A Fenchel-Lagrange Duality Approach for a Bilevel Programming Problem with Extremal-Value Function

Abdelmalek Aboussoror, Samir Adly

▶ **To cite this version:**

HAL Id: hal-00683042
https://hal.archives-ouvertes.fr/hal-00683042

Submitted on 20 Oct 2017

A Fenchel–Lagrange Duality Approach for a Bilevel Programming Problem with Extremal-Value Function

Abdelmalek Aboussoror · Samir Adly

Abstract In this paper, for a bilevel programming problem (S) with an extremal-value function, we first give its Fenchel–Lagrange dual problem. Under appropriate assumptions, we show that a strong duality holds between them. Then, we provide optimality conditions for (S) and its dual. Finally, we show that the resolution of the dual problem is equivalent to the resolution of a one-level convex minimization problem.

Keywords Two-level optimization · Convex analysis · Conjugate duality · Composed programming problems

1 Introduction

We are concerned with the following bilevel nonlinear optimization problem

$$(S) \qquad \operatorname*{Min}_{x \in X} F(x, v(x)),$$

where $v(x)$ is the optimal value of the lower level problem

$$\mathcal{P}(x) \qquad \operatorname*{Min}_{y \in Y} f(x, y).$$

A. Aboussoror

Faculté Polydisciplinaire de Safi, Département de Mathématiques et Informatique, Université Cadi Ayyad, B.P. 4162, Sidi Bouzid, Safi, Morocco
e-mail: aboussororabdel@hotmail.com

S. Adly (✉)
Laboratoire XLIM UMR-CNRS 6172, Département de Mathématiques, Université de Limoges, 123 Avenue Albert Thomas, 87060 Limoges Cedex, France
e-mail: samir.adly@unilim.fr

$F : \mathbb{R}^p \times \mathbb{R} \to \mathbb{R}$, $f : \mathbb{R}^p \times \mathbb{R}^q \to \mathbb{R}$; X and Y are two compact convex subsets of $\mathbb{R}^p$ and $\mathbb{R}^q$, respectively. The problem (S) is called a bilevel problem with the extremal-value function $v(\cdot)$. It corresponds to a two-player game in which a leader announces first a strategy $x \in X$, to minimize his objective function F, and a follower reacts optimally by selecting a strategy $y(x) \in Y$, to minimize his objective function f. We assume that the reaction of the follower to the strategies of the leader has no impact on the constraints of the leader. Then, when the leader evaluates the performance of the follower by the optimal value $v(x)$, he includes it in the objective function F, which leads to the formulation considered above. For many applications of such a problem, we refer, for example, to the interesting book [1]. However, when the performance of the follower is evaluated by his solutions, we obtain different formulations of the leader's problem, as the optimistic and the pessimistic formulations. For these interesting formulations, we refer, for example, to [1–9], respectively. Note that (S) can be viewed as a composed programming problem. This remark will be the key of our investigation. As it is well-known, the problem (S) presents difficulties in the theoretical and numerical aspects, since its formulation includes a marginal function which, in general, is not differentiable. For different approaches concerning such a class of bilevel programming problems, we refer, for example, to [1, 10–13].

In [10], for the problem (S), global optimality conditions using min-max problems are given. These optimality conditions are based on a notion of stability of optimization problems and regularity of solutions. Other optimality conditions for some classes of bilevel programming problems with the extremal-value function using different directional derivatives and the subdifferential of the marginal function $v(\cdot)$ are given in [1, 11–13]. In this paper, adopting the Fenchel–Lagrange duality, which has been first introduced by Boţ and Wanka in [14] for ordinary convex programming problems, we give the so-called Fenchel–Lagrange dual of the bilevel problem (S). Based on some results given in [15–18], we first show that a strong duality holds between them. Then, we provide necessary and sufficient optimality conditions for (S) and its dual. These optimality conditions are given in terms of conjugate functions and subdifferentials, respectively. Finally, we show that the resolution of the dual problem is equivalent to the resolution of a one-level convex minimization problem whose objective function is expressed in terms of the conjugates of the functions involved. Assume that the obtained dual problem is easier to solve than (S). Therefore, this approach may give a possibility to solve indirectly the problem (S) via a one-level convex minimization problem.

The paper is organized as follows. In Sect. 2, we recall some fundamental definitions and results relating essentially to convex analysis that are needed later. In Sect. 3, under appropriate assumptions, we show that we have strong duality between (S) and its Fenchel–Lagrange dual problem. In Sect. 4, we provide optimality conditions for (S) and its dual. Finally, Sect. 5 is devoted to reducing the dual problem to a one-level convex minimization problem.

2 Preliminaries

In this section, we recall some fundamental definitions and results relating essentially to convex analysis that we will use in the sequel.

Let A be a nonempty subset of $\mathbb{R}^n$. We shall denote by ψ_A and σ_A the indicator and the support functions of the set A, respectively, defined on $\mathbb{R}^n$ by

$$\psi_A(x) := \begin{cases} 0, & \text{if } x \in A, \\ +\infty, & \text{if } x \notin A, \end{cases} \quad \text{and} \quad \sigma_A(d) := \sup_{x \in A} \langle d, x \rangle,$$

where $\langle \cdot, \cdot \rangle$, denotes the inner product of two vectors in $\mathbb{R}^n$, i.e., for $x = (x_1, \ldots, x_n)^T$, and $y = (y_1, \ldots, y_n)^T$, $\langle x, y \rangle = \sum_{i=1}^n x_i y_i$. The relative interior of A, denoted by $\mathrm{ri}(A)$, is the interior of A relative to the smallest affine set containing A, equipped with the induced topology of $\mathbb{R}^n$.

Definition 2.1 [16] Let $g : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a function, where $\overline{\mathbb{R}} = \mathbb{R} \cup \{\pm\infty\}$, and let A be a nonempty subset of $\mathbb{R}^n$. The conjugate function of g relative to the set A is denoted by g_A^* and defined on $\mathbb{R}^n$ by

$$g_A^*(p) := \sup_{x \in A} \big\{ \langle p, x \rangle - g(x) \big\}.$$

When $A = \mathbb{R}^n$, we get the usual Legendre–Fenchel conjugate function of g, usually denoted by g^*.

Definition 2.2 Let $g : \mathbb{R}^n \to \overline{\mathbb{R}}$ be a function. The function g is said to be $\mathbb{R}_+^n$-increasing iff for any $x, y \in \mathbb{R}^n$, $x = (x_1, \ldots, x_n)^T$, $y = (y_1, \ldots, y_n)^T$, with $x_i \leq y_i$, $\forall i \in \{1, \ldots, n\}$, we have $g(x) \leq g(y)$, where $\mathbb{R}_+^n$ denotes the non-negative orthant of $\mathbb{R}^n$, and the superscript ^{T} stands for transpose.

Definition 2.3 Let $f_1, \ldots, f_m : \mathbb{R}^n \to \overline{\mathbb{R}}$ be functions. The infimal convolution of $f_1, \ldots, f_m$ is the function denoted by $f_1 \square \cdots \square f_m$, and defined on $\mathbb{R}^n$ by

$$(f_1 \square \cdots \square f_m)(x) := \inf_{\substack{x_1, \ldots, x_m \in \mathbb{R}^n \\ x_1 + \cdots + x_m = x}} \left\{ \sum_{i=1}^m f_i(x_i) \right\}.$$

For an extended real-valued function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$, we shall denote by $\mathrm{dom}\, f$ the effective domain of f, i.e., the set defined by

$$\mathrm{dom}\, f := \big\{ x \in \mathbb{R}^n / f(x) < +\infty \big\}.$$

We say that f is proper if $f(x) > -\infty$ for all $x \in \mathbb{R}^n$ and $\mathrm{dom}\, f \neq \emptyset$.

Definition 2.4 Let $f : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ be a proper convex function. Let $\bar{x} \in \mathrm{dom}\, f$. The subdifferential of f at $\bar{x}$, denoted by $\partial f(\bar{x})$, is the set defined by

$$\partial f(\bar{x}) := \big\{ x^* \in \mathbb{R}^n / f(x) \geq f(\bar{x}) + \langle x^*, x - \bar{x} \rangle, \forall x \in \mathbb{R}^n \big\},$$

and any $x^* \in \partial f(\bar{x})$ is called a subgradient of f at $\bar{x}$.

Definition 2.5 Let A be a nonempty subset of $\mathbb{R}^n$, and $\bar{x} \in A$. The normal cone to A at $\bar{x}$ denoted by $\mathcal{N}_A(\bar{x})$ is the set defined by

$$\mathcal{N}_A(\bar{x}) := \left\{ x^* \in \mathbb{R}^n / \langle x^*, x - \bar{x} \rangle \leq 0, \forall x \in A \right\}.$$

If A is convex, then $\partial \psi_A(\bar{x}) = \mathcal{N}_A(\bar{x})$ (e.g., see [19]).

Theorem 2.1 [19] *Let* $f_1, \ldots, f_m : \mathbb{R}^n \to \overline{\mathbb{R}}$ *be proper convex functions. Assume that* $\bigcap_{i=1}^{m} \mathrm{ri}(\mathrm{dom}\, f_i) \neq \emptyset$. *Then, for any* $x \in \mathbb{R}^n$, *we have*

$$\left(\sum_{i=1}^{m} f_i \right)^* (x) = \left(f_1^* \square \cdots \square f_m^* \right)(x) = \inf_{\substack{x_1, \ldots, x_m \in \mathbb{R}^n \\ x_1 + \cdots + x_m = x}} \left\{ \sum_{i=1}^{m} f_i^*(x_i) \right\},$$

and the infimum is attained.

3 The Fenchel–Lagrange Duality Approach

In this section, we give a dual problem of (S) involving conjugate functions, called the Fenchel–Lagrange dual of (S). Then, under appropriate assumptions, we show that a strong duality holds between them. This interesting duality has been first introduced in [14] for ordinary convex programming problems and afterwards used in several works (e.g., see [16, 18, 20, 21]).

Let $h : \mathbb{R}^p \to \mathbb{R}^{p+1}$ be the function defined by

$$h(x) := (h_1(x), \ldots, h_p(x), h_{p+1}(x))^T = (x, v(x))^T,$$

i.e., for $x = (x_1, \ldots, x_p)^T$,

$$\begin{cases} h_i(x) := x_i, & \text{for } i = 1, \ldots, p, \\ h_i(x) := v(x), & \text{for } i = p+1, \end{cases}$$

with $v(x)$ the optimal value of the lower level problem $\mathcal{P}(x)$. Then, the problem (S) can be written as

$$(S) \qquad \underset{x \in X}{\mathrm{Min}}\, F(h(x)),$$

which is a composed programming problem. Note that the interesting and wide class of composed programming problems was extensively studied in the last decade. For papers dealing with such problems, we refer, for example, to [16, 18] and references therein. For our approach, the problem (S) will be seen as a composed programming problem.

In the sequel, in order to give the dual of the bilevel programming problem (S), we adopt the procedure given by Boţ et al. in [16] for composed programming problems. Then, we start our duality approach by considering the following minimization problem

$$(S') \qquad \underset{\substack{x \in X, y \in \mathbb{R}^{p+1} \\ h(x) \leq y}}{\mathrm{Min}}\, F(y).$$

We will make the following assumptions.

Assumption 3.1 The function F is $\mathbb{R}_+^{p+1}$-increasing.

Assumption 3.2 The function F is convex on $\mathbb{R}^{p+1}$.

Assumption 3.3 The function f is convex on $\mathbb{R}^p \times \mathbb{R}^q$.

Remark 3.1 Under the above assumptions, the remarks are quite obvious, especially in the particular case.

1. The Assumption 3.1 is extensively used in the study of composed programming problems. It allows us to establish a relation between the problems (S) and (S'), and to preserve the convexity of functions under composition.
2. From the convexity of the set Y and the function f, we deduce that the marginal function $v(\cdot)$ is convex on $\mathbb{R}^p$. It follows that the vector function h is also convex on $\mathbb{R}^p$. If, moreover, the Assumptions 3.1 and 3.2 are satisfied, then the problem (S) is convex.
3. Under the Assumption 3.3, we deduce that, for any $x \in \mathbb{R}^p$, $v(x)$ is a finite real number. This follows from the continuity of the function $f(x, \cdot)$ and the compactness of Y. Furthermore, $v(\cdot)$ is continuous on $\mathbb{R}^p$ as a finite convex function.

   Let $\inf(S)$ and $\inf(S')$ denote the infimal values of the problems (S) and (S'), respectively. Then, we have the following result.

Proposition 3.1 *Let the Assumptions* 3.1–3.3 *hold. Then,*

1. *The problem* (S) *has at least one solution,*
2. $\inf(S) = \inf(S')$.

Proof 1. First, note that F is continuous on $\mathbb{R}^{p+1}$ as a finite convex function. Then, the result follows from the continuity of the composed function $F \circ h$, and the compactness of the set X. It follows that $\inf(S)$ is a finite real number.

2. Let $x \in X$. Then, $(x, h(x))$ is a feasible point to (S'). Hence, $\inf(S') \le F(h(x))$, and
$$\inf(S') \le \inf_{x \in X} F(h(x)) = \inf(S).$$

Conversely, let (x, y) be a feasible point of (S'). Then, $x \in X$, i.e., x is feasible for (S). Since $h(x) \le y$, it follows from Assumption 3.1 that $F(h(x)) \le F(y)$. Hence $\inf(S) \le F(h(x)) \le F(y)$. Since (x, y) is an arbitrary feasible point of (S'), it follows that
$$\inf(S) \le \inf(S'). \qquad \square$$

Let $(\mathcal{D})$ be the Lagrangian dual problem to (S'), i.e.,

$$(\mathcal{D}) \qquad \underset{\substack{\alpha \ge 0 \\ \alpha \in \mathbb{R}^{p+1}}}{\text{Max}} \quad \underset{\substack{x \in X \\ y \in \mathbb{R}^{p+1}}}{\inf} \left\{ F(y) + \langle \alpha, h(x) - y \rangle \right\},$$

and consider the following maximization problem in which the objective function is expressed in terms of the conjugate functions of F, $\alpha^T h$ and ψ_X

$$(\mathcal{D}^*) \qquad \underset{\substack{\alpha \in \mathbb{R}_+^{p+1} \\ \beta \in \mathbb{R}^p}}{\mathrm{Max}} \left\{ -F^*(\alpha) - (\alpha^T h)^*(\beta) - \psi_X^*(-\beta) \right\},$$

where for $\alpha = (\alpha_1, \ldots, \alpha_{p+1})^T \in \mathbb{R}^{p+1}$, $\alpha^T h$ denotes the function defined on $\mathbb{R}^p$ and valued in $\mathbb{R}$, by

$$\alpha^T h(x) = \langle \alpha, h(x) \rangle = \sum_{i=1}^{p+1} \alpha_i h_i(x) = \sum_{i=1}^{p} \alpha_i x_i + \alpha_{p+1} v(x),$$

with $x = (x_1, \ldots, x_p)^T$. Using the fact that $\psi_X^* = \sigma_X$, the dual problem $(\mathcal{D}^*)$ can also be written as

$$(\mathcal{D}^*) \qquad \underset{\substack{\alpha \in \mathbb{R}_+^{p+1} \\ \beta \in \mathbb{R}^p}}{\mathrm{Max}} \left\{ -F^*(\alpha) - (\alpha^T h)^*(\beta) - \sigma_X(-\beta) \right\},$$

which is the formulation that we adopt in the sequel. In the literature, $(\mathcal{D}^*)$ is called the Fenchel–Lagrange dual problem to (S) (see [14, 18]).

Remark 3.2 By weak duality, we have $\inf(S') \geq \sup(\mathcal{D})$, where $\sup(\mathcal{D})$ denotes the supremum of the dual problem $(\mathcal{D})$.

The following proposition establishes that problems $(\mathcal{D})$ and $(\mathcal{D}^*)$ have the same supremum. The proof corresponding to an ordinary composed programming problem can be found, for example, in [16, 18]. For the convenience of the reader, we give the proof corresponding to our case.

Proposition 3.2 *Let the Assumptions* 3.2 *and* 3.3 *hold. Then,*

$$\sup(\mathcal{D}) = \sup(\mathcal{D}^*).$$

Proof We have

$$\underset{\substack{x \in X \\ y \in \mathbb{R}^{p+1}}}{\inf} \left\{ F(y) + \langle \alpha, h(x) - y \rangle \right\}$$

$$= \underset{y \in \mathbb{R}^{p+1}}{\inf} \left\{ F(y) - \langle \alpha, y \rangle \right\} + \underset{x \in X}{\inf} \langle \alpha, h(x) \rangle$$

$$= - \underset{y \in \mathbb{R}^{p+1}}{\sup} \left\{ \langle \alpha, y \rangle - F(y) \right\} - \underset{x \in X}{\sup} \left\{ \langle 0, x \rangle - \langle \alpha, h(x) \rangle \right\}$$

$$= -F^*(\alpha) - (\alpha^T h)_X^*(0).$$

On the other hand, since X is a nonempty convex set, we have $\mathrm{ri}(X) \neq \emptyset$ (see [19]). So that

$$\mathrm{ri}(\mathrm{dom}\, \alpha^T h) \cap \mathrm{ri}(\mathrm{dom}\, \psi_X) = \mathbb{R}^p \cap \mathrm{ri}(X) = \mathrm{ri}(X) \neq \emptyset.$$

Therefore, by Theorem 2.1, we have

$$\left(\alpha^T h\right)^*_X(0) = \left(\alpha^T h + \psi_X\right)^*(0) = \left(\alpha^T h\right)^* \square (\psi_X)^*(0)$$

$$= \inf_{\beta \in \mathbb{R}^p} \left\{ \left(\alpha^T h\right)^*(\beta) + (\psi_X)^*(-\beta) \right\}$$

$$= \inf_{\beta \in \mathbb{R}^p} \left\{ \left(\alpha^T h\right)^*(\beta) + \sigma_X(-\beta) \right\}.$$

Then,

$$\sup_{\alpha \in \mathbb{R}^{p+1}_+} \inf_{\substack{x \in X \\ y \in \mathbb{R}^{p+1}}} \left\{ F(y) + \langle \alpha, h(x) - y \rangle \right\}$$

$$= \sup_{\alpha \in \mathbb{R}^{p+1}_+} \left\{ -F^*(\alpha) - \inf_{\beta \in \mathbb{R}^p} \left\{ \left(\alpha^T h\right)^*(\beta) + \sigma_X(-\beta) \right\} \right\}$$

$$= \sup_{\substack{\alpha \in \mathbb{R}^{p+1}_+ \\ \beta \in \mathbb{R}^p}} \left\{ -F^*(\alpha) - \left(\alpha^T h\right)^*(\beta) - \sigma_X(-\beta) \right\}.$$

That is,

$$\sup(\mathcal{D}) = \sup(\mathcal{D}^*). \qquad \square$$

Corollary 3.1 *Let the Assumptions 3.1–3.3 hold. Then, the weak duality holds between (S) and (D*).*

Proof The result is deduced by using Remark 3.2, and Propositions 3.1 and 3.2.
Let $g : \mathbb{R}^p \times \mathbb{R}^{p+1} \to \mathbb{R}^{p+1}$ be the function defined by

$$g(x, y) := \left(g_1(x, y), \ldots, g_p(x, y), g_{p+1}(x, y) \right)^T = h(x) - y,$$

that is,

$$\begin{cases} g_i(x, y) = x_i - y_i, & \text{for } i = 1, \ldots, p, \\ g_i(x, y) = v(x) - y_i, & \text{for } i = p + 1, \end{cases}$$

with $x = (x_1, \ldots, x_p)^T$ and $y = (y_1, \ldots, y_p, y_{p+1})^T$. Assume that the marginal function $v(\cdot)$ is not affine, which is the case, in general. So, the function g_{p+1} is not affine. Let

$$\begin{cases} \mathcal{L} := \{ i \in \{1, \ldots, p + 1\} / g_i \text{ is an affine function} \}, \\ \mathcal{N} := \{1, \ldots, p + 1\} \setminus \mathcal{L}. \end{cases}$$

It follows that

$$\mathcal{L} = \{1, \ldots, p\} \quad \text{and} \quad \mathcal{N} = \{p + 1\}. \qquad \square$$

Remark 3.3 If $v(\cdot)$ is affine, then $\mathcal{L} = \{1, \ldots, p+1\}$ and $\mathcal{N} = \emptyset$.

Consider the following constraint qualifications

$(\overline{\text{CQ}})$ $\exists x = (x_1, \ldots, x_p)^T \in \text{ri}(X)$, $\exists y = (y_1, \ldots, y_p, y_{p+1})^T \in \mathbb{R}^{p+1}$, $\exists z \in Y$ such that

$$x_i \le y_i, i = 1, \ldots, p \quad \text{and} \quad f(x, z) < y_{p+1},$$

(CQ) $\exists x \in \text{ri } X$, $\exists y \in \mathbb{R}^{p+1}$ such that

$$\begin{cases} g_i(x, y) \le 0, & \text{for } i \in \mathcal{L}, \\ g_i(x, y) < 0, & \text{for } i \in \mathcal{N}. \end{cases}$$

Note that the constraint qualification (CQ) is well-known in the literature and is used to ensure strong Lagrange duality for an ordinary convex program with its Lagrange dual (see [19, 22]). On the other hand, it is not difficult to show that the constraint qualifications $(\overline{\text{CQ}})$ and (CQ) are equivalent.

Remark 3.4

1. Under the Assumptions 3.2 and 3.3 and the data of problem (S), the constraint qualification $(\overline{\text{CQ}})$ is always satisfied. This follows principally from the fact that for any $x \in X$, $v(x)$ is a finite real number (Remark 3.1). Indeed, since $\text{ri}(X)$ is nonempty, let $x = (x_1, \ldots, x_p)^T \in \text{ri}(X)$, and let z be a point in the compact set Y such that $v(x) = f(x, z)$. Then, any point $y = (y_1, \ldots, y_p, y_{p+1})^T \in \mathbb{R}^{p+1}$, verifying

$$x_i \le y_i, i = 1, \ldots, p \quad \text{and} \quad f(x, z) < y_{p+1},$$

   is suitable. That is, the constraint qualification $(\overline{\text{CQ}})$ is fulfilled.
2. Let the Assumptions 3.1–3.3 hold. From Proposition 3.1, we have $\inf(S) = \inf(S')$. Since $\inf(S) \in \mathbb{R}$, it follows that $\inf(S') > -\infty$.

The following theorem establishes that a strong duality between (S) and $(\mathcal{D}^*)$ holds.

Theorem 3.1 *Let the Assumptions* 3.1–3.3 *hold. Then, a strong duality holds between* (S) *and* $(\mathcal{D}^*)$, *i.e.,*

(i) $\inf(S) = \sup(\mathcal{D}^*)$;
(ii) *The dual* $(\mathcal{D}^*)$ *admits solutions.*

Proof (i) As quoted in the above remark, the constraint $(\overline{\text{CQ}})$ is satisfied. Hence, since $(\overline{\text{CQ}})$ is equivalent to (CQ) and $\inf(S') > -\infty$ [see 2. of Remark 3.4], it follows that a strong duality holds between (S') and $(\mathcal{D})$ (e.g., see [19, 22]). Therefore,

$$\inf(S) = \inf(S') = \sup(\mathcal{D}) = \sup(\mathcal{D}^*),$$

where the first and the last inequalities follow from Proposition 3.1 and Proposition 3.2, respectively.

(ii) See, for example, [16, 18]. $\qquad \square$

4 Optimality Conditions for (S) and (𝒟*)

In this section, under the Assumptions 3.1–3.3 and exploiting our duality, we provide optimality conditions for (S) and its dual (𝒟*). These optimality conditions are expressed in terms of conjugate functions and subdifferentials, respectively. Since the problems (S) and (𝒟*) are convex, we are only concerned with global solutions. On the other hand, recall that the problem (S) admits solutions (Proposition 3.1), and that (S) and (𝒟*) are in strong duality (Theorem 3.1).

The following theorems give respectively necessary and sufficient optimality conditions for (S) and its dual. The proofs can be found in [15]. Similar proofs are given in [17] for the case of the Fenchel–Rockafellar duality.

Theorem 4.1 (Necessary optimality conditions) *Let the Assumptions* 3.1–3.3 *hold. Let* $\bar{x} \in X$ *and* $(\bar{\alpha}, \bar{\beta}) \in \mathbb{R}_+^{p+1} \times \mathbb{R}^p$ *be solutions to* (S) *and* (𝒟*)*, respectively. Then,* $\bar{x}$ *and* $(\bar{\alpha}, \bar{\beta})$ *satisfy the following necessary optimality conditions*

$$F(h(\bar{x})) + F^*(\bar{\alpha}) = \langle \bar{\alpha}, h(\bar{x}) \rangle, \tag{1}$$

$$\langle \bar{\alpha}, h(\bar{x}) \rangle + \left(\bar{\alpha}^T h\right)^*(\bar{\beta}) = \langle \bar{\beta}, \bar{x} \rangle, \tag{2}$$

$$\sigma_X(-\bar{\beta}) + \langle \bar{\beta}, \bar{x} \rangle = 0. \tag{3}$$

Theorem 4.2 (Sufficient optimality conditions) *Let the Assumptions* 3.1–3.3 *hold. Let* $\bar{x}$ *and* $(\bar{\alpha}, \bar{\beta})$ *be feasible points to* (S) *and* (𝒟*)*, respectively. Assume that they satisfy the conditions* (1)–(3) *in Theorem* 4.1. *Then,* $\bar{x}$ *solves* (S) *and* $(\bar{\alpha}, \bar{\beta})$ *solves* (𝒟*)*.*

The following corollaries give the main results of this section. They provide necessary and sufficient optimality conditions for (S) and (𝒟*).

Corollary 4.1 *Let the Assumptions* 3.1–3.3 *hold. Let* $\bar{x}$ *and* $(\bar{\alpha}, \bar{\beta})$ *be feasible points to* (S) *and* (𝒟*)*, respectively. Then,* $\bar{x}$ *and* $(\bar{\alpha}, \bar{\beta})$ *solve* (S) *and* (𝒟*)*, respectively, if and only if the conditions* (1)–(3) *are satisfied.*

Proof The result follows by using Theorems 4.1 and 4.2. □

In terms of the subdifferentials, we have the following result.

Corollary 4.2 *Let the Assumptions* 3.1–3.3 *hold. Let* $\bar{x}$ *and* $(\bar{\alpha}, \bar{\beta})$ *be feasible points to* (S) *and* (𝒟*)*, respectively. Then,* $\bar{x}$ *and* $(\bar{\alpha}, \bar{\beta})$ *solve* (S) *and* (𝒟*)*, respectively, if and only if the following conditions are satisfied*

$$\bar{\alpha} \in \partial F(h(\bar{x})), \tag{1'}$$

$$\bar{\beta} \in \partial(\bar{\alpha}^T h)(\bar{x}), \tag{2'}$$

$$-\bar{\beta} \in \partial \psi_X(\bar{x}) = \mathcal{N}_X(\bar{x}). \tag{3'}$$

Proof The conditions (1) and (2) are obviously equivalent to (1′) and (2′), respectively. Furthermore, remarking that $\psi_X(\bar{x}) = 0$ (since $\bar{x} \in X$), and using the fact that $\sigma_X = \psi_X^*$, the condition (3) can be written as

$$\psi_X(\bar{x}) + \psi_X^*(-\bar{\beta}) = \langle -\bar{\beta}, \bar{x} \rangle,$$

which is equivalent to

$$-\bar{\beta} \in \partial \psi_X(\bar{x}) = \mathcal{N}_X(\bar{x}). \qquad \square$$

5 Reduction of ($\mathcal{D}^*$) to a One-level Convex Minimization Problem

In this section, exploiting the particular form of the dual problem ($\mathcal{D}^*$), and via the conjugacy operation, we will show that the resolution of the dual problem is equivalent to the resolution of two single-level (one-level) convex minimization problems.

First, let us give the conjugate of the marginal function $v(\cdot)$ which will be needed in the sequel.

Proposition 5.1 *Let the Assumption* 3.3 *hold. Let* $x^* \in \mathbb{R}^p$. *Then,*

$$v^*(x^*) = f_{\mathbb{R}^p \times Y}^*(x^*, 0).$$

Proof First, recall that for any $x \in \mathbb{R}^p$, $v(x)$ is a finite real number (Remark 3.1). Let $x^* \in \mathbb{R}^p$. We have

$$
\begin{aligned}
v^*(x^*) &= \sup_{x \in \mathbb{R}^p} \left\{ \langle x^*, x \rangle - v(x) \right\} \\
&= \sup_{x \in \mathbb{R}^p} \left\{ \langle x^*, x \rangle - \inf_{y \in Y} f(x, y) \right\} \\
&= \sup_{(x, y) \in \mathbb{R}^p \times Y} \left\{ \left\langle \begin{pmatrix} x^* \\ 0 \end{pmatrix}, \begin{pmatrix} x \\ y \end{pmatrix} \right\rangle - f(x, y) \right\} = f_{\mathbb{R}^p \times Y}^*(x^*, 0). \qquad \square
\end{aligned}
$$

As remarked, the dual problem ($\mathcal{D}^*$) has an objective function which contains the conjugate function $(\alpha^T h)^*$. However, the function $\alpha^T h$ depends on the marginal function $v(\cdot)$ which is an output of the lower level problem. So, in the following, we try to avoid such a dependence. This allows us to formulate the dual problem ($\mathcal{D}^*$) only in terms of the data of the problem (S). Then, we have the following result.

Proposition 5.2 *Let the Assumptions* 3.1–3.3 *hold. Then, the resolution of the dual* ($\mathcal{D}^*$) *is equivalent to the resolution of the following one-level maximization problems*

$$(\mathcal{D}_1^*) \qquad \operatorname*{Max}_{\substack{\alpha_1 \geq 0, \ldots, \alpha_p \geq 0 \\ \beta = (\alpha_1, \ldots, \alpha_p)^T}} \left\{ - F^*(\alpha_1, \ldots, \alpha_p, 0) - \sigma_X(-\beta) \right\}$$

and

$$(\mathcal{D}_2^*) \quad \underset{\substack{\alpha=(\alpha_1,\ldots,\alpha_p,\alpha_{p+1})^T \\ \alpha_1\geq0,\ldots,\alpha_p\geq0 \\ \alpha_{p+1}>0 \\ \beta\in\mathbb{R}^p}}{\text{Max}} \left\{ -F^*(\alpha) - \alpha_{p+1} f^*_{\mathbb{R}^p\times Y}\left(\frac{\beta-(\alpha_1,\ldots,\alpha_p)}{\alpha_{p+1}}, 0\right) - \sigma_X(-\beta)\right\}$$

in the sense that

- $\sup(\mathcal{D}^*) = \max\{\sup(\mathcal{D}_1^*), \sup(\mathcal{D}_2^*)\}$, *and*
- (α, β) *solves* $(\mathcal{D}^*)$ *if and only if* (α, β) *solves* $(\mathcal{D}_i^*)$, *with* $\sup(\mathcal{D}_i^*) \geq \sup(\mathcal{D}_j^*)$, $i \neq j, i, j \in \{1, 2\}$.

Proof Replacing $h(x)$ by its expression $(x, v(x))$ in the formulation of the dual problem $(\mathcal{D}^*)$, we get

$$(\mathcal{D}^*) \quad \underset{\substack{\alpha=(\alpha_1,\ldots,\alpha_p,\alpha_{p+1})^T \\ \alpha_1\geq0,\ldots,\alpha_p\geq0 \\ \alpha_{p+1}\geq0 \\ \beta\in\mathbb{R}^p}}{\text{Max}} \left\{ -F^*(\alpha) - \left(\sum_{i=1}^p \alpha_i h_i + \alpha_{p+1} v(\cdot)\right)^*(\beta) - \sigma_X(-\beta)\right\}.$$

So that

$$\sup(\mathcal{D}^*) = \sup\left\{ \underset{\substack{\alpha_1\geq0,\ldots,\alpha_p\geq0 \\ \beta\in\mathbb{R}^p}}{\sup} \left\{ -F^*(\alpha_1,\ldots,\alpha_p,0) - \underbrace{\left(\sum_{i=1}^p \alpha_i h_i\right)^*(\beta)}_{\mathcal{A}_{(\alpha_1,\ldots,\alpha_p)}(\beta)} - \sigma_X(-\beta)\right\}, \right.$$

$$\left. \underset{\substack{\alpha=(\alpha_1,\ldots,\alpha_p,\alpha_{p+1})^T \\ \alpha_1\geq0,\ldots,\alpha_p\geq0 \\ \alpha_{p+1}>0 \\ \beta\in\mathbb{R}^p}}{\sup} \left\{ -F^*(\alpha) - \left(\sum_{i=1}^p \alpha_i h_i + \alpha_{p+1} v(\cdot)\right)^*(\beta) - \sigma_X(-\beta)\right\}\right\}.$$

Let us calculate the value of $\mathcal{A}_{(\alpha_1,\ldots,\alpha_p)}(\beta)$. We have

$$\begin{aligned}
\mathcal{A}_{(\alpha_1,\ldots,\alpha_p)}(\beta) &= \underset{t\in\mathbb{R}^p}{\sup} \left\{\langle\beta,t\rangle - (\alpha_1 h_1 + \cdots + \alpha_p h_p)(t)\right\} \\
&= \underset{t\in\mathbb{R}^p}{\sup} \left\{\langle\beta,t\rangle - (\alpha_1 t_1 + \cdots + \alpha_p t_p)\right\} \\
&= \underset{t\in\mathbb{R}^p}{\sup} \left\langle\beta - (\alpha_1,\ldots,\alpha_p)^T, t\right\rangle \\
&= \begin{cases} +\infty, & \text{if } \beta \neq (\alpha_1,\ldots,\alpha_p), \\ 0, & \text{if } \beta = (\alpha_1,\ldots,\alpha_p), \end{cases}
\end{aligned}$$

with $t = (t_1, \ldots, t_p)^T$. So that

$$\sup(\mathcal{D}^*) = \sup \left\{ \sup_{\substack{\alpha_1 \geq 0, \ldots, \alpha_p \geq 0 \\ \beta = (\alpha_1, \ldots, \alpha_p)^T}} \left\{ -F^*(\alpha_1, \ldots, \alpha_p, 0) - \sigma_X(-\beta) \right\}, \right.$$

$$\left. \sup_{\substack{\alpha = (\alpha_1, \ldots, \alpha_{p+1})^T \\ \alpha_1 \geq 0, \ldots, \alpha_p \geq 0 \\ \alpha_{p+1} > 0 \\ \beta \in \mathbb{R}^p}} \left\{ -F^*(\alpha) - \left(\sum_{i=1}^p \alpha_i h_i + \alpha_{p+1} v(\cdot) \right)^*(\beta) - \sigma_X(-\beta) \right\} \right\}.$$

On the other hand, for $\alpha_{p+1} > 0$, we have

$$\left(\sum_{i=1}^p \alpha_i h_i + \alpha_{p+1} v(\cdot) \right)^*(\beta) = \sup_{t \in \mathbb{R}^p} \left\{ \langle \beta, t \rangle - \alpha_1 t_1 - \cdots - \alpha_p t_p - \alpha_{p+1} v(t) \right\}$$

$$= \alpha_{p+1} \sup_{t \in \mathbb{R}^p} \left\{ \left\langle \frac{\beta - (\alpha_1, \ldots, \alpha_p)^T}{\alpha_{p+1}}, t \right\rangle - v(t) \right\}$$

$$= \alpha_{p+1} v^* \left(\frac{\beta - (\alpha_1, \ldots, \alpha_p)^T}{\alpha_{p+1}} \right)$$

$$= \alpha_{p+1} f_{\mathbb{R}^p \times Y}^* \left(\frac{\beta - (\alpha_1, \ldots, \alpha_p)^T}{\alpha_{p+1}}, 0 \right),$$

where the last equality follows from Proposition 5.1. Now, consider the following maximization problems

$$(\mathcal{D}_1^*) \quad \max_{\substack{\alpha_1 \geq 0, \ldots, \alpha_p \geq 0 \\ \beta = (\alpha_1, \ldots, \alpha_p)^T}} \left\{ -F^*(\alpha_1, \ldots, \alpha_p, 0) - \sigma_X(-\beta) \right\}$$

and

$$(\mathcal{D}_2^*) \quad \max_{\substack{\alpha = (\alpha_1, \ldots, \alpha_p, \alpha_{p+1})^T \\ \alpha_1 \geq 0, \ldots, \alpha_p \geq 0 \\ \alpha_{p+1} > 0 \\ \beta \in \mathbb{R}^p}} \left\{ -F^*(\alpha) - \alpha_{p+1} f_{\mathbb{R}^p \times Y}^* \left(\frac{\beta - (\alpha_1, \ldots, \alpha_p)}{\alpha_{p+1}}, 0 \right) - \sigma_X(-\beta) \right\}.$$

Then, we have

$$\sup(\mathcal{D}^*) = \max \left\{ \sup(\mathcal{D}_1^*), \sup(\mathcal{D}_2^*) \right\}.$$

It follows that (α, β) is a solution of $(\mathcal{D}^*)$ if and only if (α, β) is a solution of $(\mathcal{D}_i^*)$ with $\sup(\mathcal{D}_i^*) \geq \sup(\mathcal{D}_j^*)$, $i \neq j$, $i, j \in \{1, 2\}$. $\qquad \square$

Remark 5.1

1. Remark that the objective functions of problems $(\mathcal{D}_1^*)$ and $(\mathcal{D}_2^*)$ are expressed in terms of the conjugate functions of F, f and ψ_X (since $\sigma_X = \psi_X^*$).
2. As it is well-known, the functions F^*, f^* and σ_X are convex. Besides, the constraint sets of problems $(\mathcal{D}_1^*)$ and $(\mathcal{D}_2^*)$ are also convex. It follows that $(\mathcal{D}_1^*)$ and $(\mathcal{D}_2^*)$ are concave maximization problems (equivalent to convex minimization problems).

Consequently, the resolution of the problem $(\mathcal{D}^*)$ is equivalent to the resolution of the two single-level concave maximization problems $(\mathcal{D}_1^*)$ and $(\mathcal{D}_2^*)$. Therefore, the problem $(\mathcal{D}^*)$ is reduced to a single-level convex minimization problem.

For illustration, we give the following simple example.

Example 5.1 Let $X = Y = [0, 1]$, f and F be the functions defined on $\mathbb{R}^2$ by

$$f(x, y) = \frac{1}{2}x^2 + y + \frac{1}{4}y^4 \quad \text{and} \quad F(x, t) = \begin{cases} \frac{1}{2}x^2 + \frac{1}{2}t^2, & \text{if } x \geq 0, t \geq 0, \\ 0, & \text{otherwise.} \end{cases}$$

Then, the Assumptions 3.1–3.3 are satisfied. For $x \in X$, we have

$$\inf \mathcal{P}(x) = v(x) = \frac{1}{2}x^2.$$

The bilevel programming problem that we consider is

$$(S) \qquad \underset{x \in X}{\text{Min}}\, F\big(x, v(x)\big) = \frac{1}{2}x^2 + \frac{1}{8}x^4$$

whose unique solution is obviously $\bar{x} = 0$ and $\inf(S) = 0$. Let us verify this by using our duality approach. Note that, in this example, we have strong duality between (S) and its Fenchel–Lagrange dual problem $(\mathcal{D}^*)$ (Theorem 3.1). Let us calculate $f_{\mathbb{R} \times Y}^*$ and F^*. By simple calculation, we find

$$f^*(x^*, y^*) = \begin{cases} \frac{1}{2}x^{*2} + y^* - \frac{5}{4}, & \text{if } y^* > 2, x^* \in \mathbb{R}, \\ \frac{1}{2}x^{*2} + \frac{5}{4}(y^* - 1)^{\frac{4}{3}}, & \text{if } 1 \leq y^* \leq 2, x^* \in \mathbb{R}, \\ \frac{1}{2}x^{*2}, & \text{if } y^* < 1, x^* \in \mathbb{R}, \end{cases}$$

and

$$F^*(x^*, t^*) = \begin{cases} +\infty, & \text{if } (x^*, t^*) \neq (0, 0), \\ 0, & \text{if } (x^*, t^*) = (0, 0). \end{cases}$$

On the other hand, for problems $(\mathcal{D}_i^*)$, $i = 1, 2$, we have

$$\sup(\mathcal{D}_2^*) = -\infty \quad \text{and} \quad \sup(\mathcal{D}_1^*) = 0,$$

where $\sup(\mathcal{D}_1^*)$ is attained at $\left(\frac{(\bar{\alpha}_1, \bar{\alpha}_2)^T}{\bar{\beta}}\right)$. So that $\sup(\mathcal{D}^*) = \sup(\mathcal{D}_1^*) = 0$, attained at $\left(\frac{(\bar{\alpha}_1, \bar{\alpha}_2)^T}{\bar{\beta}}\right) = \left(\frac{(0,0)^T}{0}\right)$. By strong duality, we have $\sup(\mathcal{D}^*) = \inf(S) = 0$.

Let us calculate the solution of (S) by using the conditions in Corollary 4.1. We have $h(x) = (x, \frac{1}{2}x^2)^T$. Then, $\bar{x} \in X$ is a solution of (S) if and only if

$$F(h(\bar{x})) + F^*(\bar{\alpha}) = \langle \bar{\alpha}, h(\bar{x}) \rangle, \tag{4}$$

$$\langle \bar{\alpha}, h(\bar{x}) \rangle + (\bar{\alpha}^T h)^*(\bar{\beta}) = \langle \bar{\beta}, \bar{x} \rangle, \tag{5}$$

$$\sigma_{[0,1]}(-\bar{\beta}) + \langle \bar{\beta}, \bar{x} \rangle = 0, \tag{6}$$

with $\bar{\alpha} = (\bar{\alpha}_1, \bar{\alpha}_2)^T$. So that, from the first equation, we have

$$\bar{x}^2 \left(1 + \frac{1}{4}\bar{x}^2 \right) = 0,$$

whose unique solution is $\bar{x} = 0$. Besides, $\bar{x}$ satisfies with $\bar{\alpha}$ and $\bar{\beta}$ (5) and (6). Hence, $\bar{x}$ is the unique solution of (S).

Remark 5.2

1. The same duality approach can be applied via appropriate modifications to the case where the constraints of the leader depend on the optimal value of the follower, i.e., the case

$$(S) \qquad \underset{\substack{x \in X \\ G(x,v(x)) \leq 0}}{\text{Min}} \quad F(x, v(x)),$$

   where $G : \mathbb{R}^p \times \mathbb{R} \to \mathbb{R}$.
2. Viewing (S) as a composed programming problem, other duality approaches can be applied to (S), as Lagrange duality and Fenchel duality. Relations between these dualities and the Fenchel–Lagrange duality are given in [18].

6 Conclusions

The bilevel programming problem (S) which, in general, is not differentiable is difficult to solve in view of the existence of the marginal function $v(\cdot)$ in its formulation. For such a problem, we have considered a duality approach by introducing its Fenchel–Lagrange dual problem $(\mathcal{D}^*)$. This type of dual problem has been first introduced by Boţ and Wanka in [14] for ordinary convex programming problems. Under appropriate assumptions, and based on some results given in the literature for composed programming problems, we have shown that a strong duality holds between (S) and $(\mathcal{D}^*)$. Then, for problem (S) and its dual we have provided optimality conditions which are expressed in terms of conjugate functions and subdifferentials, respectively. Finally, in Proposition 5.2, we have reduced the dual problem to a one-level convex minimization problem. Therefore, this duality approach gives the possibility to solve the bilevel programming problem (S) via a one-level convex minimization programming problem.

References

1. Shimizu, K., Ishizuka, Y., Bard, J.F.: Nondifferentiable and Two-level Mathematical Programming. Kluwer Academic, Boston (1997)
2. Anandalingam, G., White, D.J.: A solution method for the linear static Stackelberg problem using penalty functions. IEEE Trans. Autom. Control **35**, 1170–1173 (1990)
3. Bard, J.F.: Practical Bilevel Optimization: Algorithms and Applications. Kluwer Academic, Dordrecht (1998)
4. Dempe, S.: A simple algorithm for the linear bilevel programming problem. Optimization **183**, 73–385 (1987)
5. Dempe, S., Dutta, J., Mordukhovich, B.: New necessary optimality conditions in optimistic bilevel programming. Optimization **56**, 577–604 (2007)
6. Dinh, N., Mordukhovich, B., Nghia, T.T.A.: Subdifferentials of value functions and optimality conditions for DC and bilevel infinite and semi-infinite programs. Math. Program. **123**, 101–138 (2010)
7. Aboussoror, A., Loridan, P.: Existence and approximation results involving regularized constrained Stackelberg problems. J. Math. Anal. Appl. **188**, 101–117 (1994)
8. Aboussoror, A., Mansouri, A.: Weak linear bilevel programming problems: existence of solutions via a penalty method. J. Math. Anal. Appl. **304**, 399–408 (2005)
9. Lucchetti, R., Mignanego, F., Pieri, G.: Existence theorems of equilibrium points in Stackelberg games with constraints. Optimization **18**, 857–866 (1987)
10. Aboussoror, A., Babahadda, H., Mansouri, A.: Bilevel programs with extremal value function: global optimality conditions. Int. J. Math. Math. Sci. 419–435 (2005)
11. Ishizuka, Y.: Optimality conditions for quasi-differentiable programs with application to two-level optimization. SIAM J. Control Optim. **26**, 1388–1398 (1988)
12. Shimizu, K., Ishizuka, Y.: Optimality conditions and algorithms for parameter design problems with two-level structure. IEEE Trans. Autom. Control **AC-30**, 986–993 (1985)
13. Tanino, T., Ogawa, T.: An algorithm for solving two-level convex optimization problems. Int. J. Syst. Sci. **15**, 163–174 (1984)
14. Wanka, G., Boţ, R.I.: On the relations between different dual problems in convex mathematical programming. In: Chamoni, P., Leisten, R., Martin, A., Minnermann, J., Stadtler, H. (eds.) Operations Research Proceedings 2001, pp. 255–262. Springer, Berlin (2002)
15. Wanka, G., Boţ, R.I., Vargyas, E.: Conjugate duality for multiobjective composed optimization problems. Acta Math. Hung. **116**, 177–196 (2007)
16. Boţ, R.I., Hodrea, I.B., Wanka, G.: Farkas-type results for inequality systems with composed convex functions via conjugate duality. J. Math. Anal. Appl. **322**, 316–328 (2006)
17. Boţ, R.I., Wanka, G.: Duality for composed convex functions with applications in location theory. In: Habenicht, W., Scheubrein, B., Scheubrein, R. (eds.) Multi-Criteria und Fuzzy-Systeme in Theorie und Praxis, pp. 1–18. Deutscher Univestitäts-Verlag, Wiesbaden (2003)
18. Wanka, G., Boţ, R.I., Vargyas, E.: On the relations between different duals assigned to composed optimization problems. Math. Methods Oper. Res. **66**, 47–68 (2007)
19. Rockafellar, R.T.: Convex Analysis. Princeton University Press, Princeton (1970)
20. Boţ, R.I.: Conjugate duality in convex optimization. In: Lecture Notes in Economics and Mathematical Systems, vol. 637. Springer, Berlin, Heidelberg (2010)
21. Boţ, R.I., Grad, S.M., Wanka, G.: Duality in Vector Optimization. Springer, Berlin, Heidelberg (2009)
22. Elster, K.-H., Reinhardt, R., Schäuble, M., Donath, G.: Einführung in die Nichtlineare Optimierung. Teubner, Leipzig (1977)