}}$ is the graphical matroid for the graph $G' = G(A_i)/A_{i-1}$, that is generated by contracting the edge set A_{i-1} in the subgraph that is induced by the edge set A_i.
3. The matroids $\mathcal{M}_{A_i/A_{i-1}}$ can have zero rank, i.e., it is possible that the empty set is the only independent set.

4. The directed sum

$$\mathcal{M}(x) = \mathcal{M}_{A_1/A_0} \oplus \mathcal{M}_{A_2/A_1} \oplus \cdots \oplus \mathcal{M}_{A_p/A_{p-1}}$$

of the p small matroids is also a matroid (see Cook et al. 1998). Then, due to the theorem, $\Psi(x)$ is the set of all bases of the matroid $\mathcal{M}(x)$.

Since $\Psi(x)$ is the set of all bases of a matroid, we obtain the following result on the computation of the optimistic/pessimistic solution function.

Corollary 1 *Let $x \geq \mathbf{0}$. Then, the values $\phi_o(x)$ and $\phi_p(x)$ can be computed in polynomial time.*

Proof Remember that the optimistic/pessimistic solution function was defined by

$$\phi_o(x) = -c^\top x + \max_{I \in \Psi(x)} w^L(I) \quad \text{and} \quad \phi_o(x) = -c^\top x + \min_{I \in \Psi(x)} w^L(I),$$

respectively. Since $\Psi(x)$ is the set of bases of the matroid $\mathcal{M}(x)$, the formulas contain matroid problems. These matroid problem can be solved with the Greedy Algorithm (see Algorithm 2). Consequently, the values $\phi_o(x)$ and $\phi_p(x)$ can be computed within polynomial time. □

In practice it is enough to apply the Greedy Algorithm only once. For the optimistic case the algorithm is as follows.

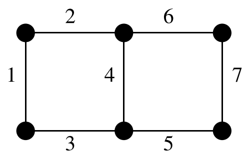
Algorithm 3 (Compute $\phi_o(x)$)

1. *Order the elements of $E = \{e_1, \ldots, e_n\}$ such that it holds*
   (a) $w^F_{e_i} + x_{e_i} \geq w^F_{e_{i+1}} + x_{e_{i+1}}$ *for all indices $i = 1, \ldots, n-1$,*
   (b) $w^L_{e_i} \geq w^L_{e_{i+1}}$ *for each index i with $w^F_{e_i} + x_{e_i} = w^F_{e_{i+1}} + x_{e_{i+1}}$.*
2. *Set $I = \emptyset$.*
3. *For $i = 1, \ldots, n$, if $I \cup \{e_i\} \in \mathcal{I}$, then set $I = I \cup \{e_i\}$.*
4. *Compute $\phi_o(x) = w^L(I) - c^\top x$.*

To compute the value $\phi_p(x)$ we have only to reverse the second ordering condition, i.e., we have to order the elements of E such that $w^L_{e_i} \leq w^L_{e_{i+1}}$ holds for each index i with $w^F_{e_i} + x_{e_i} = w^F_{e_{i+1}} + x_{e_{i+1}}$. Then a base I is computed with Step 2 and Step 3, and $\phi_p(x) = w^L(I) - c^\top x$. For optimality conditions in the optimistic case this computed base will be more important to us than the value.

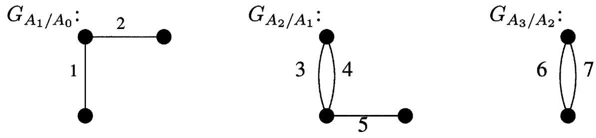
Next we want to apply the statements of this section on a small example.

Example 1 We investigate the graphical matroid with $E = \{1, 2, 3, 4, 5, 6, 7\}$ and the following graph G:



For the data $x = \mathbf{0}$, $w^L = (1, -1, 0, 1, 1, 5, 4)$, and $w^F = (4, 4, 2, 2, 2, 1, 1)$ we want to compute the set $\Psi(x)$, $\phi_o(x)$, and $\phi_p(x)$.

We start with computing the set $\Psi(x)$ as described in Theorem 2. It is $A_1 = \{1, 2\}$, $A_2 = \{1, 2, 3, 4, 5\}$, and $A_3 = E$. Thus, we have to compute all spanning forests of the following graphs:



The sets of spanning forests are $\mathcal{B}_{A_1/A_0} = \{\{1, 2\}\}$, $\mathcal{B}_{A_2/A_1} = \{\{3, 5\}, \{4, 5\}\}$, and $\mathcal{B}_{A_3/A_2} = \{\{6\}, \{7\}\}$. Consequently, we obtain

$$\Psi(\mathbf{0}) = \{\{1, 2, 3, 5, 6\}, \{1, 2, 3, 5, 7\}, \{1, 2, 4, 5, 6\}, \{1, 2, 4, 5, 7\}\}.$$

Now we want to compute the value $\phi_o(x)$. Because of $x = \mathbf{0}$ we do not need the vector c for this. It is simply $\phi_o(x) = \max\{w^L(I) : I \in \Psi(x)\}$. We apply Algorithm 3. Thus, first we have to order the elements of E. The values of w^F are already decreasing. However, the values of w^L are not. Therefore we reorder the elements of E as $E = \{1, 2, 4, 5, 3, 6, 7\}$. Then we proceed with the algorithm. As a solution we compute

$$I_o = \{1, 2, 4, 5, 6\}, \quad \text{i.e.,} \quad \phi_o(\mathbf{0}) = 7.$$

In the pessimistic case we obtain $E = \{2, 1, 3, 4, 5, 7, 6\}$ by reordering the elements. Consequently, a result of the algorithm is

$$I_p = \{2, 1, 3, 5, 7\} \quad \text{and} \quad \phi_p(\mathbf{0}) = 5.$$

4 The regions of stability

An important concept for the development of optimality conditions and solution algorithms for the problems (Opt) and (Pess) are the so-called regions of stability. They describe the set of all points $x \geq \mathbf{0}$, for which a given independent set is optimum, i.e.,

$$R(I) = \{x \in \mathbb{R}_+^E : I \in \Psi(x)\}.$$

First we want to discuss some simple properties of the regions of stability.

Lemma 1

1. *Each region of stability $R(I)$ with $I \in \mathcal{I}$ is an convex polyhedron.*
2. *Two different regions of stability overlap each other only on their boundaries, i.e., $\Psi(x) = I$ for all $x \in \operatorname{int} R(I)$ with $I \in \mathcal{I}$.*
3. *The regions of stability cover the set $\mathbb{R}_+^E$, i.e., $\bigcup_{I \in \mathcal{I}} R(I) = \mathbb{R}_+^E$.*

Proof The third property holds obviously because of $\Psi(x) \neq \emptyset$ for all $x \in \mathbb{R}_+^E$. To prove the other properties we reformulate the definition of $R(I)$. For each $I^* \in \mathcal{I}$ it is

$$\begin{aligned} R(I) &= \{x \in \mathbb{R}_+^E : \ (w^F + x)(I) \leq (w^F + x)(I^*) \text{ for all } I \in \mathcal{I}\} \\ &= \{x \in \mathbb{R}_+^E : \ w^F(I) - w^F(I^*) \leq x(I^*) - x(I) \text{ for all } I \in \mathcal{I}\}. \end{aligned}$$

Since $R(I^*)$ is determined by a finite number of linear inequalities, it is a convex polyhedron. Moreover, the interior is given if all inequalities are replaced by sharp ones. Because of $(w^F + x)(I) = (w^F + x)(I^*)$ for all $x \in R(I) \cap R(I^*)$, this implies the second property. □

An independent set with empty region of stability is not interesting for us since it is never an optimum decision of the follower. Therefore, we consider the following trivial corollary of Theorem 1.

Theorem 3 *It holds $R(I) \neq \emptyset$ if and only if $I \in \mathcal{B}$. Furthermore,*

$$R(I) = \Big\{x \in \mathbb{R}_+^E : x_q - x_e \leq w_e^F - w_q^F \quad \textit{for all } q \notin I,\ e \in I \textit{ with } (I \cup \{q\}) \backslash \{e\} \in \mathcal{I}\Big\} \quad \textit{for all } I \in \mathcal{B}.$$

In the following important corollary of Lemma 1 and Theorem 3 we investigate the interior of the regions of stability. Remark that the interior of a polyhedron is nonempty if and only if it has full dimension.

Corollary 2 *For each $I \in \mathcal{B}$ it holds* $\operatorname{int} R(I) \neq \emptyset$ *and* $\operatorname{cl}(\operatorname{int} R(I)) = R(I)$.

Proof The second property holds since $R(I)$ is a nonempty convex polyhedron. It is still open to prove $\operatorname{int} R(I) \neq \emptyset$ if I is a base of the matroid. But this is a simple consequence of the formula in Theorem 3 since sharp inequalities can be obtained from any $x \in R(I)$ by increasing all x_e by some $\epsilon > 0$ for all $e \notin I$ and by 2ϵ for all $e \in I$, respectively. □

For each base $I \in \mathcal{B}$ of the matroid we define (as in Schrijver 2003) an exchange set

$$A(I) = \{(e, q) : \ e \in I,\ q \in E \backslash I,\ (I \cup \{q\}) \backslash \{e\} \in \mathcal{I}\}.$$

Then Theorem 3 implies

$$R(I) = \{x \in \mathbb{R}_+^E : \ x_e \geq x_q + w_q^F - w_e^F \text{ for all } (e, q) \in A(I)\}.$$

Now we can prove that there exists a minimal element in the region of stability.

Corollary 3 *Let $I \in \mathcal{B}$. We consider the vector $\bar{x}^I \in \mathbb{R}^E$ that is defined by*

$$\bar{x}_e^I = \begin{cases} 0 & \textit{if } e \notin I, \\ \max\{0, \max_{q:(e,q) \in A(I)} (w_q^F - w_e^F)\} & \textit{if } e \in I. \end{cases}$$

Then it holds $\bar{x}^I \in R(I)$ and $\bar{x}^I \leq x$ for all $x \in R(I)$.

Proof Obviously, $\bar{x}^I \in R(I)$. Now, let be given some arbitrary $x \in R(I)$. Then it is $x_q \geq 0 = \bar{x}^I_q$ for all $q \notin I$. For each $e \in I$ Theorem 3 implies $0 \leq x_e$ and

$$w^F_q - w^F_e \leq x_q + w^F_q - w^F_e \leq x_e \quad \text{for all } q \text{ with } (e,q) \in A(I).$$

Thus, the inequality holds also for the maximum of the left-hand terms, i.e., $\bar{x}^I_e = \max_q\{0, w^F_q - w^F_e\} \leq x_e$. Consequently, $\bar{x}^I \leq x$. □

The points $\bar{x}^I$ will be very important for the development of optimality conditions. For each base $I \in \mathcal{B}$ we can compute $\bar{x}^I$ and the set $A(I)$ with $\mathcal{O}(n^2)$ independence tests. Since the independence test for concrete classes of matroids are quite different, it is given by an oracle with running time γ.

Next, we use the results of this section to develop optimality criteria for the leader's problem (LP).

5 Optimality criteria

5.1 Basic concepts

In the previous sections we have investigated the solution sets $\Psi(x)$ and the regions of stability $R(I)$. The results we want to use now to construct optimality conditions for the problems (Opt) and (Pess), i.e., we want to describe points $x \in \mathbb{R}^E_+$ that are local optimal solutions for the optimistic (resp. pessimistic) solution function ϕ_o (resp. ϕ_p).

In the following we use the notation $\text{Locmax}\{\phi(x) : x \in Y\}$ for the set of all local maximum solutions of a function $\phi : Y \to \mathbb{R}$.

First, we need a better description of the optimistic and pessimistic solution functions ϕ_o and ϕ_p. Remember that they are given as the maximum and minimum, respectively, of the function $w^L(I) - c^\top x$ over the solution set $\Psi(x)$. Thus, if $\Psi(x)$ is a singleton, there is nothing to compute. In this case Lemma 1 helps. It tells us that

$$\phi_o(x) = \phi_p(x) = w^L(I) - c^\top x$$

for all $x \in \operatorname{int} R(I)$ with $I \in \mathcal{B}$. In general, let

$$O(I) = \{x \in R(I) : \phi_o(x) = w^L(I) - c^\top x\} \quad \text{and}$$
$$P(I) = \{x \in R(I) : \phi_p(x) = w^L(I) - c^\top x\}$$

denote all $x \in \mathbb{R}^E_+$, such that $I \in \mathcal{B}$ is chosen by the follower in the optimistic and pessimistic case, respectively. Then, for each base I of the matroid it holds

$$\operatorname{int} R(I) \subseteq O(I) \quad \text{and} \quad \operatorname{int} R(I) \subseteq P(I).$$

Thus, Corollary 2 implies the following lemma.

Lemma 2 *The sets $O(I)$ and $P(I)$ are nonempty for all $I \in \mathcal{B}$, and*

$$R(I) = \operatorname{cl} O(I) = \operatorname{cl} P(I).$$

Proof The sets $O(I)$ and $P(I)$ are nonempty because of $\operatorname{int} R(I) \neq \emptyset$.

Due to the definitions it holds $R(I) \supseteq O(I)$ and $R(I) \supseteq P(I)$. Then, since $R(I)$ is a closed set, it has the subsets $\operatorname{cl} O(I)$ and $\operatorname{cl} P(I)$. The other inclusion holds because of $R(I) = \operatorname{cl}\operatorname{int} R(I)$ (Corollary 2), $\operatorname{int} R(I) \subseteq O(I)$, and $\operatorname{int} R(I) \subseteq P(I)$. □

Obviously, if some $x^* \in \mathbb{R}^E_+$ is a local optimal solution of the problems (Opt) and (Pess), respectively, then it is also a local maximum if the feasible set is reduced. For this reason we investigate the set

$$L(I) = \operatorname{Locmax}\{w^L(I) - c^\top x : x \in R(I)\}.$$

If $I \in \mathcal{I}$ is a base, then Lemma 2 implies

$$L(I) = \operatorname{Locmax}\{\phi_o(x) : x \in \operatorname{cl} O(I)\} = \operatorname{Locmax}\{\phi_p(x) : x \in \operatorname{cl} P(I)\}.$$

The following lemma characterizes these sets $L(I)$.

Lemma 3 *It holds $L(I) = \{\bar{x}^I\}$ for each set $I \in \mathcal{B}$.*

Proof Obviously,

$$L(I) = \operatorname{Locmax}\{-c^\top x : x \in R(I)\} = \operatorname{Locmin}\{c^\top x : x \in R(I)\}.$$

Now, remember that $\bar{x}^I \leq x$ for all $x \in R(I)$ (see Corollary 3). Then $c_e > 0$ for all $e \in E$ implies $c^\top \bar{x}^I \leq c^\top x$ for all $x \in R(I)$ with equality only in the case $\bar{x}^I = x$. Thus, $L(I) = \{\bar{x}^I\}$. □

As a result of Lemma 3 we obtain the inclusions

$$\operatorname{Locmax}\{\phi_p(x) : x \geq \mathbf{0}\} \subseteq \{\bar{x}^I : I \in \mathcal{B}\} \quad \text{and}$$
$$\operatorname{Locmax}\{\phi_o(x) : x \geq \mathbf{0}\} \subseteq \{\bar{x}^I : I \in \mathcal{B}\}.$$

Local optimality conditions for bilevel programming problems with a discrete problem of the follower are developed in Fanghänel (2006b), Fanghänel and Dempe (2009). We can specify them easily for the given bilevel matroid problem. Thus, we obtain the following optimality conditions.

Theorem 4 (Fanghänel 2006b; Fanghänel and Dempe 2009)

1. *The point $x \in \mathbb{R}^E_+$ is a local optimistic solution if and only if it holds $x = \bar{x}^I$ for all $I \in \Psi(x)$ with $\phi_o(x) = w^L(I) - c^\top x$.*
2. *The point $x \in \mathbb{R}^E_+$ is a local pessimistic solution if and only if $x = \bar{x}^I$ and $\phi_p(x) = w^L(I) - c^\top x$ for all $I \in \Psi(x)$.*

Checking local optimality conditions is a difficult task in many classes of bilevel problems. Actually, it is NP-hard to check if a given feasible point is a local optimistic solution of a linear bilevel programming problem (Vicente et al. 1994). To provide tests with polynomial running time for the bilevel matroid problem, we simplify the conditions of Theorem 4.

5.2 Local pessimistic solutions

Remember, that $x \in \mathbb{R}_+^E$ is a local pessimistic solution if and only if the following two conditions are satisfied:

1. The value $w^L(I)$ coincides for all bases $I \in \Psi(x)$.
2. It holds $x = \bar{x}^I$ for all $I \in \Psi(x)$.

Then we can prove that $x = \mathbf{0}$ is the only possible local pessimistic solution.

Theorem 5 *It holds* $\operatorname{Locmax}\{\phi_p(x) : x \geq \mathbf{0}\} \subseteq \{\mathbf{0}\}$.

Proof Assume that there exists some base $I \in \mathcal{B}$ such that $x = \bar{x}^I$ is a local pessimistic solution with $x_e > 0$ for some $e \in E$. Then, due to the definition of $\bar{x}^I$ (see Corollary 3) it holds $e \in I$, and there exists some $q \in E \backslash I$ with $x_q = 0$, $x_e = w_q^F - w_e^F > 0$, and $I' = (I \cup \{q\}) \backslash \{e\} \in \mathcal{B}$. Furthermore, because of $w_e^F + x_e = w_q^F + x_q$ it is $(w^F + x)(I) = (w^F + x)(I')$, i.e., $\{I, I'\} \subseteq \Psi(x)$. However, $e \notin I'$ implies $\bar{x}_e^I > 0 = \bar{x}_e^{I'}$. Thus, the optimality conditions of Theorem 4 are violated, i.e., x is not a local pessimistic solution. □

Now, consider the point $x = \mathbf{0}$. Then, because of $\bar{x}^I \leq x$ for all $I \in \Psi(x)$, the second condition is satisfied. The first condition can be tested with the next theorem.

Theorem 6 *The point $x = \mathbf{0}$ is a local pessimistic solution if and only if it holds* $\phi_o(\mathbf{0}) = \phi_p(\mathbf{0})$.

Proof Obviously, if the values $w^L(I)$ coincide for all $I \in \Psi(x)$, also the maximum and the minimum value coincide, i.e., $\phi_o(\mathbf{0}) = \phi_p(\mathbf{0})$.

Reversely, assume that $\phi_o(\mathbf{0}) = \phi_p(\mathbf{0})$. Then, for all $I \in \Psi(x)$ it is

$$w^L(I) - c^\top \mathbf{0} \leq \phi_o(\mathbf{0}) = \phi_p(\mathbf{0}) \leq w^L(I) - c^\top \mathbf{0},$$

i.e., it holds $\phi_o(\mathbf{0}) = \phi_p(\mathbf{0}) = w^L(I)$ for all $I \in \Psi(x)$. □

The condition of Theorem 6 can be tested with Algorithm 3 in polynomial time $\mathcal{O}(\gamma n^2)$. But, even if $x = \mathbf{0}$ is a local pessimistic solution, the set of all global pessimistic solutions can be empty.

5.3 Local optimistic solutions

Let be given some point $x \in \mathbb{R}_+^E$. We want to specify the conditions under which the given point is a local optimistic solution. For this we remember the computation of the optimistic solution function value $\phi_o(x)$ (see Corollary 1 and Algorithm 3). There we have solved the problem

$$\max\{w^L(I) : I \in \Psi(x)\}$$

where $\Psi(x)$ is the set of all bases of the matroid $\mathcal{M}(x)$. Since this is a matroid problem, we can apply Theorem 2. This theorem implies, that the set of all optimum solutions (denoted by $\Psi_o(x)$) is the set of all bases of a matroid that is a directed sum of minors of $\mathcal{M}(x)$ and therefore also of $\mathcal{M}=(E,\mathcal{I})$.

Now, remember the optimality conditions of Theorem 4. We can rewrite them by using the solution set

$$\Psi_o(x)=\{I\in\Psi(x):\ w^L(I)\geq w^L(J)\text{ for all }J\in\Psi(x)\}.$$

Then, $x\in\mathbb{R}_+^E$ is a local optimistic solution if and only if $x=\bar{x}^I$ for all $I\in\Psi_o(x)$.

Thus, as a first step in an optimality test we apply Algorithm 3 and compute a base $I\in\mathcal{B}$ with $I\in\Psi_o(x)$. This is done in polynomial time $\mathcal{O}(\gamma n^2)$.

Then, we have to compare the points $\bar{x}^J$ for different sets $J\in\Psi_o(x)$. To do this, we need to know more about the sets

$$A(J)=\{(e,q):\ e\in J,\ q\notin J,\ (J\cup\{q\})\backslash\{e\}\in\mathcal{I}\}$$

from Sect. 4. A useful result on these sets is given in Schrijver (2003) (Lemma 40.4α). Since we want to apply it, we state the lemma but omit the proof.

Lemma 4 *Let $\mathcal{M}=(E,\mathcal{I})$ be a matroid and $I\in\mathcal{I}$. Let $(t,s)\in A(I)$ and $J=(I\cup\{s\})\backslash\{t\}$. Then, $(e,q)\in A(I)\backslash A(J)$ implies*

1. *$e=t$ or $(e,t)\in A(J)$, and*
2. *$s=q$ or $(s,q)\in A(J)$.*

Using Lemma 4 we can characterize the local optimistic solutions.

Theorem 7 *Let $I\in\Psi_o(x)$. Then, $x=\bar{x}^I$ is a local optimistic solution if and only if for each pair $(e,q)\in A(I)$ with $w_e^L=w_q^L$ and $w_e^F+x_e=w_q^F$ it holds $x_e=0$.*

Proof We want to show that the given condition coincides with the condition of Theorem 4, i.e., we prove that $\bar{x}^J=x$ for all sets $J\in\Psi_o(x)$ if and only if there exists no $(e,q)\in A(I)$ with $w_e^L=w_q^L$, $w_e^F+x_e=w_q^F$, and $x_e>0$.

$\Rightarrow$: This part of the proof is analogously to the proof of Theorem 5. We assume that there exists some pair $(e,q)\in A(I)$ with $x_e>0$, $w_e^L=w_q^L$, and $w_e^F+x_e=w_q^F$. Then it holds $J=(I\cup\{q\})\backslash\{e\}\in\Psi_o(x)$. However, since $e\notin J$, it holds $\bar{x}_e^J=0<\bar{x}_e^I=x_e$. Thus, it is not $x=\bar{x}^J$ for all sets $J\in\Psi_o(x)$.

$\Leftarrow$: Assume that there exists some set $J\in\Psi_o(x)$ with $\bar{x}^J\neq x$. Because of $J\in\Psi(x)$ and Corollary 3 it holds $x\geq\bar{x}^J$, and there exists some $e\in I$ with

$$0\leq\bar{x}_e^J<x_e=\bar{x}_e^I.$$

Furthermore, there exists some $q\notin I$ with $w_q^F=w_e^F+x_e>w_e^F$ and $(e,q)\in A(I)$ (see the definition of $\bar{x}^I$ in Corollary 3).

Since $\Psi_o(x)$ is the set of bases of a matroid, we can assume, w.l.o.g., that there exists a pair of elements $(t,s)\in A(I)$ with $J=(I\cup\{s\})\backslash\{t\}$. Because of $J\in\Psi_o(x)$ it is $w_s^L=w_t^L$ and $w_t^F+x_t=w_s^F$.

If $t = e$, then we are finished. With $q = s$ we have found some $q \notin I$, such that $(e, q) \in A(I)$, $w_e^L = w_q^L$, and $w_e^F + x_e = w_q^F$. Analogously, we are finished if $x_t > 0$.

Therefore we suppose that it holds $e \neq t$ and $x_t = 0$, and thus, $e \in J$ and $w_t^F = w_s^F$. Furthermore, it is $(e, q) \in A(I)\backslash A(J)$ because of $e \in J$ and

$$\bar{x}_e^J = \max\{w_z^F - w_e^F : (e, z) \in A(J)\} < w_q^F - w_e^F = x_e.$$

Then, Lemma 4 and $e \neq t$ imply $(e, t) \in A(J)$. This yields

$$w_t^F - w_e^F \leq \bar{x}_e^J < \bar{x}_e^I = w_q^F - w_e^F, \quad \text{i.e.,} \quad w_s^F = w_t^F < w_q^F.$$

Consequently, it is also $s \neq q$. Now, applying Lemma 4 again, we obtain $(s, q) \in A(J)$. But then it is $\bar{x}_s^J = 0 = x_s \geq w_q^F - w_s^F$. This is a contradiction to $w_s^F = w_t^F < w_q^F$. □

Remark that $x = \mathbf{0}$ is always a local optimistic solution since there exists some $I \in \Psi(\mathbf{0})$ and because $0 \leq \bar{x}^J \leq \bar{x}^I$ for all $J \in \Psi_o(\bar{x}^I)$ implies the condition of Theorem 7.

We want to test if a given point x is a local optimistic solution. Because of Theorem 7 this can be done with the following algorithm.

Algorithm 4 (Optimality test for the optimistic case)

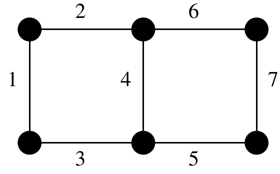
1. *Compute a base* $I \in \Psi_o(x)$ *with Algorithm* 3.
2. *Compute the set* $A = A(I)$ *and* $\bar{x}^I$. *If* $x \neq \bar{x}^I$, *then goto Step* 5.
3. *While* $A \neq \emptyset$ *do*:
   (a) *Choose some pair* $(e, q) \in A$ *and set* $A = A\backslash\{(e, q)\}$.
   (b) *If* $x_e = w_q^F - w_e^F$, $w_e^L = w_q^L$, *and* $x_e > 0$, *then goto Step* 5.
4. *Result*: *The point* $x = \bar{x}^I$ *is a local optimistic solution. Stop.*
5. *Result*: *The point* $x = \bar{x}^I$ *is not a local optimistic solution. Stop.*

Obviously, also this algorithm has a polynomial running time $\mathcal{O}(\gamma n^2)$. Remark that the vector c has no influence on local optimality. However, it is important to find a best local optimistic solution, i.e., it is needed for discussing global optimality.

5.4 An example

Finally we apply the optimality conditions on a small example. For this we consider again the situation from Example 1.

Example 2 We proceed with Example 1, i.e., we investigate the graphical matroid with $E = \{1, 2, 3, 4, 5, 6, 7\}$ and the following graph G:



We assume that for the bilevel problem the weights and costs are $w^L = (1, -1, 0, 1, 1, 5, 4)$, $w^F = (4, 4, 2, 2, 2, 1, 1)$, and $c = (1, 1, 1, 1, 1, 1, 1)$.

First we consider the pessimistic case. Due to Theorem 5 only $x = \mathbf{0}$ can be a local pessimistic solution. Local pessimistic optimality in this point can be verified by comparing the values of the optimistic and pessimistic solution function (Theorem 6). In Example 1 we have computed $\phi_o(\mathbf{0}) = 7$ and $\phi_p(\mathbf{0}) = 5$. Thus, there exists no local pessimistic solution.

Now we want to investigate some points $x \in \mathbb{R}^E_+$ for local optimistic optimality.

1. Let $x = \mathbf{0}$, i.e., the leader does not increase the weights of the follower. Then we know from Example 1 that $I_1 = \{1, 2, 4, 5, 6\} \in \Psi_o(\mathbf{0})$ and $\phi_o(x) = w^L(I_1) = 7$. This is a local optimistic solution.
2. Now, consider $x = (0, 0, 2, 2, 0, 1, 1)$. Then it holds $I_2 = \{1, 3, 4, 6, 7\} \in \Psi_o(x)$. Obviously, this base is a optimum solution of the matroid problem $\max\{w^L(I) : I \in \mathcal{B}\}$ that arises if the leader can choose an optimum base by himself. In Step 2 of Algorithm 4 we obtain

$$A = A(I_2) = \{(1, 2), (3, 2), (4, 2), (4, 5), (6, 5), (7, 5)\}$$

and $x = \bar{x}^{I_2}$. For all pairs $(e, q) \in A$ with $x_e + w^F_e = w^F_q$ it holds $w^L_e > w^L_q$. Hence, $x = (0, 0, 2, 2, 0, 1, 1)$ is a local optimistic solution. The value is

$$\phi_o(x) = w^L(I_2) - c^\top x = 11 - 6 = 5.$$

3. Let $x = (0, 0, 2, 0, 0, 1, 1)$. Then, $I_3 = \{1, 2, 3, 6, 7\} \in \Psi(x)$. But because of $\bar{x}^{I_3} = (0, 0, 0, 0, 0, 1, 1) \neq x$ the point $x = (0, 0, 2, 0, 0, 1, 1)$ is not a local optimistic solution. However, remark that $x = \bar{x}^{I_3} = (0, 0, 0, 0, 0, 1, 1)$ is not only a local optimistic solution but even a global optimistic one with $\phi_o(x) = 8$.

Consequently, there exist many local optimistic solutions in the example. The best one is neither the point $x = \mathbf{0}$, where the leader does not influence the behaviour of the follower, nor the point $x = (0, 0, 2, 2, 0, 1, 1)$ where the follower chooses the optimum base of the problem $\max\{w^L(I) : I \in \mathcal{B}\}$.

6 Conclusions

In this paper we have presented a bilevel problem with a matroid problem in the lower level and continuous variables of the leader. After discussing the solution sets we have introduced and characterized so-called regions of stability. These results we have used to develop optimality conditions for both local optimistic and pessimistic solutions. We have shown that these conditions can be verified in polynomial running time $\mathcal{O}(\gamma n^2)$.

References

Bard JF (1998) Practical bilevel optimization—algorithms and applications. Kluwer Academic, Dordrecht

Cook W, Cunningham W, Pulleyblank W, Schrijver A (1998) Combinatorial optimization. Wiley, New York

Dempe S (2002) Foundations of bilevel programming. Kluwer Academic, Dordrecht

Dempe S (2003) Annotated bibliography on bilevel programming and mathematical programs with equilibrium constraints. Optimization 52:333–359

Dempe S, Kalashnikov V, Ríos-Mercado RZ (2005) Discrete bilevel programming: application to a natural gas cash-out problem. Eur J Oper Res 166(2):469–488

Fanghänel D (2006a) Optimality criteria for bilevel programming problems using the radial subdifferential. In: Dempe S, Kalashnikov V (eds) Optimization with multivalued mappings: theory, application and algorithms. Springer, Berlin, pp 73–95

Fanghänel D (2006b) Zwei Ebenen Optimierung mit diskreter unterer Ebene. Dissertation, TU Bergakademie Freiberg. Available via http://fridolin.tu-freiberg.de/dissliste.html

Fanghänel D, Dempe S (2009) Bilevel programming with discrete lower level. Optimization 58(8):1029–1047

Fujishige S (2005) Submodular functions and optimization, 2nd edn. Annals of Discrete Mathematics, vol 58. Elsevier, Amsterdam

Heuberger C (2004) Inverse combinatorial optimization: a survey on problems, methods, and constraints. J Comb Optim 8:329–361

Kalashnikov V, Ríos-Mercado RZ (2001) A penalty-function approach to a mixed integer bilevel programming problem. In: Zozaya C et al (eds) Proceedings of the 3rd international meeting on computer science, Aquascalientes, Mexico, vol 2, pp 1045–1054

Schrijver A (2003) Combinatorial optimization: polyhedra and efficiency. Springer, Berlin

Vicente LN, Savard G, Judice JJ (1994) Descent approaches for quadratic bilevel programming. J Optim Theory Appl 81:379–399

Vicente LN, Savard G, Judice JJ (1996) The discrete linear bilevel programming problem. J Optim Theory Appl 89:597–614