BILEVEL OPTIMIZATION PROBLEMS WITH VECTORVALUED OBJECTIVE FUNCTIONS IN BOTH LEVELS

S. DEMPE, S. FRANKE

Abstract. Bilevel optimization problems with multivalued objective functions in both levels are first replaced by a problem with a parametric lower level using a convex combination of the lower level objectives. Thus a nonconvex multiobjective bilevel optimization problem arises which is then transformed into a parametric bilevel programming problem. The investigated problem has been considered in the paper [11] using ideas from fuzzy optimization. It is one goal to correct some of the wrong ideas in the paper [11].

1. Introduction

Bilevel programming problems were first introduced by von Stackelberg in 1934 [15]. They can be interpreted as hierarchical games of two decision makers where the first decision maker (the so-called leader) has the first choice, and the second one (the so-called follower) reacts optimally to the leader’s selection. It is the leader’s aim to find such a decision which, together with the optimal response of the follower, optimizes the objective function of the leader.

To model this problem, we formulate the (parametric) optimization problem of the second decision maker first:

\[
\min_{y} \{ f(x, y) : g(x, y) \leq 0 \},
\]

where \( f, g_i : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R} \), \( i = 1, \ldots, p \). If \( \Psi(x) \) denotes the set of optimal solutions of problem (1.1), the bilevel programming problem (i.e. the leader’s problem) can be formulated as follows:

\[
" \min_{x} " \{ F(x, y) : x \in X, y \in \Psi(x) \},
\]

where \( X \subseteq \mathbb{R}^n \) is a closed, nonempty set and \( F : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R} \). This problem has been topic of a large number of articles, see the bibliography [8], and of at least two monographs [2, 7]. If the optimal solution of the problem (1.1) is not uniquely determined for at least one value of the parameter \( x \), problem (1.2) is in general not well posed since the leader is not able to compute his / her objective function value. To indicate this ambiguity, we used the quotation marks in the formulation (1.2). Is this the case, the problem is usually replaced with its optimistic resp. pessimistic formulation, see e.g. [7]. We will use the optimistic formulation and replace problem (1.2) by

\[
\min_{x, y} \{ F(x, y) : x \in X, y \in \Psi(x) \}.
\]
In the paper [11], the objective functions in both levels are vectorvalued functions, i.e. \( F : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^k \) and \( f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^q \). In this case, as long as no feasible point exists which minimizes all objective functions at the same time, problem (1.1) will not have a unique solution and the model (1.3) will be considered. Here, using definitions in multicriterial optimization, see e.g. [9], \( \Psi(x) \) denotes the set of (weak) Pareto optimal solutions. The point \( y \) with \( g(x, y) \leq 0 \) is a Pareto optimal (weak Pareto optimal) solution of problem (1.1) at \( x = x^* \) provided there does not exist \( \hat{y} \) with \( g(x, \hat{y}) \leq 0 \), \( f(x, \hat{y}) \leq f(x, y) \) and \( f(x, \hat{y}) \neq f(x, y) \) \( (f_i(x, \hat{y}) < f_i(x, y) \) for \( i = 1, \ldots, q) \). In this case, the set \( \Psi(x) \) will contain in general more than one point, and problem (1.3) needs to be used.

Bilevel optimization problems with multiple objectives in the lower resp. the upper level problem have been investigated rather seldom in the literature [10]. An application in the case of linear objectives in both levels is given by transportation management, see [16]. Bonnel et al. propose a penalty approach for the semivectorial bilevel programming problem [6]. An algorithm for optimizing a linear function over a Pareto set using sequentially improved relaxations of the efficient set is proposed by [3].

In problem (1.3), the leader assumes that the follower will select that feasible solution out of his/her set of optimal solutions \( \Psi(x) \) which minimizes the leader’s objective function over this set. This assumption is not always possible (in economical applications it can be forbidden; also, if the follower is nature it is not possible to force some specific action). In some other situations there may exist other rules used by the follower to select the true response on the leader’s selection. We assume here that the leader knows the rule applied by the follower. In the paper [11] it is assumed that the follower will select a feasible solution minimizing the distance (using Tchebyshief- or \( l_\infty \)-norm) of the objective function value from the ideal point \( (Z_1^U, \ldots, Z_q^U) \). Note that the computed solution needs not to be unique (making the application of the above problem (1.3) necessary), nor is the computed solution Pareto optimal in general. The first remark is implied by the norm which is not strictly convex, the second one is a consequence of using fixed values for the parameters in computing the ideal point.

The content of the paper is as follows. We will investigate the problem (1.3) with multiple objective functions in both levels in the linear case. In Section 2, some known results on multicriterial optimization are collected which will then be used in Section 3 to formulate an ordinary optimization problem that helps to solve the bilevel programming problem (1.3). In Section 4, some properties of the initial problem are derived which are used to solve it. Then, a respective solution algorithm is presented. We will illustrate this algorithm using a small example. Section 5 concludes the paper.

The notation is standard. \( \mathbb{R}^n_+ \) denotes the \( n \)-dimensional nonnegative orthant.

2. Auxiliary results on multicriterial optimization

First, consider a multicriterial optimization problem

\[
\text{min} \{ \alpha(x) : x \in M \},
\]

where \( \emptyset \neq M \subseteq \mathbb{R}^n \) is a closed set and \( \alpha : \mathbb{R}^n \rightarrow \mathbb{R}^q \). Then, the common optimality definitions need to be adjusted since usually no feasible point \( x \in M \) optimizes
every objective function simultaneously. We begin by introducing a suitable partial order.

**Definition 2.1.** Let $y, z \in \mathbb{R}^q$. $y \succeq z$ if it holds that
\[ y_i \geq z_i \quad \forall i = 1, \ldots, q \quad \text{and} \quad \exists q \in \{1, \ldots, q\} : y_q > z_q. \]

This allows us to formulate a definition for an optimal solution of a vector optimization (or multicriterial optimization) problem.

**Definition 2.2.** A feasible point $x$ of (2.1) is called Pareto optimal if there does not exist a feasible point $x \in M$ with $\alpha(x) \succeq \alpha(x)$.

Pareto optimality can be interpreted as the nonexistence of a point in the image space having a "better" objective function value in the sense of the introduced partial order. For our investigations, we will only use this optimality definition. There exist various other common definitions like weak or proper Pareto optimality.

The following results are useful for computing Pareto optimal solutions. They can be found e.g. in [9].

**Theorem 2.3.** Let $\lambda \in \mathbb{R}^q$, $\lambda_i > 0$ for $i = 1, \ldots, q$. Let $x \in M$ be an optimal solution of the problem
\[ (2.2) \min \{\lambda^\top \alpha(x) : x \in M\}. \]
Then, $\lambda$ is Pareto optimal for problem (2.1).

**Theorem 2.4.** Let $y < \alpha(x)$ for all $x \in M$ (e.g. let $y_i < \min \{\alpha_i(x) : x \in M\}$, $i = 1, \ldots, q$) and $x \in M$ be an optimal solution of the problem
\[ (2.3) \min \{\|\alpha(x) - y\| : x \in M\}, \]
where $\|\cdot\|$ denotes the Euclidean norm in $\mathbb{R}^q$. Then, $x$ is Pareto optimal for problem (2.1).

It should be noted that the Euclidean norm in Theorem 2.4 can be replaced by any strictly convex function. In the paper [11], a fuzzy optimization approach [17] is used to replace problem (2.1). This reduces to solving the problem
\[ (2.4) \max \{\lambda(x, y) : x \in M\}, \]
where $y < \alpha(x)$ for all $x \in M$ and $\lambda(x, y) = \min \{t_i(\alpha_i(x) - y_i)\}$ with $t_i > 0$ for $i = 1, \ldots, q$. The function $x \mapsto \lambda(x, y)$ is concave, but not strictly concave. The solution $x$ of problem (2.4) is, hence, a weak Pareto optimal solution, so there does not exist $x \in M$ such that $\alpha_i(x) > \alpha_i(\bar{x})$ for all $i = 1, \ldots, q$.

The possibility to compute all Pareto optimal solutions by using convex combinations of the objective functions is restricted to convex multicriterial optimization problems:

**Theorem 2.5.** Consider problem (2.1) with $M$ being a convex set and $\alpha_i$ being convex functions for $i = 1, \ldots, q$. Then, for each Pareto optimal solution $x \in M$ there exists a vector $\lambda \in \mathbb{R}^q, \lambda_i \geq 0$, $i = 1, \ldots, q$, $\sum_{i=1}^q \lambda_i = 1$ such that $x$ is an optimal solution of problem (2.2).

Using parametric linear optimization and the vertex property of optimal solutions in linear optimization, it can be shown that $\lambda$ can be found such that $\lambda_i > 0$ for $i = 1, \ldots, q$ if $M$ is a polyhedral set and all functions $\alpha_i$ are linear.
Theorem 2.6. If $\pi \in M$ is a Pareto optimal solution of problem (2.1), then there exists $y \in \mathbb{R}^q$ such that $\pi$ is an optimal solution of problem (2.3).

Using Theorems 2.3 and 2.5, it is easy to prove that the set of Pareto optimal solutions of linear multicriterial optimization problems equals the union of certain faces of the polyhedron $M$ (which equals the intersection of a finite number of half spaces bounded by faces of the polyhedron). This set is connected (i.e. not equal to the union of two or more sets which can be strongly separated). Moreover, this set is in general not convex.

3. Reformulation of the lower level as a nonlinear optimization problem

In this paper, we want to investigate the linear bilevel programming problem in combination with vector optimization. In both levels, there is a certain number of objective functions, possibly conflicting with each other, that need to be optimized. The leader solves (note that we maximize the problems in both levels as it is usual in linear optimization)

$$\max_{x \in X} (c_1^T x, c_2^T x, \ldots, c_k^T x) + (f_1^T y, f_2^T y, \ldots, f_k^T y)$$

s.t. $A_1 x + B_1 y \leq b_1$

$x \geq 0$

$y \in \Psi(x)$

with the solution set mapping $\Psi(x)$ of the follower’s problem

$$\max_{y \in Y} (d_1^T y, d_2^T y, \ldots, d_p^T y)$$

$$A_2 x + B_2 y \leq b_2$$

$y \geq 0$,

for $c_i \in \mathbb{R}^n$ ($i = 1, \ldots, k$), $f_l \in \mathbb{R}^m$ ($l = 1, \ldots, k$), $d_j \in \mathbb{R}^m$ ($j = 1, \ldots, p$), $X \subseteq \mathbb{R}^n$ and $Y \subseteq \mathbb{R}^m$ being polyhedral sets and matrices $A_1, B_1, A_2, B_2$ as well as vectors $b_1, b_2$ of appropriate dimensions. Let $C^T = (c_1 \ldots c_k)$, $F^T = (f_1 \ldots f_k)$ and $D^T = (d_1 \ldots d_p)$ be the matrices composed of the objective function coefficients of all the objective functions of both the leader and the follower. Then, for fixed $x$, and using Theorems 2.3 and 2.5, the set $\Psi(x)$ of Pareto optimal solutions of problem (3.2) equals

$$\Psi_s(x, \lambda) = \arg\max_{y \in Y} \{ \lambda^T D y : A_2 x + B_2 y \leq b_2, \ y \geq 0 \}.$$ 

Here

$$\Psi_s(x, \lambda) = \arg\max_{y \in Y} \{ \lambda^T D y : A_2 x + B_2 y \leq b_2, \ y \geq 0 \}.$$ 

Now, if the optimistic approach (1.3) is used, problem (3.1) reduces to

$$\max_{x,y} \{ C x + F y : A_1 x + B_1 y \leq b_1, \ x \geq 0, \ y \in \Psi(x) \}$$

$$= \max_{x,y,\lambda} \{ C x + F y : A_1 x + B_1 y \leq b_1, \ x \geq 0, \ y \in \Psi_s(x, \lambda), \ \lambda \in \Lambda \}.$$ 

The last problem is a multiobjective nonconvex optimization problem. To compute a Pareto optimal solution for it, Theorem 2.4 can be used.
Remark 3.1. Problem (3.1) has been formulated using so-called coupling constraints $A_1x + B_1y \leq b_1$. To prove their satisfaction, the leader needs the solution selected by the follower. This implies that the leader can recognize his/her selection as being feasible only after being told the follower’s selection. Mathematically, this implies that the feasible solution set

$$\{(x, y) \in \mathbb{R}^n \times \mathbb{R}^m : A_1x + B_1y \leq b_1, \ x \geq 0, \ y \in \Psi(x)\}$$

of the leader’s problem is now disconnected [12]. Shifting these constraints to the lower level problem will modify the feasible set of the leader’s problem [1, 12]. Hence, those constraints are difficult to handle and are, therefore, usually avoided.

4. Properties of the linear multiobjective bilevel optimization problem

We will now attempt to transform the multiobjective linear bilevel programming problem into a linear multicriterial optimization problem. We need the following property of a parametric linear multicriterial optimization problem.

**Theorem 4.1.** Consider the parametric multicriterial optimization problem (3.2) and let

$$\Psi := \bigcup_{x \in \mathbb{R}^n_+ \cap X} \Psi(x)$$

denote the union over all sets of Pareto optimal solutions for every the parametric optimization problem (3.2). Then, $\Psi$ equals the union of faces of the set $T := \{(x, y) \in (\mathbb{R}^n_+ \cap X) \times \mathbb{R}^m : A_2x + B_2y \leq b_2, \ y \geq 0\}$.

**Proof.** Let $\overline{y} \in \Psi(\overline{x})$ for some $\overline{x} \in \mathbb{R}^n$. Then, by Theorem 2.5, there exists $\lambda \in \Lambda$ such that $\overline{y} \in \Psi(\lambda, \overline{x})$. Hence, by convex optimization, the gradient $\lambda^\top D$ of the objective function in problem (3.3) belongs to the normal cone of the feasible set of this problem [14, Theorem 3.24]. If $(\overline{x}, \overline{y})$ is an element of the relative interior of one face of the set $T$, then $\overline{y}$ is an element in the relative interior of the set $T(\overline{x}) := \{y : B_2y \leq b_2 - A_2\overline{x}, \ y \geq 0\}$ and, hence (by the definition of the normal cone), $\lambda^\top D$ is parallel to the normal vector (which is unique) of $T(\overline{x})$. Since this is true for every element $y \in T(\overline{x})$, we derive the well-known property that $T(\overline{x}) \subseteq \Psi(\lambda, \overline{x})$ which implies by the remark after Theorem 2.5 that $\Psi(\overline{x})$ equals the union of faces of the set $T(\overline{x})$. The latter property is maintained if the set $T(x)$ is shifted parallel which happens if the parameter $x$ is changed. This now implies that every element of the face of $T$ containing $(\overline{x}, \overline{y})$ in its relative interior belongs to the set $\Psi$. This verifies the statement. □

Now, consider the bilevel programming problem (3.1), (3.2) and assume that the coupling constraints disappear and that $X$ is a polyhedral set. Then, in the optimistic formulation, this problem is equivalent to

$$\max_{x, y} \{Cx + Fy : x \in X, \ x \geq 0, \ y \in \Psi(x)\}$$

(4.1)

$$= \max_{x, y} \{Cx + Fy : x \in X, \ x \geq 0, \ (x, y) \in \text{gph } \Psi\},$$

where $\text{gph } \Psi := \{(x, y) : y \in \Psi(x)\}$ denotes the graph of the point-to-set mapping $\Psi$. 
Problem (4.1) is the problem of maximizing a linear function over a (nonempty) set. This equals maximizing this function over the convex hull of the set
\[
\max \{ Cx + y : x \in X, (x, y) \in \text{conv gph } \Psi \}.
\]
The feasible set of problem (4.2) is a polyhedron, the problem itself is a linear multiobjective optimization problem. Unfortunately, its feasible set is not known. Nevertheless, the set of Pareto optimal solutions of this problem equals the union of faces of the set \( \{(x, y) \in \text{conv gph } \Psi \} \).

To compute one Pareto optimal solution of problem (4.1) we can use a convex combination of the upper level objective functions and solve
\[
\max_{x,y} \mu^\top (Cx + y) : x \in X, x \geq 0, (x, y) \in \text{gph } \Psi \}
\]
with \( \mu \in \mathcal{M} := \{ \mu \in \mathbb{R}^k : \mu_i > 0, i = 1, \ldots, k, \sum_{i=1}^k \mu = 1 \} \). This problem can e.g. be solved using the k-th best algorithm [5] tailored to problem (4.3). The algorithm is based on the following important property.

**Lemma 4.2.** [4] Let \( K := \{(x, y) \in X \times \mathbb{R}^m : A_2 x + B_2 y \leq b_2, x \geq 0, y \geq 0 \} \) be nonempty and compact. Then, a solution of (4.3) occurs at an extreme point of \( K \).

This motivates the following approach. First for some \( \mu \in \mathcal{M} \), a solution of
\[
\max_{x,y} \mu^\top (Cx + y) : x \in X, x \geq 0, (x, y) \in \text{gph } \Psi \}
\]
is computed. Then it is verified whether the solution is Pareto optimal for the lower level (3.2) with fixed upper level parameter \( x \). If this is not the case, the adjacent extreme points of the current solution are considered in order of non-increasing function values. Since \( K \) is compact, a feasible point with respect to the lower level and, hence, a globally optimal solution of (4.3) is found after a finite number of steps. An extensive study of this approach has been done in [13].

In the \( i \)-th iteration, \( W^i \) denotes the adjacent extreme points \( x \) of \( \mathcal{F}^i \) with \( \mu^\top (Cx + Fy) \leq \mu^\top (Cx + Fy) \) for some \( \mu \). \( T \) denotes the set of already examined, infeasible points, while \( W \) contains the points that still need to be considered.

**A k-th best algorithm**

**Initialization.** Set \( i := 1 \) and \( W = T = W^i := \emptyset \). Choose \( \mu \in \mathcal{M} \).

**Step 1.** Solve (4.4) using the simplex method; the optimal solution is \((\mathcal{F}^i, \mathcal{F}^i)\). Set \( W := \{(\mathcal{F}^i, \mathcal{F}^i)\} \).

**Step 2.** Solve the lower level problem with fixed upper level parameter \( \mathcal{F}^i \). If \( \mathcal{F}^i \in \mathcal{U}_\mathcal{F} \), stop. The solution \((\mathcal{F}^i, \mathcal{F}^i)\) is globally optimal for (4.3). Otherwise, go to Step 3.
Step 3. Set $T := T \cup \{(\pi^i, \overline{y}^i)\}$. $W := (W \cup W^i) \setminus T$, increase $i$. Choose $(\pi^i, \overline{y}^i)$ with $\mu^\top (C\pi^i + F\overline{y}^i) = \max_{(x, y) \in W} \mu^\top (Cx + Fy)$ and go to Step 2.

Example 1. Consider the multicriterial bilevel programming problem
\[
\begin{align*}
\max_{x \in \mathbb{R}, y \in \mathbb{R}^2} & (x - 2y_2, 2x) \\
\text{s.t.} & \quad 2 \leq x \leq 5, \\
& \max_{y \in \mathbb{R}^2} (y_1 + 2y_2, y_1 - y_2) \\
& \quad (x, y) \in K
\end{align*}
\]
with
\[
K := \{(x, y) \in \mathbb{R}^3 : \begin{array}{l}
y_1 \leq 6, \\
x + y_1 + y_2 \leq 10, \\
-x + y_2 \leq 0, \\
y_1, y_2 \geq 0 \end{array}\}.
\]

Since the lower level feasible set is compact provided that it is not empty, we only have to guarantee the latter property. A quick look shows that $K$ is not empty for all $x \in [2, 5]$. We can apply the $k$-th best algorithm.

Initialization. Set $i := 1$ and $W = T = W^i := \emptyset$, choose $\mu := (0.5, 0.5)$.

First iteration

Step 1. Solving $\max_{(x, y) \in K} \mu^\top (Cx + Fy) = \max_{(x, y) \in K} (1.5x - y_2)$ leads to the optimal solution $(\pi^1, \overline{y}^1) = (5, 0, 0)$ with $\mu^\top (C\pi^1 + F\overline{y}^1) = 7.5$.

Step 2. The Pareto set of the lower level with $\pi^1 = 5$ is
\[
\bigcup_{\lambda \in \Lambda} \Psi_s(5, \lambda) = \text{conv} \{(5, 0), (0, 5)\}
\]
which means that $\overline{y}^1 = (0, 0)$ is not a Pareto optimal point. Hence, the solution obtained in Step 1 is not globally optimal for the bilevel problem.

Step 3. Set $T := \{(5, 0, 0)\}$. The adjacent extreme points of $(\pi^1, \overline{y}^1)$ are added to $W$, namely $W := \{(0, 0, 0), (5, 0, 0), (5, 5, 0), (5, 0, 5)\}$. We choose $(\pi^2, \overline{y}^2) = (5, 5, 0)$ with $\mu^\top (C\pi^2 + F\overline{y}^2) = 7.5$.

Second iteration

Step 2. Now, solving the lower level problem again for $\pi^2 = 5$ leads to
\[
\overline{y}^2 = (5, 0) = \bigcup_{\lambda \in \Lambda} \Psi_s(5, \lambda) = \text{conv} \{(5, 0), (0, 5)\}.
\]
The algorithm terminates with the globally optimal solution $(\pi^2, \overline{y}^2) = (5, 5, 0)$.

5. Conclusion

We considered a linear bilevel programming problem with multivalued objective functions in both upper and lower level. The problem was suitably reformulated...
using vector optimization theory. The possibility of solving the resulting problem with a k-th best algorithm has been demonstrated.

References


TU Bergakademie Freiberg, Germany
E-mail address: {dempe; susanne.franke}@math.tu-freiberg.de