Yokohama Publishers
ISSN 1349-8169 ONLINE JOURNAL

Pacific Journal of Optimization
Vol. 2, No. 3, pp. 447 – 467, September, 2006

OPTIMALITY CONDITIONS FOR THE SEMIVECTORIAL BILEVEL OPTIMIZATION PROBLEM

Henri Bonnel

Abstract: The paper studies a bilevel optimization problem with a lower level given by a convex vector minimization problem. In this case the lower level usually has multiple solutions, thus the so called optimistic formulation is considered. Necessary optimality conditions are obtained when dealing with properly efficient solutions as well as weakly efficient solutions for the lower level. An application to the case when the lower level is given by a linear multiobjective optimization problem (with data having a C^1-dependence on parameters) is given.

Key words: *bilevel optimization, multiobjective optimization, parameterized linear programming, optimization over the efficient set*

Mathematics Subject Classification: *90C29, 90C30, 90C48, 90C31, 90C05*

1 Introduction

Bilevel programming problems are very important for economic planning, network design and other applications. There are many research papers dealing with theoretical and/or practical approaches of these problems. These useful strategic economic problems can be considered as a static noncooperative asymmetric games (see e.g. [7]). Two players seek to optimize their individual objective function. The first player (the leader) must take into account any possible reaction of the second player (the follower).

When there are several reactions of the follower, but the follower is assumed to choose in favor of the leader we have the so called "optimistic case". This kind of program is usually a quite difficult nonconvex optimization problem. In some works, in order to simplify the model, stronger assumptions are made such as uniqueness of the solution of the lower level, or, if there are several solutions, they provide the same (unique) upper level objective value. In this case, using equivalent first order optimality condition for the lower level, the problem becomes a mathematical program with equilibrium constraints.

The more realistic and more difficult situation, the so called "pessimistic case", happens when the follower has different possible reactions and chose one which is the most unfavorable for the leader.

The main characteristic feature of this paper is that the objective of the follower is a vector valued function. More precisely, the follower solves a convex multiobjective problem looking for weakly or properly efficient solutions. Recall that the Pareto set contains the properly efficient set and is contained in the weakly efficient set. In the particular case of a linear vector optimization problem properly efficient and Pareto solutions coincide.

This paper deals mainly with the optimistic case, and sketches one issue for the pessimistic case, case which will be studied in a subsequent paper.

Therefore, the difficulty of this problem is due to two facts : the (vector) minimizers of the lower level are not unique, and a vector objective in the lower level (instead of a standard case with a scalar objective). This problem, generalizes not only usual (scalar) bilevel programms, but also the difficult vector optimization problem of optimizing a scalar function over the efficient (Pareto) set, as it is shown later in this section.

The paper presents optimality conditions for the following bilevel optimization problem with a convex vector (multiobjective) lower level optimization problem

$$(\mathrm{BL}_\sigma)\qquad \begin{cases} \min\limits_{x,\,y}\text{imize } f(x,y) \qquad \text{subject to :} \\ (x,y)\in U\subseteq X\times Y \qquad \text{(upper level constraints)} \\ y\in\sigma\text{-}\underset{y'}{\mathrm{ARGMIN}}_C\{F(x,y')|\, y'\in S(x)\} \qquad \text{(lower level } \sigma\text{-efficiency)} \end{cases}$$

where X, Y and Z are three real Banach spaces, $C\subset Z$ is a pointed convex cone, i.e. $\mathbb{R}_+C+C\subseteq C$ and $C\cap(-C)=\{0\}$, closed, with topological interior $\operatorname{int} C\neq\emptyset$, f is a real-valued map on $X\times Y$ (the upper level objective), U is a subset of $X\times Y$ (the upper level feasible set), F (the lower level *vector* objective) is a map from $X\times Y$ to $Z\cup\{+\infty_C\}$ (see Section 2 for details) such that, for each $x\in X$, the map $y\mapsto F(x,y)$ is C-convex proper, and $S : X\rightrightarrows Y$ is a set-valued map with convex values. The symbol $\sigma\in\{\mathrm{w},\mathrm{e},\mathrm{p}\}$, and, for each $x\in X$, $S(x)$ stands for the lower level feasible set, and (see Section 2 for details) :

- $\mathrm{w}\text{-}\underset{y'}{\mathrm{ARGMIN}}_C\{F(x,y')|\, y'\in S(x)\}$ is the set of the weakly-efficient solutions ($\sigma=$w);
- $\mathrm{e}\text{-}\underset{y'}{\mathrm{ARGMIN}}_C\{F(x,y')|\, y'\in S(x)\}$ is the set of the efficient solutions ($\sigma=$e);
- $\mathrm{p}\text{-}\underset{y'}{\mathrm{ARGMIN}}_C\{F(x,y')|\, y'\in S(x)\}$ is the set of the properly-efficient solutions ($\sigma=$p)

associated to the convex lower level vector optimization problem :

$$(\mathrm{LL})(x)\qquad \underset{y'}{\mathrm{MINIMIZE}}_C F(x,y') \qquad \text{subject to } \; y'\in S(x).$$

Note that the lower level is a *parameterized vector optimization problem,* the parameter x representing the strategy of the leader. Note also that, in fact there are considered (separately) three different bilevel problems ($\mathrm{BL_W}$), ($\mathrm{BL_e}$) and ($\mathrm{BL_p}$).

To synthesize the ideas, let us consider the set-valued map $\Psi : X\rightrightarrows Y$, defined for each $x\in X$ by

$$\Psi(x)=\{y\in Y|\ (x,y)\in U,\;\; y\in\sigma\text{-}\underset{y'}{\mathrm{ARGMIN}}_C\{F(x,y')|\, y'\in S(x)\}\}.$$

Thus, we have the following equivalent abstract formulation of problem (BL_σ) such as a MPEC problem:

$$\text{minimize}\, f(x,y)\;\; \text{subject to}\;\; y\in\Psi(x),\; x\in X.$$

We can write this problem equivalently as

$$\min_{x\in X}\min_{y\in\Psi(x)} f(x,y),$$

hence it represents the optimistic case.

The following problem represents the pessimistic case

$$(PBL_\sigma)\qquad \min_{x\in X}\max_{y\in\Psi(x)} f(x,y).$$

We can see that for both problems (optimistic and pessimistic), the inner optimization, i.e. the optimization with respect to y,

$$\min_{y\in\Psi(x)} f(x,y)$$

or

$$\max_{y\in\Psi(x)} f(x,y)$$

represents, for fixed x, the difficult program to optimize a scalar function over a (weakly or properly) efficient (called also Pareto) set.

Assuming that the partial map $y \mapsto f(x,y)$ is a quasi-convex function for each $x \in X$, and that the lower level is given by a linear vector optimization problem, it is possible to solve the inner optimization for the pessimistic case which becomes a combinatorial optimization problem (see e.g. [9]). Then one has to find necessary conditions with respect to x. This research will be done in a subsequent paper.

Let us go back to the optimistic case. If we consider also the set-valued function $\varphi : X \rightrightarrows \mathbb{R}$, given, for each $x \in X$, by $\varphi(x) = \{f(x,y)|y \in \Psi(x)\}$, we get another equivalent formulation of problem (BL_σ) as a *set-valued optimization problem*[(i)]

$$\text{minimize}\,\varphi(x)\;\text{ subject to }\; x \in X.$$

The fact that the problem (BL_σ) is a set-valued optimization problems comes from the fact that, even in the standard scalar case when $Z = \mathbb{R}$, for a given x, the optimal solution set of the lower level is not in general a singleton, and distinct optimal solutions yield distinct upper level objective values.

The problem ($\mathrm{BL_W}$) (without convexity assumptions) has been studied in [15] using a penalty approach, where its solutions are approximated by sequences of solutions to penalized problems.

The present paper deals not only with weakly-efficient solutions, but also with properly-efficient solutions of the lower level vector optimization problem. First order necessary conditions for the solutions to ($\mathrm{BL_W}$) and ($\mathrm{BL_P}$) are given. An application is proposed for the case when the lower level is a multiobjective linear programming problem having continuously differentiable data with respect to parameters. The conditions obtained are related to the simplex tableau.

The main motivation of this research is that the semivectorial bilevel problem (BL_σ) covers in particular the following important problems.

(i) For the set-valued optimization problem "minimize $\varphi(x)$ subject to $x \in X$", where $\varphi : X \rightrightarrows \mathbb{R}$, is a set-valued function, a solution is a point $x_0 \in X$, such that there exists $y_0 \in \varphi(x_0)$ verifying $y_0 = \min \cup_{x\in X}\varphi(x)$.

- *Optimization over a σ-efficient set* (OES) :

$$\underset{y}{\text{minimize}}\ f_1(y) \ \text{ subject to } y \in \Psi_1,$$

where Ψ_1 is the set of σ-efficient solutions associated to the vector optimization problem

$$\underset{y}{\text{MINIMIZE}}_C F_1(y) \quad \text{subject to } y \in S_1,$$

and $f_1 : Y \to \mathbb{R}$, $F_1 : Y \to Z$, and $S_1 \subseteq Y$. This problem is an important tool in the decision making theory. The main difficulty to solve (OES) is due to its non convexity. Even when all the functions are linear and S_1 is a polyhedron, the problem (OES) is not convex. Another difficulty is given by the fact that Ψ_1, the feasible set of (OES), is not known explicitly. For these reasons the problem (OES) was studied intensively during the last two decades by many authors. Thus, the "all linear case", i.e. when Ψ_1 is a linear functional and the vector optimization problem has linear objectives and constraints has been studied theoretically and some algorithms have been proposed in the papers [5, 6, 17, 23, 28]. Some nonlinear cases dealing also with theoretical aspects and algorithms may be found in [3, 8, 9, 10, 11, 19, 20, 24].

For a survey paper and an extensive bibliography see [30].

The problem (OES) is obtained as a particular case of the bilevel problem (BL_σ) considering $X = \{0\}$, $U = X \times Y$, and, for every $y \in Y$,

$$f(0, y) = f_1(y), \qquad F(0, y) = F_1(y), \qquad S(0) = S_1.$$

- *the usual (scalar) bilevel programming problem* (SBP) :

$$\underset{x,\,y}{\text{minimize}}\ f(x, y) \quad \text{subject to} \quad G(x, y) \leqslant 0,\ y \in \Psi(x),$$

where $\Psi(x)$ is the solution set of the (lower level) scalar minimization problem :

$$\underset{y'}{\text{minimize}}\ F(x, y') \quad \text{subject to} \quad H(x, y') \leqslant 0,$$

with $x \in \mathbb{R}^{n_x}$, $y \in \mathbb{R}^{n_y}$, $f,\, F : \mathbb{R}^{n_x+n_y} \to \mathbb{R}$, $G : \mathbb{R}^{n_x+n_y} \to \mathbb{R}^{n_u}$, $H : \mathbb{R}^{n_x+n_y} \to \mathbb{R}^{n_l}$.

We obtain the problem (SBP) from (BL_w) considering $X = \mathbb{R}^{n_x}$, $Y = \mathbb{R}^n$, $Z = \mathbb{R}$, $C = \mathbb{R}_+$, $U = \{(x, y) \in X \times Y |\ G(x, y) \leqslant 0\}$, and for each $x \in X$, $S(x) = \{y \in Y |\ H(x, y) \leqslant 0\}$.

This problem has been investigated by many authors (see, for example, [21] and [22] for an extensive and recent bibliography with more than 400 references!).

2 Preliminaries

Let $\|\cdot\|$ denote the norm of Z, and let $\langle\cdot,\cdot\rangle$ denote the duality scalar product between Z^* (the topological dual of Z) and Z. The extended space $\bar{Z} = Z \cup \{-\infty_C, +\infty_C\}$ is introduced in [12]. Recall that a neighborhood of $+\infty_C$ is a set $N \subseteq \bar{Z}$ containing $a + C \cup \{+\infty_C\}$ for some $a \in Z$, and its opposite $-N$ is a neighborhood of $-\infty_C$. The partial order relation, compatible with the linear structure of Z, given by

$$\forall z, z' \in Z, \qquad z \leqslant_C z' \Longleftrightarrow z' - z \in C,$$

and the transitive relation

$$\forall z, z' \in Z, \qquad z <_C z' \Longleftrightarrow z' - z \in \operatorname{int} C,$$

are extended to $\bar{Z}$ by

$$\forall z \in Z, \qquad -\infty_C <_C z <_C +\infty_C, \qquad -\infty_C \leqslant_C z \leqslant_C +\infty_C.$$

Note that the embedding $Z \subseteq \bar{Z}$ is continuous and dense.

Let G be an extended-valued map from Y to $Z \cup \{+\infty_C\}$, which is proper, i.e. not identically equal to $+\infty_C$, and let $S_1 \subseteq Y$ be a nonempty set.

The effective domain of G is dom $(G) = \{y \in Y |\ G(y) \neq +\infty_C\}$, the positive polar cone of C is $C^+ = \{\lambda \in Z^* |\ \langle \lambda, z\rangle \geq 0\ \forall z \in C\}$, and denote $C_0^+ = \{\lambda \in Z^* |\ \langle \lambda, z\rangle > 0\ \forall z \in C \setminus \{0\}\}$. We extend by continuity every $\lambda \in C^+ \setminus \{0\}$ to $\bar{Z}$, setting (see [12, 13] for more details),

$$\langle \lambda, \pm\infty_C \rangle = \pm\infty.$$

We denote by $\bar{T}$ the topological closure in the topological space $\bar{Z}$ of the set $T \subseteq \bar{Z}$. The *infimal set* (resp. *weakly infimal set* or *properly infimal set*) of a subset $T \subseteq \bar{Z}$, is the set

$$\mathrm{INF}_C(T) = \{z \in \bar{T} |\ \nexists v \in T \setminus \{z\},\ \ v \leqslant_C z\}$$

$$\text{(resp.} \qquad \text{w-INF}_C(T) = \{z \in \bar{T} |\ \nexists v \in T,\ \ v <_C z\} \qquad \text{or}$$

$$\text{p-INF}_C(T) = \{z \in \bar{T} |\ \exists K \subseteq Z \text{ convex pointed cone, } C \setminus \{0\} \subseteq \operatorname{int} K,\ z \in \mathrm{INF}_K(T)\}).$$

According to the definition, we obtain immediately that

$$\text{p-INF}_C(T) \subseteq \mathrm{INF}_C(T) \subseteq \text{w-INF}_C(T) \tag{1}$$

We recall the following result (see [12]).

Proposition 1. *Let T be a nonempty subset of $\bar{Z}$. The following statements are equivalent :*

(i) $-\infty_C \in \text{w-INF}_C(T)$

(ii) $\text{w-INF}_C(T) = \{-\infty_C\}$

(iii) *there exists a sequence $(z_n)_{n\in\mathbb{N}}$ in T such that $z_n \to -\infty_C$ (for the topology of $\bar{Z}$)*

(iv) $-\infty_C \in \bar{T}$.

For the vector optimization problem of the form

$$\text{MINIMIZE}_C\, G(y) \ \text{ s.t. } y \in S_1,$$

a point $a \in Y$ is called :

- *efficient* (or *Pareto*) solution, if $a \in S_1$, and there is no $y \in S_1$ verifying $G(y) \leqslant_C G(a)$, $G(y) \neq G(a)$ (in other words $G(a) \in \mathrm{INF}_C(G(S_1))$);
- *weakly-efficient* (or *weakly-Pareto*) solution if $a \in S_1$, and there is no $y \in S_1$ verifying $G(y) <_C G(a)$ (in other words $G(a) \in \text{w-INF}_C(G(S_1))$);

- *properly-efficient* (or *properly-Pareto*) solution, if $a \in S_1$ and $G(a) \in \text{p-INF}_C(G(S_1))$.

The set of efficient solutions (resp. weakly-efficient, or properly-efficient solutions) will be denoted $\text{e-ARGMIN}_C\{G(y) \mid y \in S_1\}$ (resp. $\text{w-ARGMIN}_C\{G(y) \mid y \in S_1\}$ or $\text{p-ARGMIN}_C\{G(y) \mid y \in S_1\}$).

According to (1) we obtain immediately the inclusions :

$$\begin{aligned}\text{p-ARGMIN}_C\{G(y) \mid y \in S_1\} &\subseteq \text{e-ARGMIN}_C\{G(y) \mid y \in S_1\}\\ &\subseteq \text{w-ARGMIN}_C\{G(y) \mid y \in S_1\}.\end{aligned} \tag{2}$$

To solve the vector optimization problem (VOP) :

$$\text{MINIMIZE}_C\, G(y) \ \text{ s.t. } y \in Y$$

means to find the set of the (weak or proper) efficient solutions.
Notice that a particular *constrained* vector optimization problem (CVOP)

$$\text{MINIMIZE}_C\, G_0(y) \ \text{ s.t. } y \in S_1$$

where $S_1 \subseteq Y$ is the feasible set, and G_0 is a map from S_1 to Z, is equivalent to the unconstrained extended-valued vector optimization problem (VOP) with

$$G(y) = \begin{cases} G_0(y) & \text{if } y \in S_1 \\ +\infty_C & \text{if } y \in Y \setminus S_1. \end{cases}$$

The problems (CVOP) and (VOP) are equivalent in the sense that they have the same efficient solutions (resp. weakly efficient and properly efficient solutions)), and the same infimal set (resp. weakly and properly infimal set).

Suppose Y be a topological vector space. A map $G : S_1 \subseteq Y \to Z \cup \{+\infty_C\}$, where S_1 is a convex set, is called $C-$*convex* if :

$$\forall y_1,\, y_2 \in S_1,\ \alpha \in]0,1[, \qquad G((1-\alpha)y_1 + \alpha y_2) \leqslant_C (1-\alpha)G(y_1) + \alpha G(y_2)$$

(with the conventions $\infty_C + \infty_C = \infty_C,\ \alpha \cdot (+\infty_C) = +\infty_C$, for each positive number α).

We recall the following scalarization result (see e.g. [25]) known for finite vector valued functions but which could be immediately generalized to extended valued vector functions.

Theorem 1. *Let Y be a topological vector space. If $S_1 \subseteq Y$ is a convex set and $G : S_1 \to Z \cup \{+\infty_C\}$ is a $C-$convex proper map, then we have*

$$\text{w-ARGMIN}_C \{G(y) \mid y \in S_1\} = \bigcup_{\lambda \in C^+ \setminus \{0\}} \operatorname{argmin}\{\langle \lambda, G(y)\rangle \mid y \in S_1\}$$

$$\text{p-ARGMIN}_C \{G(y) \mid y \in S_1\} = \bigcup_{\lambda \in C_0^+} \operatorname{argmin}\{\langle \lambda, G(y)\rangle \mid y \in S_1\}.$$

3 Optimality Conditions

3.1 Some Results about Set-Valued Map

Let $\mathcal{X}$, $\mathcal{Y}$ be real Banach spaces and let $G : \mathcal{X} \rightrightarrows \mathcal{Y}$ be a set-valued map, $A \subseteq \mathcal{X}$ and $B \subseteq \mathcal{Y}$. Denote

$$\operatorname{dom} G = \{x \in \mathcal{X} |\ G(x) \neq \emptyset\}, \ \ \operatorname{Gr}(G) = \{(x, y) \in \mathcal{X} \times \mathcal{Y} |\ y \in G(x)\},$$

$$G(A) = \bigcup_{x \in A} G(x) \qquad G^{-}(B) = \{x \in \mathcal{X} |\ G(x) \cap B \neq \emptyset\},$$

$G_A : \mathcal{X} \rightrightarrows \mathcal{Y}$, $G_A(x) := G(x)$ $(\forall x \in A)$, $\operatorname{dom} G_A := (\operatorname{dom} G) \cap A$.

The *contingent cone* $T(A, x)$ of the set A at the point $x \in A$ is the set of the elements $h \in \mathcal{X}$ such that there exists a sequence $(x_n)_{n \geqslant 1}$ of elements of A and a sequence $(t_n)_{n \geqslant 1}$ of positive real numbers such that

$$x = \lim x_n \quad \text{and} \quad h = \lim t_n(x_n - x).$$

The *contingent derivative* $DG(x_0, y_0) : \mathcal{X} \rightrightarrows \mathcal{Y}$ of G at $(x_0, y_0) \in \operatorname{Gr}(G)$ is defined by

$$\operatorname{Gr}(DG(x_0, y_0)) = T(\operatorname{Gr}(G), (x_0, y_0)).$$

This is equivalent to say that, for each $x \in \mathcal{X}$, $y \in DG(x_0, y_0)(x) \Longleftrightarrow$

$$\exists t_n > 0,\ (x_n, y_n) \in \operatorname{Gr}(G)\ :\ \lim(x_n, y_n) = (x_0, y_0) \ \text{ and } \ (x, y) = \lim t_n(x_n - x_0, y_n - y_0).$$

Let $\mathcal{Y} \supset \mathcal{C}$ be a closed convex pointed cone with nonempty interior. A point $(x_0, y_0) \in \operatorname{Gr}(G)$ is a *σ-solution* $(\sigma \in \{\mathrm{p}, \mathrm{e}, \mathrm{w}\})$ for the problem :

$$\sigma\text{-MINIMIZE}_{\mathcal{C}} G(x) \quad \text{s.t.} \quad x \in A \tag{3}$$

if $x_0 \in A$ and $y_0 \in \sigma\text{-INF}_{\mathcal{C}} G(A)$.

In the sequel we need the following results which we recall for reader's convenience.

Proposition 2. [4, pp. 442-443] *Let $\mathcal{X}, \mathcal{Y}, \mathcal{Z}$ be real Banach spaces, let $G : \mathcal{X} \rightrightarrows \mathcal{Y}$ be a set valued map, and let $g : \Omega \to \mathcal{Z}$ be a continuously differentiable (single-valued) function in an open neighborhood Ω of the image $G(\mathcal{X})$ of G. Then for all $x_0 \in \mathcal{X}$, $y_0 \in G(x_0)$,*

$$\forall x \in \mathcal{X}, \qquad D(g \circ G)(x_0, g(y_0))(x) \supset \nabla g(y_0) \cdot DG(x_0, y_0)(x), \tag{4}$$

where $\nabla g(y_0)$ denotes the Fréchet derivative of g at y_0.

Theorem 2. [18] *If (x_0, y_0) is a σ-solution of problem (3) then*

$$\forall x \in \mathcal{X}, \qquad DG_A(x_0, y_0)(x) \cap \operatorname{int}(-\mathcal{C}) = \emptyset. \tag{5}$$

3.2 Back to the Bilevel Problem (BL_σ)

Let $\mathcal{E} : X \times Z^* \rightrightarrows Y$ be given by

$$\mathcal{E}(x,\lambda) = \begin{cases} \operatorname*{argmin}_{y\in S(x)} \langle \lambda, F(x,y)\rangle & \text{if } (x,\lambda) \in X \times C^+ \setminus \{0\} \\ \emptyset & \text{if } (x,\lambda) \notin X \times C^+ \setminus \{0\} \end{cases}$$

Denote

$$\Lambda_\sigma = \begin{cases} C^+ \setminus \{0\} & \text{if } \sigma = w, \\ C_0^+ & \text{if } \sigma = p. \end{cases}$$

Then, according to Theorem 1, we have :

$$\forall x \in X,\ \ \sigma\text{-}\mathrm{ARGMIN}_{\substack{C \\ y \in S(x)}} F(x,y) = \mathcal{E}(x,\Lambda_\sigma), \quad \sigma \in \{\mathrm{w},\, \mathrm{p}\}.$$

Consider the set-valued map $\Phi_\sigma : X \times Z^* \rightrightarrows X \times Y$ given by

$$\Phi_\sigma(x,\lambda) = \begin{cases} \Big(\{x\} \times \mathcal{E}(x,\lambda)\Big) \cap U & \text{if } (x,\lambda) \in X \times \Lambda_\sigma \\ \emptyset & \text{if } (x,\lambda) \notin X \times \Lambda_\sigma \end{cases} \quad \sigma \in \{\mathrm{w,p}\}, \tag{6}$$

and the ***scalar set-valued minimization problem :***

$$(\mathrm{SSM}_\sigma) \qquad \boxed{\text{minimize}\,(f \circ \Phi_\sigma)(x,\lambda) \ \text{ subject to } (x,\lambda) \in X \times Z^*.}$$

Recall that a point $((x_0,\lambda_0),t_0) \in \mathrm{Gr}\,(f\circ\Phi_\sigma)$ is a *solution* of (SSM_σ)

$$\Longleftrightarrow \qquad t_0 = \inf\,(f\circ\Phi_\sigma)(X \times Z^*).$$

Proposition 3. *Let $\sigma \in \{\mathrm{w,p}\}$. The problem ($BL_\sigma$) is equivalent to the problem (SSM_σ) in the following sense.*

If (x_0,y_0) is a solution of (BL_σ), then $\mathcal{E}(x_0,\cdot)^-_{\Lambda_\sigma}(\{y_0\}) \neq \emptyset$ and, for each $\lambda_0 \in \mathcal{E}(x_0,\cdot)^-_{\Lambda_\sigma}(\{y_0\})$, the point $((x_0,\lambda_0), f(x_0,y_0))$ is a solution of (SSM_σ).

Conversely, if $((x_0,\lambda_0),t_0)$ is a solution of (SSM_σ), then there exists $y_0 \in \mathcal{E}(x_0,\lambda_0)$ such that (x_0,y_0) is a solution of (BL_σ) and $t_0 = f(x_0,y_0)$.

Proof. Let (x_0,y_0) be a solution of (BL_σ). Then $(x_0,y_0) \in U$, $y_0 \in \sigma\text{-}\mathrm{ARGMIN}_{\substack{C \\ y \in S(x_0)}} F(x_0,y) = \mathcal{E}(x_0,\Lambda_\sigma)$ and, for all $(x,y) \in U$ such that $y \in \sigma\text{-}\mathrm{ARGMIN}_{\substack{C \\ y \in S(x)}} F(x,y)$, we have $f(x_0,y_0) \leqslant f(x,y)$. There exists $\bar\lambda \in \Lambda_\sigma$ such that $y_0 \in \mathcal{E}(x_0,\bar\lambda)$, hence $(x_0,y_0) \in \Phi_\sigma(x_0,\bar\lambda)$. Let $\lambda_0 \in \mathcal{E}(x_0,\cdot)^-_{\Lambda_\sigma}(\{y_0\}) \neq \emptyset$ (because $\bar\lambda \in \mathcal{E}(x_0,\cdot)^-_{\Lambda_\sigma}(\{y_0\})$. It follows that $(x_0,y_0) \in \Phi_\sigma(x_0,\lambda_0)$, hence $((x_0,\lambda_0), f(x_0,y_0)) \in \mathrm{Gr}\,(f\circ\Phi_\sigma)$. Let $(x,\lambda) \in \mathrm{dom}\,(f\circ\Phi_\sigma) = \mathrm{dom}\,\mathcal{E}$ and $t \in (f\circ\Phi_\sigma)(x,\lambda)$. There exists $y \in \mathcal{E}(x,\lambda) \subseteq \sigma\text{-}\mathrm{ARGMIN}_{\substack{C \\ y \in S(x)}} F(x,y)$ such that $(x,y) \in U$ and $t = f(x,y)$. Thus (x,y) is feasible for problem (BL_σ), hence $f(x_0,y_0) \leqslant f(x,y)$. We conclude that $f(x_0,y_0) = \inf(f\circ\Phi_\sigma)(X\times Z^*)$, therefore $((x_0,y_0), f(x_0,y_0))$ is a solution of (SSM_σ).

Conversely, let $((x_0,\lambda_0),t_0)$ be a solution of (SSM_σ). We have $t_0 \in (f\circ\Phi_\sigma)(x_0,\lambda_0)$, hence there exists $y_0 \in \mathcal{E}(x_0,\lambda_0)$ such that $(x_0,y_0) \in \Phi_\sigma(x_0,\lambda_0)$ and $t_0 = f(x_0,y_0)$. Thus

$(x_0, y_0) \in U$ and $y_0 \in \sigma\text{-}\underset{y \in S(x_0)}{\mathrm{ARGMIN}_C} F(x_0, y)$, i.e. (x_0, y_0) is feasible for problem (BL_σ). On the other hand, since $t_0 = \inf(f \circ \Phi_\sigma)(X \times Z^*)$ and any feasible solution (x, y) of problem (BL_σ) verifies $(x, y) \in \Phi_\sigma(x, \lambda)$ for some $\lambda \in \Lambda_\sigma$, we obtain $f(x_0, y_0) \leqslant f(x, y)$, hence (x_0, y_0) is a solution of (BL_σ). □

Now we can state the results about the optimality conditions.

Theorem 3. *Let f be continuously differentiable and let (x_0, y_0) be a minimizer for problem (BL_σ), $\sigma \in \{\mathrm{w, p}\}$. Then $\mathcal{E}(x_0, \cdot)^-_{\Lambda_\sigma}(\{y_0\}) \neq \emptyset$ and, for each $\lambda_0 \in \mathcal{E}(x_0, \cdot)^-_{\Lambda_\sigma}(\{y_0\})$, and for each $(x, \lambda) \in X \times Z^*$,*

$$\nabla f(x_0, y_0) \cdot D\Phi_\sigma((x_0, \lambda_0), (x_0, y_0))(x, \lambda) \subseteq [0, +\infty[.$$

Proof. Let (x_0, y_0) be a solution to (BL_σ). According to Proposition 3, the set $\mathcal{E}(x_0, \cdot)^-_{\Lambda_\sigma}(\{y_0\})$ is not empty and, for any $\lambda_0 \in \mathcal{E}(x_0, \cdot)^-_{\Lambda_\sigma}(\{y_0\})$, the point $((x_0, \lambda_0), f(x_0, y_0))$ is a solution to (SSM_σ). By Theorem 2, since $f \circ \Phi_\sigma$ is real set-valued and the order cone in $\mathbb{R}$ is the half-line $[0, +\infty[$, we have that for all $(x, \lambda) \in X \times Z^*$

$$D(f \circ \Phi_\sigma)((x_0, \lambda_0), f(x_0, y_0))(x, \lambda) \subseteq [0, +\infty[.$$

Using Proposition 2 we have that

$$\forall (x, \lambda) \in X \times Z^*,$$
$$\nabla f(x_0, y_0) \cdot D\Phi_\sigma((x_0, \lambda_0), (x_0, y_0))(x, \lambda) \subseteq D(f \circ \Phi_\sigma)((x_0, \lambda_0), f(x_0, y_0))(x, \lambda)$$

which completes the proof. □

Next we will give a more explicit form of the above theorem.

Notice first that, using the definition of Φ_σ and of its contingent derivative, it is easy to see that for all $(x, \lambda, x', y) \in X \times Z^* \times X \times Y$, we have

$$(x', y) \in D\Phi_\sigma((x_0, \lambda_0), (x_0, y_0))(x, \lambda) \iff$$
$$x' = x, \;\; (x, y) \in T(U, (x_0, y_0)) \;\text{ and } y \in D\mathcal{E}(x_0, \lambda_0)(x, \lambda).$$

Thus, denoting by $\nabla_1 f(x_0, y_0) \in X^*$ and $\nabla_2 f(x_0, y_0) \in Y^*$ the partial Fréchet derivatives of f in (x_0, y_0)[(ii)], we can state the following.

Theorem 4. *Let f be continuously differentiable and let (x_0, y_0) be a minimizer for (BL_σ), $\sigma \in \{\mathrm{w, p}\}$. Then $\mathcal{E}(x_0, \cdot)^-_{\Lambda_\sigma}(\{y_0\}) \neq \emptyset$ and, for each $\lambda_0 \in \mathcal{E}(x_0, \cdot)^-_{\Lambda_\sigma}(\{y_0\})$, and for all $(x, y, \lambda) \in X \times Y \times Z^*$,*

$$(x, y) \in T(U, (x_0, y_0)), \;\; y \in D\mathcal{E}(x_0, \lambda_0)(x, \lambda) \Longrightarrow \nabla_1 f(x_0, y_0)(x) + \nabla_2 f(x_0, y_0)(y) \geqslant 0. \quad (7)$$

3.3 The Case when the Lower Level is a Linear Multiobjective Optimization Problem

In this subsection is studied the important particular case when X is a finite dimensional normed linear space, Ω is an open subset of X, $Y = \mathbb{R}^p$, $Z = \mathbb{R}^r$, $C = \mathbb{R}^r_+$, $U = \Omega \times Y$, and the lower level is linear in y, i.e. we deal with the semilinear semivectorial bilevel problem :

$$\text{(BLL)} \qquad \min_{x,y} f(x, y) \;\text{ subject to}$$

[(ii)] i.e., the Fréchet derivatives of the functions $x \mapsto f(x, y_0)$ and $y \mapsto f(x_0, y)$

$$(x, y) \in \Omega \times \mathbb{R}^p, \text{ and for any } x \in \Omega,\ y \text{ is an efficient solution to}$$
$$\text{e-MINIMIZE}_{y'} {}_{\mathbb{R}^r_+} \{C(x)y' | A(x)y' = b(x),\ y' \geqslant 0\}.$$

The functions[(iii)] $f : \Omega \times \mathbb{R}^p \to \mathbb{R}$, $A : \Omega \to \mathbb{R}^{m \times p}$, $b : \Omega \to \mathbb{R}^m$, $C : \Omega \to \mathbb{R}^{r \times p}$ are continuously differentiable. Of course, $m < p$, and we suppose that

$$(H1) \qquad \text{rank of } A(x) = m \text{ for all } x \in \Omega.$$

It is well known that for a linear multiobjective problem the efficient set and the properly efficient set coincide. So, we have the same problem if we replace the symbol "e" by "p" in the lower level.

For any matrix $M \in \mathbb{R}^{l \times q}$, we denote m_i its i^{th} column, and for any index subset $I = \{i_1, \ldots, i_k\} \subset \{1, \ldots, q\}$ with $i_1 < \ldots < i_k$, we denote $M_I = [m_{i_1} \ \ldots m_{i_k}] \in \mathbb{R}^{l \times k}$. Also, if $z = [z_1 \cdots z_q]^T \in \mathbb{R}^q$ is a (column) vector, we denote $z_I = [z_{i_1} \ \cdots \ z_{i_k}]^T \in \mathbb{R}^k$.

Recall that an index set $B = \{i_1, \ldots, i_m\} \subset \{1, \ldots, p\}$, $i_1 < \cdots < i_m$ is called a *basis* of the matrix $A(x)$, $(x \in \Omega)$ if the matrix $A(x)_B$ is invertible. For any basis B of $A(x)$, denoting $\hat{B} = \{1, \ldots, p\} \setminus B$, we can write the system $A(x)y = b(x)$ in an equivalent form as

$$y_B = A(x)_B^{-1}(b(x) - A(x)_{\hat{B}} y_{\hat{B}}). \tag{8}$$

For any solution y to the above system, we have

$$C(x)y = (C(x)_{\hat{B}} - C(x)_B A(x)_B^{-1} A(x)_{\hat{B}}) y_{\hat{B}} + C(x)_B A(x)_B^{-1} b(x). \tag{9}$$

The solution y^B with $y^B_{\hat{B}} = 0$, hence $y^B_B = A(x)_B^{-1} b(x)$, is called *basic solution* of the system (8).

We say that the basis B of $A(x)$ is *feasible* for the convex polyhedron $S(x) := \{y \in \mathbb{R}^p |\ A(x)y = b(x),\ y \geqslant 0\}$ iff $A(x)_B^{-1} b(x) \geqslant 0$. The associated basic solution is called *basic feasible solution* of $S(x)$. It is well known that, for any point y belonging to $S(x)$, we have

$$y \text{ is a basic feasible solution } \iff y \text{ is an extremal point (vertex) of } S(x)^{(\text{iv})}.$$

For each $x \in \Omega$, denote $V(x)$ the set of vertices of $S(x)$,

$$\mathcal{B}(x) = \{B \subset \{1, \ldots, n\} \mid B \text{ is a feasible basis for } S(x)\}.$$

Denote also

$$\mathcal{B} = \bigcup_{x \in \Omega} \mathcal{B}(x).$$

Notice that $\mathcal{B}$ is a finite set.

Let $x \in \Omega$, $B \in \mathcal{B}(x)$, and $\lambda \in \operatorname{int} \mathbb{R}^r_+$. Since $S(x) = \{y \in \mathbb{R}^p |\ y_B \text{ satisfies (8)}, y_B \geqslant 0,\ y_{\hat{B}} \geqslant 0\}$, it is clear from (9) that the basic feasible solution y^B associated to B is an optimal solution to the scalarized problem associated to (x, λ) :

$$\min_y \lambda^T C(x) y \quad \text{s.t.} \quad y \in S(x),$$

[(iii)] We can consider everywhere in this paper that the function f is defined on a set of the form $\Omega \times Y$ where Ω is some open set in X (modifying accordingly the domain of Φ_σ replacing X by Ω in (6)).

[(iv)] y is an extremal point of the convex set $S(x)$ (called also vertex since $S(x)$ is a polyhedron) if there is no distinct points $y', y'' \in S(x)$ such that $y = \alpha y' + (1 - \alpha) y''$ for some $\alpha \in]0, 1[$.

iff all the coefficients of the non basic variables in the reduced objective are positive. Hence

$$y^B \in \mathcal{E}(x,\lambda) \Longleftrightarrow \lambda^T(C(x)_{\hat{B}} - C(x)_B A(x)_B^{-1} A(x)_{\hat{B}}) \geqslant 0. \tag{10}$$

For every $(x,\lambda) \in \Omega \times \operatorname{int}(\mathbb{R}^r_+)$, let us denote

$$\mathcal{B}(x,\lambda) = \{B \in \mathcal{B}(x) :\ y^B \in \mathcal{E}(x,\lambda)\}.$$

Thus, relation (10) may be rewritten in the form

$$B \in \mathcal{B}(x,\lambda) \Longleftrightarrow \lambda^T(C(x)_{\hat{B}} - C(x)_B A(x)_B^{-1} A(x)_{\hat{B}}) \geqslant 0, \text{ and } A(x)_B^{-1} b(x) \geqslant 0. \tag{11}$$

Moreover, it is easy to see that, for any $B \in \mathcal{B}(x,\lambda)$ and $y \in \mathbb{R}^p$, we have

$$y \in \mathcal{E}(x,\lambda) \Longleftrightarrow y \in S(x) \text{ and } \lambda^T(C(x)_{\hat{B}} - C(x)_B A(x)_B^{-1} A(x)_{\hat{B}}) y_{\hat{B}} = 0. \tag{12}$$

Since the solution set of a linear programming problem is a face of the feasible polyhedron which coincide with the convex hull of the basic optimal solutions (see [31, Theorem 2.4.12]), we have :

$$\forall (x,\lambda) \in \Omega \times \operatorname{int} \mathbb{R}^r_+, \qquad \mathcal{E}(x,\lambda) = \operatorname{conv}\,(V(x) \cap \mathcal{E}(x,\lambda)) = \operatorname{conv}\,\{y^B |\ B \in \mathcal{B}(x,\lambda)\}. \tag{13}$$

Recall that a basis $B \in \mathcal{B}(x)$ is called *not degenerate* if $A(x)_B^{-1} b(x) > 0$.

We make the following assumption :

$(H2)$ For every $x \in \Omega$, every basis $B \in \mathcal{B}(x)$ is not degenerate.

In other words, for each $B \in \mathcal{B}$ and for all $x \in \Omega$,

$$\det A(x)_B \neq 0,\ \ A(x)_B^{-1} b(x) \geqslant 0 \Longrightarrow A(x)_B^{-1} b(x) > 0.$$

We obtain easily the following.

REMARK 1. For each $B \in \mathcal{B}$, the set $D_B := \{x \in \Omega |\ B \in \mathcal{B}(x)\}$ is open.

Indeed, denoting $S_B = \{x \in \Omega |\ \det A(x)_B \neq 0\}$ (which is an open set), we have $D_B = \{x \in S_B |\ A(x)_B^{-1} b(x) \geqslant 0\}$, and $(H2)$ implies $\{x \in S_B |\ A(x)_B^{-1} b(x) \geqslant 0\} = \{x \in S_B |\ A(x)_B^{-1} b(x) > 0\}$. The last set is obviously open.

Lemma 1. *Let $(x_0,\lambda_0) \in \Omega \times \operatorname{int} \mathbb{R}^r_+$. Then there exists an open neighborhood U of (x_0,λ_0) in $\Omega \times \operatorname{int}(\mathbb{R}^r_+)$ such that, for each $(x,\lambda) \in U$ we have*

$$\mathcal{B}(x,\lambda) \subseteq \mathcal{B}(x_0,\lambda_0).$$

Proof. Let us assume by contradiction that, for each open neighborhood U of (x_0,λ_0), there exists $(x_U,\lambda_U) \in U$ such that $\mathcal{B}(x_U,\lambda_U) \not\subset \mathcal{B}(x_0,\lambda_0)$. Thus we can find a sequence $(x_n,\lambda_n)_{n\geqslant 1}$ converging to (x_0,λ_0) and a sequence (B_n) with $B_n \in \mathcal{B}(x_n,\lambda_n)$ such that, for all n we have $B_n \notin \mathcal{B}(x_0,\lambda_0)$. Since $B_n \in \mathcal{B}$ which is a finite set, we can find a strictly increasing function $\varphi : \mathbb{N} \to \mathbb{N}$ such that the subsequence $(B_{\varphi(n)})_{n\geqslant 1}$ is constant. Thus $B_{\varphi(n)} = B$ for all n, and we have

$$\lambda^T_{\varphi(n)}(C(x_{\varphi(n)})_{\hat{B}} - C(x_{\varphi(n)})_B A(x_{\varphi(n)})_B^{-1} A(x_{\varphi(n)})_{\hat{B}}) \geqslant 0, \text{ and } A(x_{\varphi(n)})_B^{-1} b(x_{\varphi(n)}) > 0.$$

Letting $n \to +\infty$, we obtain

$$\lambda_0^T(C(x_0)_{\hat{B}} - C(x_0)_B A(x_0)_B^{-1} A(x_0)_{\hat{B}}) \geqslant 0, \text{ and } A(x_0)_B^{-1} b(x_0) \geqslant 0,$$

which means that $B \in \mathcal{B}(x_0,\lambda_0)$. Contradiction! □

For any $B \in \mathcal{B}$ define the continuously differentiable maps

$$\begin{aligned} x \mapsto R(B;x) &:= C(x)_{\hat{B}} - C(x)_B A(x)_B^{-1} A(x)_{\hat{B}} \\ x \mapsto S(B;x) &:= A(x)_B^{-1} A(x)_{\hat{B}} \\ x \mapsto U(B;x) &:= A(x)_B^{-1} b(x) \end{aligned}$$

from the open set D_B (see Remark 1) to $\mathbb{R}^{r\times(p-m)}$, $\mathbb{R}^{m\times(p-m)}$ and $\mathbb{R}^m$ respectively.

Notice that $(H2)$ is equivalent to say that, for any $B \in \mathcal{B}$ and $x \in D_B$, we have $U(B;x) > 0$.

Relation (10) implies that, for any $(x_0, \lambda_0) \in (\Omega \times \operatorname{int} \mathbb{R}^r_+) \cap \operatorname{dom} \mathcal{E}$ and $\mathcal{B} \in \mathcal{B}(x_0)$,

$$B \in \mathcal{B}(x_0, \lambda_0) \Longleftrightarrow \lambda_0^T R(B;x_0) \geqslant 0. \tag{14}$$

Also it is obvious that, for any $B \in \mathcal{B}(x_0, \lambda_0)$ and $y_0 \in \mathbb{R}^p$, we have $y_0 \in \mathcal{E}(x_0, \lambda_0)$ iff

$$\begin{aligned} &S(B;x_0) y_{0\hat{B}} \leqslant U(B;x_0), \\ &y_{0B} = U(B;x_0) - S(B;x_0) y_{0\hat{B}}, \\ &y_{0\hat{B}} \geqslant 0, \text{ and } \lambda_0^T R(B;x_0) y_{0\hat{B}} = 0. \end{aligned} \tag{15}$$

Notice that in this case

$$\lambda_0^T R(B;x_0) y_{0\hat{B}} = 0 \Longleftrightarrow \lambda_0^T R(B;x_0)_i y_{0i} = 0 \ \ \forall i \in \hat{B}.$$

Let us recall that the *relative interior* of a convex set K in $\mathbb{R}^p$, denoted ri K is its interior relative to its affine hull. We have (see e.g. [16, page 8])

$$\text{ri } K = \{y \in K | \ (\forall z \in K)(\exists \varepsilon > 0) \quad y + \varepsilon(y - z) \in K\}.$$

Lemma 2. *Let* $(x_0, \lambda_0) \in (\Omega \times \operatorname{int} \mathbb{R}^r_+) \cap \operatorname{dom} \mathcal{E}$, $y_0 \in \mathcal{E}(x_0, \lambda_0)$ *and* $\mathcal{B} \in \mathcal{B}(x_0, \lambda_0)$. *Denote*

$$I(x_0, \lambda_0, B) = \{i \in \hat{B} | \ \lambda_0^T R(B;x_0)_i = 0\}. \tag{16}$$

Then, $y_0 \in \text{ri } \mathcal{E}(x_0, \lambda_0)$ *iff*

$$\forall i \in I(x_0, \lambda_0, B), \qquad y_{0i} > 0 \tag{17}$$

$$S(B;x_0) y_{0\hat{B}} < U(B;x_0). \tag{18}$$

Proof. Suppose $y_0 \in \text{ri } \mathcal{E}(x_0, \lambda_0)$. Since $y_0 \in \mathcal{E}(x_0, \lambda_0)$, from (15) and the definition of $I(x_0, \lambda_0, B)$ we obtain

$$S(B;x_0) y_{0\hat{B}} = \sum_{i \in I(x_0, \lambda_0, B)} A(x_0)_B^{-1} A(x_0)_i y_{0i}.$$

Let $i \in I(x_0, \lambda_0, B)$. Denote $v^i \in \mathbb{R}^{\hat{B}}$ the vector defined by $v^i_j = 0$ for all $j \in \hat{B} \setminus \{i\}$, and $v^i_i = 1$. Since assumption $(H2)$ holds, there exists $\delta > 0$ such that $\delta S(B;x_0) v^i < U(B;x_0)$. Then the vector $y \in \mathbb{R}^p$ defined by $y_{\hat{B}} = y_{0\hat{B}} + \delta v^i$ and $y_B = U(B;x_0) - S(B;x_0) y_{\hat{B}}$ belongs to $\mathcal{E}(x_0, \lambda_0)$. It follows that there exists $\varepsilon > 0$ such that $y_0 + \varepsilon(y_0 - y) \in \mathcal{E}(x_0, \lambda_0)$. This implies $y_{0i} - \varepsilon \cdot \delta \geqslant 0$, hence (17) holds. To prove (18), assume by contradiction that $S(B;x_0) y_{0\hat{B}} \not< U(B;x_0)$. Since $S(B;x_0) y_{0\hat{B}} \leqslant U(B;x_0)$, it follows thar there exists $j \in \{1, \ldots, m\}$ such that $L_j y_{0\hat{B}} = \beta_j$ where L_j is the j^{th} row of the matrix $S(B;x_0)$ and β_j

is the j^{th} coordinate of the column vector $U(B;x_0)$, hence $\beta_j > 0$ from $(H2)$. Let $y \in \mathbb{R}^p$ be defined by $y_{\hat{B}} = \frac{1}{2} y_{0\hat{B}}$ and $y_B = U(B;x_0) - S(B;x_0)y_{\hat{B}}$. Then y belongs to $\mathcal{E}(x_0,\lambda_0)$. Thus, there exists $\varepsilon > 0$ such that $y_0 + \varepsilon(y_0 - y) \in \mathcal{E}(x_0,\lambda_0)$. Since $(y_0 + \varepsilon(y_0 - y))_{\hat{B}} = (1 + \frac{\varepsilon}{2})y_{0\hat{B}}$, we obtain $(1 + \frac{\varepsilon}{2})\beta_j = L_j(1 + \frac{\varepsilon}{2})y_{0\hat{B}} \leqslant \beta_j$ which is a contradiction! Hence (18) holds.

Conversely, let $y_0 \in \mathcal{E}(x_0,\lambda_0)$ verify (17, 18). For any $y \in \mathcal{E}(x_0,\lambda_0)$, there exists $\varepsilon > 0$ such that $S(B;x_0)(y_0 + \varepsilon(y_0 - y))_{\hat{B}} < U(B;x_0)$ and $(y_0 + \varepsilon(y_0 - y))_{I(x_0,\lambda_0,B)} > 0$. Moreover $y_{0i} = y_i = 0$ for all $i \in \hat{B} \setminus I(x_0,\lambda_0,B)$. Therefore $y_0 + \varepsilon(y_0 - y) \in \mathcal{E}(x_0,\lambda_0)$. □

For each $(x_0,\lambda_0) \in (\Omega \times \text{int}\,\mathbb{R}^r_+) \cap \text{dom}\,\mathcal{E}$, $B \in \mathcal{B}(x_0,\lambda_0)$, and $y_0 \in \mathcal{E}(x_0,\lambda_0)$, denote

$$\mathcal{G}(B;x_0,\lambda_0,y_0) = \{(x,\lambda,y) \in X \times \mathbb{R}^r \times \mathbb{R}^p \mid \quad \text{relations } (19-23) \text{ are verified}\}$$

$$\begin{aligned}
\lambda^T R(B;x_0)_i + \lambda_0^T(\nabla R(B;x_0)(x))_i &\geqslant 0, \qquad \forall i \in I(x_0,\lambda_0,B), && (19)\\
\forall i \in \hat{B}, \qquad y_{0i} > 0 &\implies \lambda^T R(B;x_0)_i + \lambda_0^T(\nabla R(B;x_0)(x))_i\\
&\quad = 0. && (20)\\
y_B &= \nabla U(B;x_0)(x) - \nabla S(B;x_0)(x)y_{0\hat{B}}\\
&\quad -S(B;x_0)y_{\hat{B}} && (21)\\
\lambda_0^T R(B;x_0)y_{\hat{B}} &= 0, && (22)\\
\forall i \in \{1,\dots,p\}, \quad y_{0i} = 0 &\implies y_i \geqslant 0. && (23)
\end{aligned}$$

Notice that we can replace in the definition of $\mathcal{G}(B,x_0,\lambda_0,y_0)$ (22) by the relation

$$\forall i \in \hat{B}, \qquad \lambda_0^T R(B;x_0)_i y_i = 0. \tag{24}$$

Indeed, for any $i \in \hat{B}$ such that $\lambda_0^T R(B;x_0)_i \neq 0$, (14) implies $\lambda_0^T R(B;x_0)_i > 0$, hence $y_{0i} = 0$ (from (15)). Then, from (23) we obtain $y_i \geqslant 0$. Thus it follows easily that (22) implies (24). The converse is obvious.

Theorem 5. *Let $(x_0,\lambda_0) \in (\Omega \times \text{int}\,\mathbb{R}^r_+) \cap \text{dom}\,\mathcal{E}$[(v)]. Then, for all $y_0 \in \mathcal{E}(x_0,\lambda_0)$, we have that*

$$\text{Gr}\,(D\mathcal{E}((x_0,\lambda_0),y_0)) \subseteq \bigcup_{B \in \mathcal{B}(x_0,\lambda_0)} \mathcal{G}(B,x_0,\lambda_0,y_0). \tag{25}$$

Moreover, suppose

$$y_0 \in \text{ri}\,(\mathcal{E}(x_0,\lambda_0)). \tag{26}$$

Then,

A. $$\text{Gr}\,(D\mathcal{E}((x_0,\lambda_0),y_0)) = \bigcup_{B \in \mathcal{B}(x_0,\lambda_0)} T(E_{(x_0,\lambda_0,B)};(x_0,\lambda_0)) \times \{y \in \mathbb{R}^p \mid y_{\hat{B}\setminus I(x_0,\lambda_0,B)} = 0\}, \tag{27}$$

where

$$E_{(x_0,\lambda_0,B)} = \{(x,\lambda) \in D_B \times \text{int}\,\mathbb{R}^r_+ \mid \lambda^T R(B;x)_{I(x_0,\lambda_0,B)} = 0\}.$$

(v) Notice that if $S(x)$ is compact for all $x \in \Omega$, then $\text{dom}\,\mathcal{E} = \Omega \times \mathbb{R}^r_+$

B. *For each* $B \in \mathcal{B}(x_0, \lambda_0)$ *such that*

$$\text{rank } \left[R(B;x_0)_i^T \qquad \lambda_0^T \nabla R(B;x_0)_i\right]_{i\in I(x_0,\lambda_0,B)} = |I(x_0,\lambda_0,B)|, \tag{28}$$

we have

$$\text{Gr } (D\mathcal{E}((x_0,\lambda_0),y_0)) \supset \mathcal{G}(B,x_0,\lambda_0,y_0). \tag{29}$$

Proof. Let $(x,\lambda,y) \in \text{Gr}\,(D\mathcal{E}((x_0,\lambda_0),y_0))$. There exists sequences $(t_n)_{n\geqslant 1}$ of positive real numbers, and $(x_n,\lambda_n,y_n) \in \text{Gr}\,\mathcal{E}((x_0,\lambda_0),y_0)$ such that $(x_n,\lambda_n,y_n) \to (x_0,\lambda_0,y_0)$ and $t_n(x_n - x_0, \lambda_n - \lambda_0, y_n - y_0) \to (x,\lambda,y)$. Let us denote $\varepsilon_n = 1/t_n$. Assuming $(x,\lambda,y) \neq (0,0,0)$ (if not, relations (19,20,21-23) are obviously fulfilled), we obtain that $\varepsilon_n \to 0$. We have

$$x_n = x_0 + \varepsilon_n x + o(\varepsilon_n); \quad \lambda_n = \lambda_0 + \varepsilon_n \lambda + o(\varepsilon_n), \quad y_n = y_0 + \varepsilon_n y + o(\varepsilon_n), \tag{30}$$

where $o(\varepsilon_n)$ stands for any sequence of X, $\mathbb{R}^r$ or $\mathbb{R}^p$ such that $\frac{1}{\varepsilon_n} o(\varepsilon_n) \to 0$. Using Lemma 1, for sufficiently large n, $\mathcal{B}(x_n,\lambda_n) \subset \mathcal{B}(x_0,\lambda_0)$. Thus, taking eventually a subsequence without relabelling, we obtain that there is $B \in \mathcal{B}(x_0,\lambda_0)$ such that, $B \in \mathcal{B}(x_n,\lambda_n)$. Since $y_n \in \mathcal{E}(x_n,\lambda_n)$,

$$\lambda_n^T R(B;x_n) \geqslant 0, \;\; \lambda_n^T R(B;x_n) y_{n\hat{B}} = 0, \;\; y_{nB} = U(B;x_n) - S(B;x_n) y_{n\hat{B}} \text{ and } y_n \geqslant 0. \tag{31}$$

Using the differentiability of $R(B;\cdot)$, $S(B;\cdot)$ and $U(B;\cdot)$ at x_0, we obtain

$$\lambda_0^T R(B;x_0) + \varepsilon_n(\lambda^T R(B;x_0) + \lambda_0^T \nabla R(B;x_0)) + o(\varepsilon_n) \geqslant 0, \tag{32}$$

$$\lambda_0^T R(B;x_0) y_{0\hat{B}} + \varepsilon_n \Big((\lambda^T R(B;x_0) + \lambda_0^T (\nabla R(B;x_0)(x))) y_{0\hat{B}} + \lambda_0^T R(B;x_0) y_{\hat{B}} \Big) + o(\varepsilon_n) = 0, \tag{33}$$

$$\begin{aligned} y_{0B} + \varepsilon_n y_B = U(B;x_0) - S(B;x_0) y_{0\hat{B}} + \varepsilon_n(\nabla U(B;x_0)(x) - \nabla S(B;x_0)(x) y_{0\hat{B}} \\ - S(B;x_0) y_{\hat{B}}) + o(\varepsilon_n), \end{aligned} \tag{34}$$

and

$$y_0 + \varepsilon_n y + o(\varepsilon_n) \geqslant 0. \tag{35}$$

For each $i \in \hat{B}$ such that $\lambda_0^T R(B;x_0)_i = 0$ (in particular this holds if $y_{0i} > 0$ since relations (14, 15) hold), taking the i^{th} coordinate in (32), dividing by ε_n and letting $n \to +\infty$, we obtain (19).

Letting $n \to +\infty$ in (33), we obtain (22). Moreover, let us divide by ε_n the i^{th} coordinate in relation (33), and letting $n \to +\infty$, we obtain $(\lambda^T R(B;x_0)_i + \lambda_0^T(\nabla R(B;x_0)(x))) y_{0i} = 0$, hence (20) holds.

Since $y_{0B} = U(B;x_0) - S(B;x_0) y_{0\hat{B}}$, dividing by ε_n and letting $n \to +\infty$ in (34), we obtain (21).

Finally, let $i \in \{1,\ldots,p\}$ such that $y_{0i} = 0$. Dividing again by ε_n the coordinate number i of (35), and letting $n \to +\infty$ we obtain (23). Hence $(x,\lambda,y) \in \mathcal{G}(B,x_0,\lambda_0,y_0)$.

Suppose now that (26) holds.

A. Let $B \in \mathcal{B}(x_0,\lambda_0)$. Consider $(x,\lambda,y) \in T(E_{(x_0,\lambda_0,B)};(x_0,\lambda_0)) \times \{y \in \mathbb{R}^p |\; y_{\hat{B}\setminus I(x_0,\lambda_0,B)} = 0\}$. There exists sequences $\varepsilon_n \downarrow 0$, $(x_n,\lambda_n) \in D_B \times \text{int } \mathbb{R}^r_+$ such that $(x_n,\lambda_n) = (x_0,\lambda_0) + \varepsilon_n(x,\lambda) + o(\varepsilon_n)$ and $\lambda_n^T R(B;x_n)_{I(x_0,\lambda_0,B)} = 0$ for all $n \geqslant 1$. Consider the sequence (y_n) defined by $y_{n\hat{B}} = y_0 + \varepsilon_n y_{\hat{B}}$, and $y_{nB} = U(B;x_n) - S(B;x_n)_{\hat{B}} y_{n\hat{B}}$. For all $i \in \hat{B} \setminus I(x_0,\lambda_0,B)$ we have $\lambda_0^T R(B;x_0)_i > 0$ (according to the definition of $I(x_0,\lambda_0,B)$), hence $y_{ni} = 0$ (since $y_{0i} = y_i = 0$), and $\lambda_n^T R(B; x_n)_i > 0$ for sufficiently large n. Let $i \in I(x_0, \lambda_0, B)$. Using Lemma 2, we have $y_{0i} > 0$, hence $y_{ni} > 0$ for sufficiently large n. Also, Lemma 2 implies $y_{nB} > 0$ for sufficiently large n. Thus, for sufficiently large n, we have $y_n \in \mathcal{E}(x_n, \lambda_n)$, hence $(x, \lambda, y) \in \mathrm{Gr}\,(D\mathcal{E}((x_0, \lambda_0), y_0))$.

Conversely, let $(x, \lambda, y) \in \mathrm{Gr}\,(D\mathcal{E}((x_0, \lambda_0), y_0))$. There exists sequences $\varepsilon_n \downarrow 0$ and $(x_n, \lambda_n, y_n) \in \mathrm{Gr}\,\mathcal{E}((x_0, \lambda_0), y_0)$ such that $(x_n, \lambda_n, y_n) = (x_0, \lambda_0, y_0) + \varepsilon(x, \lambda, y) + o(\varepsilon_n)$. Using the same argument as in the proof of (25), we can find $B \in \mathcal{B}(x_0, \lambda_0)$ such that for a subsequence (which will not be relabelled) we have $B \in \mathcal{B}(x_n, \lambda_n)$. Since $y_n \in \mathcal{E}(x_n, \lambda_n)$, we must have (31). Since $\lambda_n^T R(B; x_n)_i y_{ni} = 0$ for all $i \in \hat{B}$, $\lambda_0^T R(B; x_0)_i > 0$ for all $i \in \hat{B} \setminus I(x_0, \lambda_0, B)$ and $y_{0i} > 0$ for all $i \in I(x_0, \lambda_0, B)$, we obtain that for sufficiently large n, $\lambda_n^T R(B; x_n)_i > 0$ for all $i \in \hat{B} \setminus I(x_0, \lambda_0, B)$ and $y_{ni} > 0$ for all $i \in I(x_0, \lambda_0, B)$. Hence

$$\lambda_n^T R(B; x_n)_{I(x_0, \lambda_0, B)} = 0, \quad \text{and} \quad y_{ni} = 0 \quad \text{for all } i \in \hat{B} \setminus I(x_0, \lambda_0, B).$$

Thus, $(x_n, \lambda_n) \in E_{(x_0, \lambda_0, B)}$ and $y_{\hat{B} \setminus I(x_0, \lambda_0, B)} = 0$. Finally

$$(x, \lambda, y)) \in T(E_{(x_0, \lambda_0, B)}; (x_0, \lambda_0)) \times \{y \in \mathbb{R}^p |\, y_{\hat{B} \setminus I(x_0, \lambda_0, B)} = 0\}.$$

B. Let $B \in \mathcal{B}(x_0, \lambda_0)$ such that (28) holds. Consider the map $\Phi : (x, \lambda) \mapsto \lambda^T R(B; x)_{I(x_0, \lambda_0, B)}$ defined from $D_B \times \mathrm{int}\, \mathbb{R}_+^r$ to $\mathbb{R}^{I(x_0, \lambda_0, B)}$. Notice that $\Phi(x_0, \lambda_0) = 0$, and for all $(x, \lambda) \in X \times \mathbb{R}^r$,

$$\begin{aligned} \nabla\Phi(x_0, \lambda_0)(x, \lambda) &= \lambda^T R(B; x_0)_{I(x_0, \lambda_0, B)} + \lambda_0^T \nabla R(B; x_0)_{I(x_0, \lambda_0, B)}(x) \\ &= \begin{bmatrix} R(B; x_0)_i^T & \lambda_0^T \nabla R(B; x_0)_i \end{bmatrix}_{i \in I(x_0, \lambda_0, B)} \begin{bmatrix} \lambda \\ [x] \end{bmatrix} \end{aligned}$$

where $[x]$ stands for the column vector of the coordinates of x in the canonical basis of X. Assumption (28) implies that $\nabla\Phi(x_0, \lambda_0)$ is surjective, hence we can apply Liusternik theorem (see [1, pp. 165-167]). Thus, for each $(x, \lambda) \in \ker \nabla\Phi(x_0, \lambda_0)$, there exist $\varepsilon_0 > 0$, and a map $(\alpha, \beta) : [0, \varepsilon_0] \to X \times \mathbb{R}^r$ such that $(\alpha(\varepsilon), \beta(\varepsilon)) = o(\varepsilon)$ when $\varepsilon \downarrow 0$, and $\Phi(x_0 + \varepsilon \cdot x + \alpha(\varepsilon), \lambda_0 + \varepsilon \cdot \lambda + \beta(\varepsilon)) = 0$ for all $\varepsilon \in [0, \varepsilon_0]$.

Let $(x, \lambda, y) \in \mathcal{G}(B, x_0, \lambda_0, y_0)$. Then (20) implies that $(x, \lambda) \in \ker \nabla\Phi(x_0, \lambda_0)$, and (24) implies that $y_{\hat{B} \setminus I(x_0, \lambda_0, B)} = 0$. Hence $(x, \lambda, y) \in T(E_{(x_0, \lambda_0, B)}; (x_0, \lambda_0)) \times \{y \in \mathbb{R}^p |\, y_{\hat{B} \setminus I(x_0, \lambda_0, B)} = 0\}$. According to (27) the last set is contained in $\mathrm{Gr}\,(D\mathcal{E}((x_0, \lambda_0), y_0))$. □

Remark 2. Let $(x_0, \lambda_0) \in (\Omega \times \mathrm{int}\, \mathbb{R}_+^r) \cap \mathrm{dom}\, \mathcal{E}$, $y_0 \in \mathcal{E}(x_0, \lambda_0)$. There exists a basis $B \in \mathcal{B}(x_0, \lambda_0)$ such that $y_{0B} > 0$.

Indeed, using (13), there exist some distinct basis $B_1, \ldots, B_k \in \mathcal{B}(x_0, \lambda_0)$, and strictly positive reals $\alpha_1, \ldots, \alpha_k$, $\sum_{j=1}^k \alpha_j = 1$, such that $y_0 = \sum_{j=1}^k \alpha_j y^{B_j}$. For each B_j we have $y_{0B_j} > 0$.

Using Lemma 2, we obtain immediately the following.

Proposition 4. *Let* $(x_0, \lambda_0) \in (\Omega \times \mathrm{int}\, \mathbb{R}_+^r) \cap \mathrm{dom}\, \mathcal{E}$, $y_0 \in \mathcal{E}(x_0, \lambda_0)$, *and (according to Remark 2)* $B \in \mathcal{B}(x_0, \lambda_0)$ *such that* $y_{0B} > 0$. *Let*

$$J = \{j \in \hat{B} | y_{0j} > 0\}.$$

If

$$\exists \lambda \in \mathrm{int}\, \mathbb{R}_+^r, \qquad \lambda^T R(B; x_0)_J = 0; \quad \lambda^T R(B; x_0)_{\hat{B} \setminus J} > 0 \tag{36}$$

then

$$y_0 \in \mathrm{ri}\,(\mathcal{E}(x_0,\lambda)).$$

In this case

$$I(x_0,\lambda,B) = J.$$

Notice that if $J=\emptyset$, then $y_0 = y^B$, and (36) is equivalent to $\mathcal{E}(x_0,\lambda) = \{y^B\}$. In this case condition (28) is fulfilled (with λ instead of λ_0).

The following property may be useful for applications.

Proposition 5. *Let* $(x_0,\lambda_0) \in (\Omega \times \mathrm{int}\, R^r_+) \cap \mathrm{dom}\,\mathcal{E}$ *Then, for all* $y_0 \in \mathcal{E}(x_0,\lambda_0)$*, we have that*

$$D\mathcal{E}((x_0,\lambda_0),y_0)(0,0) = \mathrm{Cone}\;(\mathcal{E}(x_0,\lambda_0) - y_0).$$

Proof. Let $y \in D\mathcal{E}((x_0,\lambda_0),y_0)(0,0)$. It follows that $(0,0,y) \in \mathrm{Gr}\,(D\mathcal{E}((x_0,\lambda_0),y_0))$, hence from Theorem 5 we must have $(0,0,y) \in \mathcal{G}(B;x_0,\lambda_0,y_0)$ for some $B \in \mathcal{B}(x_0,\lambda_0)$. This implies that

$$A(x_0)y = 0, \quad \lambda_0^T R(B;x_0)y_{\hat{B}} = 0, \quad \text{and } y_0 + \varepsilon y \geqslant 0,$$

for sufficiently small $\varepsilon > 0$. Thus, $y_0 + \varepsilon y \in \mathcal{E}(x_0,\lambda_0)$, hence $y \in \mathrm{Cone}\;(\mathcal{E}(x_0,\lambda_0) - y_0)$.

Conversely, let $y \in \mathrm{Cone}\;(\mathcal{E}(x_0,\lambda_0) - y_0)$. Thus there exists $\varepsilon > 0$ such that $y_0 + \varepsilon y \in \mathcal{E}(x_0,\lambda_0)$, hence for every $B \in \mathcal{B}(x_0,\lambda_0)$, relations (19-23) are satisfied by $(0,0,y)$. Let $\varepsilon_n \downarrow 0$ and $(x_n,\lambda_n,y_n) = (x_0,\lambda_0,y_0+\varepsilon_n y)$. Then we have $(x_n,\lambda_n,y_n) \in \mathrm{Gr}\,\mathcal{E}$, hence $(0,0,y) \in T(\mathrm{Gr}\,(\mathcal{E});(x_0,\lambda_0,y_0))$. Therefore $y \in D\mathcal{E}((x_0,\lambda_0),y_0)(0,0)$. □

¿From now on we denote $\nabla\Phi(x_0,y_0) = [\nabla_1\Phi(x_0,y_0)\;\;\nabla_2\Phi(x_0,y_0)] \in L(X,E)\times L(\mathbb{R}^p;E)$ the Fréchet derivative (identified with a row matrix) of a map $(x,y)\mapsto\Phi(x,y)$ from $X\times\mathbb{R}^p$ to some Euclidean space E.

A straightforward consequence of Theorems 4 and 5 is the following.

Theorem 6. *Let* $(x_0,y_0) \in \Omega\times\mathbb{R}^p$ *be a minimizer for* (BLL)*. Suppose that* $y_0 \in \mathrm{ri}\,(\mathcal{E}(x_0,\lambda_0))$ *for some* $\lambda_0 \in \mathrm{int}\;\mathbb{R}^r_+$*, and condition (28) is fulfilled for some* $B \in \mathcal{B}(x_0,\lambda_0)$*. Then the following relations are satisfied :*

$$\begin{aligned} y_{0\hat{B}} &\geqslant 0,\\ y_{0B} = U(B;x_0) - S(B;x_0)_{\hat{B}} y_{0\hat{B}} &> 0,\\ \lambda_0^T R(B;x_0)_i y_{0i} &= 0 \;\; (\forall i \in \hat{B}),\\ \lambda_0^T R(B;x_0) &\geqslant 0,\\ y_{0i} &> 0 \;\; (\forall i \in I(x_0,\lambda_0,B)),\end{aligned}$$

and

$$\nabla_1 f(x_0,y_0)(x) + \nabla_2 f(x_0,y_0)(y) \geqslant 0 \;\; \textit{for all} \;\; (\lambda,x,y)\in\mathbb{R}^r\times X\times\mathbb{R}^p \;\; \textit{such that :} \qquad (37)$$

$$\begin{aligned} \lambda^T R(B;x_0)_i + \lambda_0^T(\nabla_1 R(B;x_0)(x))_i &= 0 \qquad (\forall i \in I(x_0,\lambda_0,B)), && (38)\\ y_B &= \nabla_1 U(B;x_0)(x) - \nabla_1 S(B;x_0)(x) y_{0\hat{B}} \\ &\quad - S(B;x_0) y_{\hat{B}}, && (39)\\ \lambda_0^T R(B;x_0) y_{\hat{B}} &= 0, && (40)\\ y_i &\geqslant 0 \qquad (\forall i \in \hat{B}\setminus I(x_0,\lambda_0,B)). && (41)\end{aligned}$$

Finally we obtain the following theorem which may be useful in connection with the simplex tableau used in linear programming.

Theorem 7. *Let $(x_0, y_0) \in \Omega\times\mathbb{R}^p$ be a minimizer for* (BLL)*. Suppose that $y_0 \in \mathrm{ri}\,(\mathcal{E}(x_0,\lambda_0))$ for some $\lambda_0 \in \mathrm{int}\ \mathbb{R}^r_+$, and condition (28) is fulfilled for some $B \in \mathcal{B}(x_0,\lambda_0)$. Then the following relations are satisfied :*

$$\begin{aligned}
y_{0\hat{B}} &\geqslant 0, && (42)\\
y_{0B} = U(B;x_0) - S(B;x_0)_{\hat{B}} y_{0\hat{B}} &> 0, && (43)\\
\lambda_0^T R(B;x_0)_i y_{0i} &= 0 \ (\forall i \in \hat{B}), && (44)\\
\lambda_0^T R(B;x_0) &\geqslant 0, && (45)\\
y_{0i} &> 0 \ (\forall i \in I(x_0,\lambda_0,B)), && (46)
\end{aligned}$$

and there exist vectors $\mu \in \mathbb{R}^{I(x_0,\lambda_0,B)}$, $\rho \in \mathbb{R}^B$, $\pi \in \mathbb{R}^{\hat{B}}_+$, with $\pi_i = 0$ for all $i \in I(x_0,\lambda_0,B)$, and there exists $\eta \in \mathbb{R}$, such that[vi]

$$\sum_{i\in I(x_0,\lambda_0,B)} \mu_i R(B;x_0)_i^T = 0 \qquad (47)$$

$$\begin{aligned}
\nabla_1 f(x_0,y_0) + &\sum_{i\in I(x_0,\lambda_0,B)} \mu_i \lambda_0^T \nabla_1 R(B;x_0)_i\\
&+ \sum_{j\in B} \rho_j \left(\nabla_1 U(B;x_0) - \nabla_1 S(B;x_0)^j y_{0\hat{B}}\right) = 0 \qquad (48)
\end{aligned}$$

$$\nabla_2 f(x_0,y_0) + \begin{bmatrix} -\rho^T & \eta\lambda_0^T R(B;x_0) - \sum_{j\in B} \rho_j S(B;x_0)^j + \pi^T \end{bmatrix} = 0^{(\mathrm{vii})}. \qquad (49)$$

Proof. Consider the linear maps $\Phi : \mathbb{R}^r \times X \times \mathbb{R}^p \to \mathbb{R}^{I(x_0,\lambda_0,B)} \times \mathbb{R}^B \times \mathbb{R}$ and $\Psi : \mathbb{R}^r \times X \times \mathbb{R}^p \to \mathbb{R}^{\hat{B}\setminus I(x_0,\lambda_0,B)}$ defined by

$$\begin{aligned}
\Phi(\lambda,x,y) &= \Big((\lambda^T R(B;x_0) + \lambda_0^T \nabla_1 R(B;x_0)(x))_{I(x_0,\lambda_0,B)},\\
&\qquad \nabla_1 U(B;x_0)(x) - \nabla_1 S(B;x_0)(x) y_{0\hat{B}} - S(B;x_0) y_{\hat{B}} - y_B,\ \lambda^T R(B;x_0) y_{\hat{B}} \Big)\\
\Psi(\lambda,x,y) &= -y_{\hat{B}\setminus I(x_0,\lambda_0,B)}
\end{aligned}$$

Consider the convex cone

$$K = \{(\lambda,x,y) \in \mathbb{R}^r \times X \times \mathbb{R}^p |\ \Phi(\lambda,x,y) = 0,\ \ \Psi((\lambda,x,y) \leqslant 0\}.$$

Theorem 6 implies that

$$[0_{\mathbb{R}^r}\ \ \nabla_1 f(x_0,y_0)\ \ \nabla_2 f(x_0,y_0)] \in K^-,$$

where

$$K^- = \{(l,\xi,\eta) \in \mathbb{R}^r \times X \times \mathbb{R}^p |\ \lambda^T l + [x]^T[\xi] + y^T\eta \leqslant 0,\ \ \forall(\lambda,x,y) \in K\}$$

is the (negative) polar cone. Then a standard result (see e.g. [16, Lemma 7.2.8]) implies relations (47-49). □

(vi) We denote $S(B;x_0)^j$ the j^{th} row of the matrix $S(B;x_0)$, $j \in B$.

(vii) We consider here $y_0 = [y_{0B}\ \ y_{0\hat{B}}]$ so the gradient $\nabla_2 = [\nabla_{y_{0B}}\ \ \nabla_{y_{0\hat{B}}}]$.

We end up the paper with the following example.

Example 1. Consider the problem

$$\min_{x,y}\Big((x-1)^2+(y_1-2)^2+(y_2-1)^2\Big)\quad\text{subject to}$$

$(x,y)\in(\mathbb{R}\setminus\{-1\})\times\mathbb{R}^4$, and for any $x\in]-\infty,-1[\cup]-1,+\infty[$, y is an efficient solution to

$$\text{e-}\underset{y'}{\text{MINIMIZE}}_{\mathbb{R}^2_+}\begin{bmatrix}-y'_1\\-y'_2\end{bmatrix}\quad\text{subject to}$$

$$\begin{aligned}-xy'_1+y'_2+y'_3&=1-x\\ y'_1+y'_2+y'_4&=2\\ y'_1,\,y'_2,\,y'_3,\,y'_4&\geqslant 0.\end{aligned}$$

Thus we have the particular case of problem (BBL) with $X=\mathbb{R}$, $Y=\mathbb{R}^4$, $Z=\mathbb{R}^2$, $\Omega=\mathbb{R}\setminus\{-1\}$, $f(x,y)=(x-1)^2+(y_1-2)^2+(y_2-1)^2$, $C(x)=-\begin{bmatrix}1&0\\0&1\end{bmatrix}$, $A(x)=\begin{bmatrix}-x&1&1&0\\1&1&0&1\end{bmatrix}$ and $b(x)=\begin{bmatrix}1-x\\2\end{bmatrix}$. It is a simple exercice to see[(viii)] that :

1. for each $x\geqslant 0$, the efficient set is the segment $[(2,0,0,0),\,(1,1,0,0)]$;
2. for each $x\in]-1,0[$, the efficient set is $[(2,0,0,0),\,(1,1,0,0)]\cup[(1,1,0,0),(0,1-x,0,0)]$;
3. for each $x\in]-\infty,-1[$, the efficient set is $[(0,2,0,0),\,(1,1,0,0)]\cup[(1,1,0,0),(\frac{1-x}{x},0,0,0)]$.

It is easy to see that hypotheses $(H1)$ and $(H2)$ are fulfilled.

Let us consider the basis $B=(1,2)$. We have

$$R(B;x)=\begin{bmatrix}\dfrac{-1}{1+x}&\dfrac{1}{1+x}\\[2ex]\dfrac{1}{1+x}&\dfrac{x}{1+x}\end{bmatrix},\quad S(B;x)=R(B;x),\quad U(B;x)=\begin{bmatrix}1\\[2ex]1\end{bmatrix}.$$

Since $\mathcal{E}(x_0,\lambda_0)=\mathcal{E}(x_0,\alpha\cdot\lambda_0)$ for all $(x_0,\lambda_0)\in\operatorname{dom}(\mathcal{E})$ and $\alpha\in]0,+\infty[$, we can normalise $\lambda_0\in\operatorname{int}\mathbb{R}^2_+$, so we can take $\lambda_0=[\theta_0\ \ 1-\theta_0]^T$ with $\theta_0\in]0,1[$. We have

$$\lambda_0^T R(B;x_0)=\begin{bmatrix}\dfrac{1-2\theta_0}{1+x_0}&\dfrac{\theta_0+x_0(1-\theta_0)}{1+x_0}\end{bmatrix}.$$

Also $I(x_0,\lambda_0,B)\subsetneqq\hat{B}=\{3,4\}$ (because, $I(x_0,\lambda_0,B)=\hat{B}$ implies $\mathcal{E}(x_0,\lambda_0)=S(x_0)$). Thus,

---

(viii) Notice that y_3 and y_4 are slack variables associated to a polygonal set in $\mathbb{R}^2$, so it is possible to draw a picture.

equations (43, 44, 47-49) become respectively :

$$\begin{bmatrix} y_{01} \\ \\ y_{02} \end{bmatrix} = \begin{bmatrix} 1 \\ \\ 1 \end{bmatrix} - \begin{bmatrix} \dfrac{-y_{03}+y_{04}}{1+x_0} \\ \\ \dfrac{y_{03}+x_0 y_{04}}{1+x_0} \end{bmatrix} > 0 \quad (50)$$

$$(1-2\theta_0)y_{03} = 0 \quad (51)$$

$$(\theta_0 + x_0(1-\theta_0))\, y_{04} = 0 \quad (52)$$

$$I(x_0,\lambda_0,B) = \{i_0\} \Longrightarrow \mu_{i_0} = 0 \quad (53)$$

$$2(x_0-1) + \rho_1\left(\frac{y_{03}}{(1+x_0)^2} - \frac{y_{04}}{(1+x_0)^2}\right) + \rho_2\left(-\frac{y_{03}}{(1+x_0)^2} + \frac{y_{04}}{(1+x_0)^2}\right) = 0 \quad (54)$$

$$2(y_{01}-2) - \rho_1 = 0 \quad (55)$$

$$2(y_{02}-1) - \rho_2 = 0 \quad (56)$$

$$\eta\frac{1-2\theta_0}{1+x_0} + \frac{\rho_1}{1+x_0} - \frac{\rho_2}{1+x_0} + \pi_3 = 0 \quad (57)$$

$$\eta\frac{\theta_0 + x_0(1-\theta_0)}{1+x_0} - \frac{\rho_1}{1+x_0} - \frac{\rho_2 x_0}{1+x_0} + \pi_4 = 0 \quad (58)$$

Note that $\pi_{i_0} = 0$. We have the following interesting cases :

1. $\theta_0 = 1/2$ and $\theta_0 + x_0(1-\theta_0) \neq 0$. In this case $I(x_0,\lambda_0,B) = \{3\}$, and simple computations show that the only solution is $x_0 = 1$, $y_{01} = 3/2$, $y_{02} = 1/2$, $y_{03} = 1$, $y_{04} = 0$ (which has an obvious geometrical explanation).
2. $\theta_0 \neq 1/2$ and $\theta_0 + x_0(1-\theta_0) = 0$. In this case there is no solution (satisfying also the inequalities from Theorem 7).

If $\theta_0 = 1/2$ and $\theta_0 + x_0(1-\theta_0) = 0$, then $x_0 = -1 \notin \Omega$.

In the degenerate case $\theta_0 \neq 1/2$ and $\theta_0 + x_0(1-\theta_0) \neq 0$, then $y_{03} = y_{04} = 0$, hence $y_{01} = y_{02} = 1$ and $x_0 = 1$. In this case we have $\mathcal{E}(x_0,\lambda_0) = \{(1,1,0,0)\}$ (we can consider that the conditions of Theorem 7 are satisfied with $I(x_0,\lambda_0,B) = \emptyset$, but we have less information than in the case $\theta_0 = 1/2$ when $\mathcal{E}(x_0,\lambda_0) = [(1,1,0,0),\ (2,0,0,0)]$).

References

[1] V. Alexeev, V. Tikhomirov and S. Fomine, *Commande Optimale,* Ed. Mir, Moscow, 1982.

[2] P. Armand and C. Malivert, Determination of the efficient set in multiobjective linear programming, *J. Optim. Theory Appl.* 70 (1991) 467–489.

[3] L.T.H. An, P.D. Tao and L.D. Muu, Numerical solution for optimization over the efficient set by d.c. optimization algorithms, *Oper. Res. Lett.* 19 (1996) 117–128.

[4] J.P. Aubin and I. Ekeland, *Applied Nonlinear Analysis,* John Wiley & Sons, New York, 1984.

[5] H.P. Benson, Optimization over the efficient set, *J. Math. Anal.Appl.* 98 (1984) 562–580.

[6] H.P. Benson, A finite, non-adjacent extreme point search algorithm for optimization over the efficient set, *J. Optim. Theory Appl.* 73 (1992) 47–64.

[7] M. Bergounioux and M. Haddou, A regularisation method for ill posed bilevel optimization problems, preprint CNRS, ccsd-00021999, 2006. `http://hal.ccsd.cnrs.fr`

[8] S. Bolintinéanu, Optimality conditions for Minimization over the (Weakly or Properly) Efficient Set, *J. Math. Anal. Appl.* 173: 2 (1993) 523–541.

[9] S. Bolintinéanu, Minimization of a quasi-concave function over an efficient set, *Math. Program.* 61 (1993) 89–110.

[10] S. Bolintinéanu, Necessary conditions for nonlinear suboptimization over the weakly-efficient set, *J. Optim. Theory Appl.* 78 (1993) 579–598.

[11] S. Bolintinéanu and M. El Maghri, Pénalisation dans l'optimisation sur l'ensemble faiblement efficient, *RAIRO Oper. Res.* 31: 3 (1997) 295–310.

[12] S. Bolintinéanu, Approximate efficiency and scalar stationarity in unbounded nonsmooth convex vector optimization problems, *J. Optim. Theory Appl.* 106: 2 (2000) 265–296.

[13] S. Bolintineanu, Vector variational principles; ε-efficiency and scalar stationarity, *J. Convex Anal.* 8: 1 (2001) 71–85.

[14] F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems,* Springer, 2000.

[15] H. Bonnel and J. Morgan, Semivectorial bilevel optimization problem : Penalty approach, *J. Optim. Theory Appl.* 131 (2006).

[16] J.M. Borwein and A.S. Lewis, *Convex Analysis and Nonlinear Optimization,* Springer, New York, 2000.

[17] B.D. Craven, Aspects of multicriteria optimization, in *Recent Prospects in Mathamatical Programming*, Gordon and Breach, Philadelphia, 1991.

[18] H.W. Corley, Optimality conditions for maximization of set-valued functions, *J. Optim. Theory Appl.* 58 (1988) 1–10.

[19] J.P. Dauer, Optimization over the efficient set using an active constraint approach, Z. Oper. Res. 35 (1991) 185–195.

[20] J.P. Dauer and T.A. Fosnaugh, Optimization over the efficient set, *J. Global Optim.* 7 (1995) 261–277.

[21] S. Dempe, *Foundations of Bilevel Programming,* Kluwer Academic Publishers, Dordrecht, 2002.

[22] S. Dempe, Annotated bibliography on bilevel programming and mathematical programs with equilibrium constraints, *Optimization* 52 (2003) 333–359.

[23] J. Fülöp, A Cutting Plane Algorithm for Linear Optimization over the Efficient Set, in *Generalized Convexity*, Lecture notes in Economics and Mathematical System 405, Springer-Verlag, Berlin, 1994, pp. 374–385.

[24] R. Horst and N.V. Thoai, Maximizing a concave function over the efficient or weakly-efficient set, *European J. Oper. Res.* 117 (1999) 239–252.

[25] D.T. Luc, *Theory of Vector Optimization*, Springer-Verlag, 1989.

[26] D.G. Luenberger, *Linear and Nonlinear Programming,* 2nd ed., Addison-Wesley Publishing Company, Reading, Massachusetts, 1984.

[27] B.S. Mordukhovich, *Variational Analysis and generalized differentiation II,* Springer-Verlag Berlin Heidelberg, 2006.

[28] J. Philip, Algorithms for the vector maximization problem, *Math. Program.* 2 (1972) 207–229.

[29] Y. Sawaragi, H. Nakayama and T. Tanino, *Theory of Multiobjective Optimization,* Academic Press, Orlando, 1985.

[30] Y. Yamamoto, Optimization over the efficient set: overview, *J. Global Optim.* 22 (2002) 285–317.

[31] J. Stoer and C. Witzgall, *Convexity and Optimization in Finite Dimensions I,* Springer-Verlag, New-York, 1970.

Manuscript received 26 May 2006
revised 26 June 2006
accepted for publication 26 June 2006

Henri Bonnel (former name : S. Bolintineanu)
Université de la Nouvelle-Calédonie, ERIM, BP. R4, F98851 Nouméa Cédex, New Caledonia
E-mail address: bonnel@univ-nc.nc