On first order optimality conditions for vector optimization

L. M. Graña Drummond* A. N. Iusem† B. F. Svaiter‡

May 21, 2001

Abstract

We develop first order optimality conditions for constrained vector optimization. The partial orders for the objective and the constraints are induced by closed and convex cones with nonempty interior. After presenting some well known existence results for these problems, based on a scalarization approach, we establish necessity of the optimality conditions under a Slater-like constraint qualification, and then sufficiency for the K-convex case. We present two alternative sets of optimality conditions, with the same properties in connection with necessity and sufficiency, but which are different with respect to the dimension of the spaces to which the dual multipliers belong. We introduce a duality scheme, with a point-to-set dual objective, for which strong duality holds. Some examples and open problems for future research are also presented.

Key words: Cone constraints, vector optimization, Pareto minimization, first order optimality conditions, convex programming, duality.

*Programa de Engenharia de Sistemas de Computação, COPPE–UFRJ, CP 68511, Rio de Janeiro–RJ, 21945–970, Brazil (**lmgd@cos.ufrj.br**). The work of this author was partially supported by a post-doctoral fellowship within the Department of Mathematics of the University of Haifa and by FAPERJ grant no. E-26/152.107/1999-Bolsa.

†Instituto de Matemática Pura e Aplicada (IMPA), Estrada Dona Castorina 110, Rio de Janeiro, RJ, CEP 22460-320, Brazil (**iusp@impa.br**). The work of this author was partially supported by CNPq grant no. 301280/86.

‡Instituto de Matemática Pura e Aplicada (IMPA), Estrada Dona Castorina 110, Rio de Janeiro, RJ, CEP 22460-320, Brazil (**benar@impa.br**).

1. Introduction

In this paper we study first order optimality conditions for constrained vector optimization problems. Given a closed and convex cone K in a finite dimensional real vector space V, we consider the partial order "$\preceq$" defined in V as $x \preceq y$ if and only if $y - x \in K$. Assuming that K has nonempty interior, we say that $x \prec y$ if and only if $y - x \in \text{int}(K)$. Then, given finite dimensional real vector spaces $U, V_0, \ldots, V_p$, cones $K_i \subset V_i$ $(0 \leq i \leq p)$, and functions $f : U \to V_0$, $g_i : U \to V_i$ $(1 \leq i \leq p)$, we consider the problem

$$\min f(x)$$

$$\text{s.t.} \quad g_i(x) \preceq 0, \qquad (1 \leq i \leq p)$$

where the weak Pareto minimization is understood as follows: $\bar{x}$ solves the problem if it is feasible and there exists no feasible x such $f(x) \prec f(\bar{x})$ (we recall that the definition of Pareto optimality in a general partial order "$\preceq$" requires just inexistence of a feasible x such that $f(x) \preceq f(\bar{x})$, $f(x) \neq f(\bar{x})$).

The issue of necessary and/or sufficient conditions for optimality of vector optimization problems has been considered in different settings. For instance, in the multiobjective case (i.e., when the cone is given by the nonnegative orthant), Karush-Kuhn-Tucker type conditions were studied in [19] and [14]. For the case of convex-like functions taking values in topological real vector spaces we refer to [12], where the authors establish optimality conditions for problems in which both inequality and equality constraints are considered. We point out that there is no overlapping between our work and those just mentioned, because our two sets of optimality conditions are different from the one proposed in [12], and can be seen as generalizations of those presented in [19] and [14], since, as theirs, they are Karush-Kuhn-Tucker-like conditions and hold for a wider category of cones.

Among the most popular applications of the above problem considered in the literature, we have the following two particular cases. If the objective function f is real-valued (i.e. $V_0 = \mathbb{R}$ and K_0 is the nonnegative halfline of $\mathbb{R}$), and the cone K_i consists of the symmetric positive semidefinite matrices in $\mathbb{R}^{r_i \times r_i}$ $(1 \leq i \leq p)$, then the problem reduces to a nonlinear positive semidefinite problem, which has been widely studied in recent years, particularly in connection with interior-point methods (e.g. [21]). First order optimality conditions for this problem can be found, for instance, in [17], [18].

The second particular case considers vector-valued objective functions, but with $V_0 = \mathbb{R}^n$ and K_0 consisting of the nonnegative orthant of $\mathbb{R}^n$ (i.e. multicriteria optimization). We remark that if the constraint cones K_i, for $i \geq 1$, are also the nonnegative orthant of, say, $\mathbb{R}^{r_i}$, then each constraint $g_i(x) \preceq 0$ reduces to r_i real-valued inequalities (this is not the case for the cone of symmetric positive semidefinite matrices in $\mathbb{R}^{r_i \times r_i}$, for which the constraint $g_i(x) \preceq 0$ is equivalent to the following infinite set of real-valued constraints: $u^t g_i(x) u \leq 0$ for each $u \in \mathbb{R}^{r_i}$). With nonnegative orthants as cones it makes little sense to consider explicitly vector-valued constraints, and thus the literature on Pareto optimization problems considers only real-valued constraints (e.g. Chapter VI of [3], [8], [14], [15], [22], [23]). In some of these references the definition of optimality is the usual Pareto one,

mentioned above, so that some additional requirements are imposed (e.g., "proper efficiency" in [8]) for obtaining appropriate first order conditions. Other references, e.g. [3], deal with weak Pareto optima in terms of "$\prec$", as we do, making such requirements unnecessary. We remark that when the constraints are real-valued, the multipliers in the first order optimality conditions are scalars, as in the standard Karush-Kuhn-Tucker conditions for nonlinear programming. With vector-valued constraints the issue of the right space for the multipliers becomes nontrivial (e.g. Theorem 2 and Proposition 8 below, and also [9], where the same problem with a different definition of minimizers of vector-valued functions is analyzed).

We present our work in a self contained way: many of the preliminary results are in fact well known (and some of them even hold in more general settings) but, for the sake of completeness, we state each and every result that we need, quoting references for proofs and, in some cases, giving our own alternative proofs.

The paper is organized as follows: in section 2 we present a brief review of some elementary results related to the definition of "$\prec$", and we also recall some basic properties of convexity related to cones. In section 3 we introduce some standard existence results for weak Pareto optimization. We present a classical scalarization result, which guarantees that in most cases weak Pareto minimizers can be found by solving suitably chosen real-valued optimization problems. This reduction, however, is not fully constructive, and thus it is more useful for obtaining existence results than for algorithmic purposes. In this context, as pointed out in section 7, our first order optimality conditions open the road for primal-dual methods for weak Pareto optimization which work directly with the vector-valued problem, without reducing it to a (possibly unbounded) problem with real-valued objective. In section 4 we present our first order optimality conditions. For the sake of clarity and completeness, we analyze first the well known unconstrained case and then the constrained case, though the first one, of course, can be seen as a particular case of the second one. We establish necessity of the conditions, under a standard Slater-like constraint qualification and convexity of the vector-valued constraints respect to (fixed) cones. We prove then that these conditions are sufficient when both the objective and the constraint functions are convex with respect to cones in their codomains. As a by-product of our optimality conditions, we establish a conical version of Farkas' Lemma, of some interest on its own. In these results, we have $p+1$ dual multipliers (for the problem with p vector-valued constraints), which are elements of V_i $(0 \leq i \leq p)$, i.e. of the codomains of the objective and the constraint functions. We present also an alternative formulation with only p multipliers, but then each one of them is a linear transformation from V_i to V_0 (i.e. from the codomain of each constraint function to the codomain of the objective function). In section 5 we use our result on necessity of the first order conditions to develop a duality scheme for the convex version of the problem, with a point-to-set mapping as the dual objective, and both weak and strong duality results. Section 6 contains two illustrative examples, and some suggestions for future research are commented upon in section 7.

Throughout the paper, $\mathbb{R}^n_+$ denotes the nonnegative orthant of $\mathbb{R}^n$, $\mathbb{R}^n_{++}$ the open positive orthant of $\mathbb{R}^n$, S^n the subspace of symmetric matrices in $\mathbb{R}^{n\times n}$, S^n_+ the cone of symmetric positive semidefinite matrices in $\mathbb{R}^{n\times n}$ and S^n_{++} the open cone of symmetric positive definite matrices in

$\mathbb{R}^{n\times n}$.

2. Preliminary results

Let V be a finite dimensional real vector space with an inner product $\langle\cdot,\cdot\rangle$, $K \subset V$ a closed and convex cone and K^* the positive dual of K (i.e. $K^* = \{z \in V : \langle z, x\rangle \geq 0 \text{ for all } x \in K\}$). We assume from now on that $\text{int}(K) \neq \emptyset$ and that $\text{int}(K^*) \neq \emptyset$. We consider the partial order $\preceq$ in V induced by K, given by $x \preceq y$ (or equivalently, $y \succeq x$) if and only if $y - x \in K$. We say that $x \prec y$ (or equivalently, $y \succ x$) if and only if $y - x \in \text{int}(K)$. The next proposition presents some extremely elementary properties of "$\prec$", which are very natural and whose proofs may be found in almost any vector optimization book. Throughout this work we will use these results without quoting them specifically.

Proposition 1. *i) There exists no $x \in V$ such that $x \prec x$.*

ii) For all x, $y \in V$ and all $t \in \mathbb{R}_{++}$, if $x \prec y$ then $tx \prec ty$.

iii) For all w, x, y, $z \in V$, if $w \prec x$ and $y \preceq z$ then $w + y \prec x + z$.

iv) For all a, b, $c \in V$, if $a \prec b$ and $b \preceq c$ then $a \prec c$.

The next proposition presents a well known property of "$\prec$" in connection with differentiable functions. From now on U will be a finite dimensional vector space.

Proposition 2. *Take a differentiable function $f : U \to V$ and x, $y \in U$. If $df(x)y \succ 0$ then there exists $\bar{t} \in (0, 1]$ such that $f(x - ty) \prec f(x)$ for all $t \in (0, \bar{t}\,]$.*

We recall that a suitable extension to the vector-valued case of the classical notion of convexity of scalar functions is given by the following definition. A function $f : U \to V$ is said to be *K-convex* if

$$f(\alpha x + (1-\alpha)y) \preceq \alpha f(x) + (1-\alpha)f(y)$$

for all x, $y \in U$ and all $\alpha \in [0, 1]$. These functions are analyzed in detail in [13]. Other vector extensions of convexity are studied in [7] and [2]. We need in the sequel the following properties of K-convex functions.

Proposition 3. *i) The function f is K-convex if and only if the real-valued function $\psi_z(x) = \langle z, f(x)\rangle$ is convex for all $z \in K^*$.*

ii) If f_i is K-convex $(1 \leq i \leq m)$ then $f(x) = \sum_{i=1}^m f_i(x)$ is K-convex.

iii) If f is K-convex and differentiable then, for all x, $y \in U$, it holds that

$$f(x) \succeq f(y) + df(y)(x - y).$$

iv) If f is K-convex and differentiable, then it is continuously differentiable.

Proof. The proofs of (i),(ii) and (iii) can be found in [13], Propositions 6.2., 6.7 and Lemma 5.2, respectively. Observe that (iii) can be seen as a generalization to vector-valued functions of the classical gradient inequality for real-valued functions.
iv) Let $m = \dim(V)$. Since $\mathrm{int}(K^*)$ is nonempty, it contains a basis $\{v^j\}$ $(1 \leq j \leq m)$ of V. Since f is K-convex, by Proposition 3, the functions $\psi_j(x) = \langle f(x), v^j \rangle$ $(1 \leq j \leq m)$ are convex, and they are obviously differentiable, with $d\psi_j(\cdot) = \langle df(\cdot), v^j \rangle$. Since differentiable convex functions are continuously differentiable (see Theorem 7.8 in [20]), we conclude that $\langle df(\cdot), v^j \rangle$ is continuous for all j. Let $\{e^i\}$ $(1 \leq i \leq m)$ be an orthonormal basis of V, and take α_{ij} $(1 \leq i, j \leq m)$ such that $e^i = \sum_{j=1}^m \alpha_{ij} v^j$ $(1 \leq i \leq m)$. Then

$$df(\cdot) = \sum_{i=1}^m \langle df(\cdot), e^i \rangle e^i = \sum_{i=1}^m \sum_{j=1}^m \alpha_{ij} \langle df(\cdot), v^j \rangle e^i. \tag{1}$$

Since $\langle df(\cdot), v^j \rangle$ is continuous for all j, it follows that $\langle df(\cdot), v^j \rangle e^i$ is continuous for all i, j, and therefore $df(\cdot)$ is continuous, being a linear combination of continuous functions, in view of (1). □

3. Review of results on existence of Pareto minimal elements and scalarization

In this section we review some results on existence of Pareto minimizers and its connection with the scalarization procedure for finding such minimizers. Though they are not esential for our presentation of first order optimality conditions, they provide a more comprehensive background for them. For the sake of completeness, we include the proofs of these results.

We recall that for a given a subset A of V, $\bar{v}$ is said to be a *weak Pareto minimal element* (or *weak efficient element*) of A if $\bar{v} \in A$ and there exists no $v \in A$ such that $v \prec \bar{v}$.

In a general partial order "$\preceq$" defined on V, i.e. not necessarily induced by a cone K, $\bar{v}$ is said to be a *Pareto minimal element* (or *efficient element*) of A if and only if $\bar{v} \in A$ and there exists no $v \in A$ such that $v \preceq \bar{v}$, $v \neq \bar{v}$.

Clearly, for partial orders "$\preceq$" and "$\prec$" induced by a cone $K \subset V$ and its interior $\mathrm{int}(K)$, respectively, it holds that a Pareto minimal element of a set is a weak Pareto minimal element of that set.

In order to study existence of weak Pareto minimal elements, we need the following elementary result.

Proposition 4. *i) $int(K^*) = \{w \in V : \langle w, v \rangle > 0 \ \forall \ 0 \neq v \in K\}$.*

ii) $int(K^*) \cap K \neq \emptyset$.

Proof. i) If $w \in \text{int}(K^*)$ then $\langle w, v\rangle \geq 0$ for all $v \in K$ by definition of K^*, and we cannot have $\langle w, v\rangle = 0$ for some nonnull $v \in K$ because in such a case $w' = w - \alpha v$ belongs to K^* for $\alpha \in \mathbb{R}_{++}$ small enough, and then

$$\langle w', v\rangle = \langle w, v\rangle - \alpha||v||^2 = -\alpha||v||^2 < 0,$$

contradicting the definition of K^*. If $\langle w, v\rangle > 0$ for all $0 \neq v \in K$, then $\langle w, \cdot\rangle$ attains its minimum on the compact intersection of K with the unit sphere of V, with a positive value, say β. Is is easy to check that $\langle u, v\rangle \geq 0$ for all $v \in K$ and all u in the ball with center at w and radius β. Thus, this ball is contained in K^* and therefore $w \in \text{int}(K^*)$.
ii) If the result does not hold, then, by the nonstrict Convex Separation Theorem (see [1], Theorem 4.7) there exists $0 \neq w \in V$ such that

$$\langle w, v\rangle \geq 0 \ \ \forall\ v \in K, \tag{2}$$

$$\langle w, v\rangle \leq 0 \ \ \forall\ v \in \ \text{int}(K^*). \tag{3}$$

By (2), $w \in K^*$. By continuity of $\langle w, \cdot\rangle$, (3) holds for all $v \in K^*$, including w itself, so that we get $||w||^2 = \langle w, w\rangle \leq 0$, in contradiction with $w \neq 0$. □

We now study some existence criteria. We will follow the well known scalarization approach (see, for instance, [8]). For a given subset A of V we consider the following scalar problem, which we wil call (RV):

$$\begin{aligned} &\min \ \langle w, v\rangle \\ &\text{s.t.} \ \ v \in A. \end{aligned}$$

We wil denote by $\text{co}(A)$ the closed convex hull of A, and by A_∞ the recession cone of $\text{co}(A)$, i.e. the set of directions of halflines contained in $\text{co}(A)$.

Proposition 5. *i) If $w \in int(K^*)$ and $\bar{v} \in V$ solves problem (RV), then $\bar{v}$ is a weak Pareto minimal element of A.*

ii) If A is convex and $\bar{v}$ is a weak Pareto minimal element of A, then there exists $w \in K^*$ such that $\bar{v}$ is a solution of min $\langle w, v\rangle$ subject to $v \in A$.

iii) If A is closed and w belongs to $int[(A_\infty)^*]$, then the problem min $\langle w, v\rangle$ subject to $v \in A$ has solutions.

iv) If A is closed and $int(K^*) \cap int[(A_\infty)^*] \neq \emptyset$ then there exist weak Pareto minimal elements of A.

Proof. i) If the result does not hold, then there exists $v \in A$ such that $\bar{v} - v \in \text{int}(K)$. Thus $\bar{v} - v \neq 0$, and, since $\bar{v}$ solves (RV), $\langle w, \bar{v}\rangle \leq \langle w, v\rangle$, implying $\langle w, \bar{v} - v\rangle \leq 0$. Since $w \in \text{int}(K^*)$, we have a contradiction, in view of Proposition 4(i).
ii) This result can be seen as a sort of converse of (i), namely it says that each weak Pareto minimal element of A minimizes a linear real-valued functional in A. By our hypothesis, $\bar{v} - v \notin \text{int}(K)$ for all $v \in A$, i.e. $(\bar{v} - A) \cap \text{int}(K) = \emptyset$. Since $\bar{v} - A$ is convex by convexity of A, the nonstrict version of the Convex Separation Theorem (see [1], Theorem 4.7) implies that there exist $w \in V$ such that

$$\langle w, \bar{v} - v\rangle \leq 0 \qquad \forall\ v \in A, \tag{4}$$

$$\langle w, v\rangle \geq 0 \qquad \forall\ v \in \text{int}(K). \tag{5}$$

By (4), $\bar{v}$ is a solution of min $\langle w, v\rangle$ subject to $v \in A$, and it follows easily from (5) that $w \in K^*$.
iii) Take $v_0 \in A$. We claim that $L(v_0) := \{v \in \text{co}(A) : \langle w, v\rangle \leq \langle w, v_0\rangle\}$ is compact. It is certainly closed and convex, and if it were unbounded it would contain a halfline with direction $0 \neq y \in A_\infty$. Then, taking $v = v_0 + y \in \text{co}(A)$, we get easily that $\langle w, y\rangle \leq 0$, which, in view of Proposition 4(i), contradicts the fact that $w \in \text{int}[(A_\infty)^*]$. Thus, the claim holds. Now, since $A \subset \text{co}(A)$, we have that $\hat{L}(v_0) := \{v \in A : \langle w, v\rangle \leq \langle w, v_0\rangle\} \subset L(v_0)$. Thus $\hat{L}(v_0)$ is bounded, and also closed because A is closed, henceforth compact. It follows that the continuous function $\langle w, \cdot\rangle$ attains its minimum in A.
iv) Take $w \in \text{int}(K^*) \cap \text{int}[(A_\infty)^*]$. By (iii), problem (RV) with such a w has solutions, and by (i) such solutions are weak Pareto minimal elements of A. □

We point out that some of the preceding results also hold if we replace "weak Pareto minimal elements" by "Pareto minimal elements" (where, of course, the Pareto order "$\preceq$" is induced by a cone like those we are considering in this work).

Given $f : U \to V$, we say that $\bar{x} \in U$ is a (weak) Pareto minimizer of f if $f(\bar{x})$ is a (weak) Pareto minimal element of $\text{Im}(f)$.

We say that f is *K-coercive* if there exists $w \in \text{int}(K^*)$ such that $\varphi : U \to \mathbb{R}$ defined as $\varphi(x) = \langle w, f(x)\rangle$ is coercive, i.e. if $lim_{||x|| \to \infty} \langle w, f(x)\rangle = +\infty$.

We have the following K-coercivity criterion for existence of weak Pareto minimizers of f.

Proposition 6. *If $f : U \to V$ is K-coercive then there exist weak Pareto minimizers of f.*

Proof. By Proposition 5(i) with $A = \text{Im}(f)$, it suffices to exhibit $w \in \text{int}(K^*)$ such that the function $\varphi(x) = \langle w, f(x)\rangle$ has minimizers. But this is the case, because φ is coercive, in view of the definition of K-coerciveness. □

4. First order optimality conditions for vector minimization problems

We give next the first order necessary optimality condition for unconstrained Pareto minimizers, which is also sufficient if f is K-convex. This necessary condition is quite folkloric in vector optimization and its proof can be found in almost any introductory book on the subject (actually, it follows immediately from Propositions 1 and 2). The sufficiency for the K-convex case is just a simple exercise, which follows easily from Propositions 1 and 3(iii).

Theorem 1. *Let $f : U \to V$ be differentiable.*

i) If $\bar{x} \in U$ is a Pareto minimizer of f then $\bar{x}$ is a K-critical point, i.e. $Im(df(\bar{x})) \cap int(K) = \emptyset$.

ii) If f is K-convex and $\bar{x} \in U$ is such that $Im(df(\bar{x})) \cap int(K) = \emptyset$, then $\bar{x}$ is a weak Pareto minimizer of f.

Next we analyze the case of constrained weak Pareto minimizers. Take finite dimensional real vector spaces $U, V_0, \ldots, V_p$ with inner products, and closed and convex cones $K_i \subset V_i$ such that $\text{int}(K_i) \neq \emptyset$, $\text{int}(K_i^*) \neq \emptyset$ $(0 \leq i \leq p)$. In each space V_i we consider the partial order induced by K_i. We will denote all these partial orders with the same symbol "$\preceq$", in the hope that the involved space and cone will be clear from the context.

Take now $f : U \to V_0$, $g_i : U \to V_i$ $(1 \leq i \leq p)$ and consider the problem of finding a constrained weak Pareto minimizer of f:

$$\min f(x) \tag{6}$$

$$\text{s.t.} \quad g_i(x) \preceq 0 \qquad (1 \leq i \leq p). \tag{7}$$

We recall the meaning of (6)–(7), which we call Problem (P). Let $F \subset U$ be the the feasible set for (P), i.e. $F = \{x \in U : -g_i(x) \in K_i \ (1 \leq i \leq p)\}$. Then, $\bar{x} \in U$ solves (P) if and only if $\bar{x} \in F$ and there exists no $x \in F$ such that $f(\bar{x}) - f(x) \in \text{int}(K_0)$.

We have now the following elementary result concerning the existence of solutions of (P)

Proposition 7. *Consider (P) with f and all the g_i's continuous. If either F is compact or f is K_0-coercive, then (P) has solutions.*

Proof. If F is compact, then $f(F)$ is compact by continuity of f, and it is immediate that $\text{co}(f(F))$ is compact. It follows that $f(F)_\infty = \{0\}$, so that $[f(F)_\infty]^* = V_0$. Since $\text{int}(K_0^*) \neq \emptyset$, the hypothesis of Proposition 5(iv) is satisfied with $A = f(F)$, and the result follows.

If f is K_0-coercive, take $w \in \text{int}(K^*)$, whose existence is guaranteed by definition of K_0-coerciveness, so that $\langle w, f(\cdot)\rangle$ is coercive, and therefore it attains its minimum on F, which is closed by continuity of the g_i's. Since $w \in \text{int}(K_0^*)$, we conclude from Proposition 5(i) that $f(F)$ has weak Pareto minimal elements, which evidently are solutions of (P). □

In connection with Proposition 5(ii), we mention that $f(A) \subset V_0$ may fail to be a convex set even when $f : U \to V_0$ is K_0-convex and $A \subset U$ is convex. Take, for instance, $A = U = \mathbb{R}$, $V_0 = S^2$, $K_0 = S^2_+$ and $f(t) = \begin{pmatrix} t & 0 \\ 0 & t^2 \end{pmatrix}$. It is easy to check that f is S^2_+-convex, but $\mathrm{Im}(f) \subset S^2$ is not convex: the midpoint between $f(1)$ and $f(2)$ is $\begin{pmatrix} 3/2 & 0 \\ 0 & 5/2 \end{pmatrix}$, which does not belong to $\mathrm{Im}(f)$. Thus, even when f is K_0-convex and all the g_i's are K_i-convex, we cannot ensure that each one of the solutions of (P) can be found by minimizing on F a real-valued function of the form $\langle w, f(x)\rangle$, with $w \in K_0^*$.

Next we present the first order optimality conditions for problem (P).

We say that $\bar{x} \in U$ and $z_i \in V_i$ $(0 \leq i \leq p)$ satisfy the first order optimality conditions for (P) if:

$$\mathrm{Im}\left[df(\bar{x}) + \left(\sum_{i=1}^{p} \langle z_i, dg_i(\bar{x})\rangle\right) z_0\right] \cap \mathrm{int}(K_0) = \emptyset \qquad \text{Lagrangian condition,} \tag{8}$$

$$g_i(\bar{x}) \preceq 0 \quad (1 \leq i \leq p) \qquad \text{primal feasibility,} \tag{9}$$

$$z_0 \in K_0 \cap K_0^*, \qquad z_i \in K_i^* \ \ (1 \leq i \leq p) \qquad \text{dual feasibility,} \tag{10}$$

$$\langle z_i, g_i(\bar{x})\rangle = 0 \quad (1 \leq i \leq p) \qquad \text{complementarity,} \tag{11}$$

We clarify the meaning of (8): $df(\bar{x}) + \left(\sum_{i=1}^{p} \langle z_i, dg_i(\bar{x})\rangle\right) z_0$ is the linear transformation from U to V_0 which maps $y \in U$ to $df(\bar{x})y + \left(\sum_{i=1}^{p} \langle z_i, dg_i(\bar{x})y\rangle\right) z_0 \in V_0$. We also mention that (10) can be seen as a nonnegativity condition in the z_i's, considered as elements of the dual spaces of the V_i's, because it means that $\langle z_i, v_i\rangle \geq 0$ for all $v_i \succeq 0$.

We mention that for the particular case of $V_0 = \mathbb{R}^n$, $K_0 = \mathbb{R}^n_+$, $V_i = \mathbb{R}$ $(1 \leq i \leq p)$, (i.e., real-valued constraints) sets of conditions related to (8)–(11) have been presented, e.g. in Chapter VI of [3], [8], [23]. Since the constraints in all these references are real-vualed, the multipliers z_i are scalars.

In order to establish necessity of (8)–(11), we need the following lemma, of some interest on its own. Its first conclusion can be seen as a conical version of Farkas' Lemma, and is in fact a particular case of some more general alternative theorems; we decided to give our own proof because we believe it gives some insight on the subject. The second one is a variant which fits precisely our needs for the necessity result, both in the assumptions and in the conclusions. For other alternative/Farkas type theorems for convex-like functions we refer to [12] and [7].

Lemma 1. *Take finite dimensional real vector spaces with inner product $V_0, \ldots, V_p$, closed and convex cones $K_i \subset V_i$ $(0 \leq i \leq p)$, linear transformations $h_i : U \to V_i$ $(0 \leq i \leq p)$ and vectors $b_i \in V_i$ $(0 \leq i \leq p)$. Then,*

i) If there exists no $y \in U$ such that

$$-b_i + h_i(y) \in K_i \qquad (0 \leq i \leq p), \tag{12}$$

then there exists $z_i \in K_i^*$ $(0 \leq i \leq p)$ such that

$$\sum_{i=0}^{p} \langle z_i, b_i \rangle \geq 0, \tag{13}$$

$$\sum_{i=0}^{p} \langle z_i, h_i(y) \rangle = 0 \tag{14}$$

for all $y \in U$.

ii) If $b_0 = 0$, $-b_i \in K_i$ $(1 \leq i \leq p)$, there exists no $y \in U$ such that

$$h_0(y) \in \mathrm{int}(K_0), \tag{15}$$

$$-b_i + h_i(y) \in K_i \quad (1 \leq i \leq p), \tag{16}$$

but there exists $\bar{y} \in U$ such that

$$-b_i + h_i(\bar{y}) \in \mathrm{int}(K_i) \qquad (1 \leq i \leq p), \tag{17}$$

then there exists $z_i \in V_i$ $(0 \leq i \leq p)$ such that

$$z_0 \in K_0 \cap K_0^*, \qquad z_i \in K_i^* \ \ (1 \leq i \leq p), \tag{18}$$

$$\langle z_i, b_i \rangle = 0 \qquad (1 \leq i \leq p), \tag{19}$$

$$h_0(y) + \left(\sum_{i=1}^{p} \langle z_i, h_i(y) \rangle \right) z_0 \notin \mathrm{int}(K_0) \quad \forall\ y \in U. \tag{20}$$

Proof. i) Define $\bar{h} : U \to V_0 \times V_1 \times \cdots \times V_p$ and $K \subset V_0 \times V_1 \times \cdots \times V_p$ as

$$\bar{h}(y) = (-b_0 + h_0(y), \ldots, -b_p + h_p(y)),$$

$$K = K_0 \times \cdots \times K_p.$$

Note that $\bar{h}$ is affine, so that its image is closed and convex, and K is a cone. Asssumption (12) means precisely that $\mathrm{Im}(\bar{h}) \cap K = \emptyset$. Thus we can apply the nonstrict version of the Convex Separation Theorem (see [1], Theorem 4.7), and conclude that there exists $(z_0, \ldots, z_p) \in V_0 \times \cdots \times V_p$ such that

$$\sum_{i=0}^{p} \langle z_i, -b_i + h_i(y) \rangle \leq 0 \quad \forall\ y \in U, \tag{21}$$

$$\sum_{i=0}^{p} \langle z_i, v_i \rangle \geq 0 \quad \forall\ (v_0, v_1, \ldots v_p) \in K. \tag{22}$$

Fix $i \in \{0, \ldots, p\}$, and look at (22) with $v_j = 0$ for $j \in \{0, \ldots, p\}$, $j \neq i$. It follows that $\langle z_i, v_i \rangle \geq 0$ for all $v_i \in K_i$, so that

$$z_i \in K_i^* \qquad (0 \leq i \leq p). \tag{23}$$

Taking $y = 0$ in (21), we get, since the h_i's are linear, $\sum_{i=1}^{p} \langle z_i, b_i \rangle \geq 0$, establishing (13). Now rewrite (21) as

$$-\sum_{i=0}^{p} \langle z_i, b_i \rangle + \sum_{i=0}^{p} \langle z_i, h_i(y) \rangle \leq 0 \ \ \forall y \in U. \tag{24}$$

If (14) does not hold, then there exists $\hat{y} \in U$ such that $\sum_{i=0}^{p} \langle z_i, h_i(\hat{y}) \rangle \neq 0$. Assume that $\sum_{i=0}^{p} \langle z_i, h_i(\hat{y}) \rangle > 0$. Looking at (24) with $y = t\hat{y}$ ($t \in \mathbb{R}$) and letting t go to $+\infty$, we get that the second term in the left hand side goes to $+\infty$ by linearity of the h_i's, while the first term does not depend on t, contradicting the inequality. If $\sum_{i=0}^{p} \langle z_i, h_i(\hat{y}) \rangle < 0$, we achieve a similar contradiction by letting now t go to $-\infty$. We conclude that (14) holds.
ii) Now we consider $\bar{h}$ as in (i) and $K = \mathrm{int}(K_0) \times K_1 \times \ldots, K_p$. K is now a (nonclosed) convex cone, and the nonstrict Convex Separation Theorem still applies, since $\mathrm{Im}(\bar{h}) \cap K = \emptyset$ by (15) and (16). If we call now w_0 the vector which was called z_0 in (i), we get that (21) and (22) hold in this case, with w_0 substituting for z_0. Condition (23) also holds, noting that now we cannot take $v_0 = 0$, which does not belong to $\mathrm{int}(K_0)$, but we can let $v_0 \in \mathrm{int}(K_0)$ approach 0. By the same token, (24) holds too, with the summation extending from 1 to p, because $b_0 = 0$. Since $z_i \in K_i^*$

by (23), and $-b_i \in K_i$, we have that $\langle z_i, b_i \rangle \leq 0$ $(1 \leq i \leq p)$, and therefore we conclude from (22) that

$$\langle z_i, b_i \rangle = 0 \qquad (1 \leq i \leq p), \tag{25}$$

establishing (19). We claim now that

$$w_0 \neq 0. \tag{26}$$

We proceed to establish (26). Let $\bar{h}_i(y) = -b_i + h_i(y)$. If $w_0 = 0$, then (21) and (22) become

$$\sum_{i=1}^{p} \langle z_i, \bar{h}_i(y) \rangle \leq 0 \quad \forall\ y \in U,$$

$$\sum_{i=1}^{p} \langle z_i, v_i \rangle \geq 0 \quad \forall\ (v_1, \ldots, v_p) \in K_1 \times \cdots \times K_p,$$

meaning that $(z_1, \ldots, z_p)$ is a nonstrict separator between the convex sets $\mathrm{Im}(\bar{h}_1) \times \cdots \times \mathrm{Im}(\bar{h}_p)$ and $K_1 \times \cdots \times K_p$. It follows that the interior of the second set does not intersect the first one, and therefore

$$[\mathrm{Im}(\bar{h}_1) \times \cdots \times \mathrm{Im}(\bar{h}_p)] \cap [\mathrm{int}(K_1) \times \cdots \times \mathrm{int}(K_p)] = \emptyset,$$

which contradicts (17), thus proving that $w_0 \neq 0$.

Take now $\hat{z}_0 \in K_0 \cap \mathrm{int}(K_0^*)$, which exists by Proposition 4(ii). Since $0 \neq w_0 \in K_0^*$ by (23) with w_0 instead of z_0 and (26), it follows from Proposition 4(i) that $\langle w_0, \hat{z}_0 \rangle > 0$. Define then $z_0 \in V_0$ as

$$z_0 = \frac{\hat{z}_0}{\langle w_0, \hat{z}_0 \rangle}. \tag{27}$$

Since $\hat{z}_0 \in K_0 \cap \mathrm{int}(K_0^*)$ and $\langle w_0, \hat{z}_0 \rangle > 0$, it follows from (27) that $z_0 \in K_0 \cap K_0^*$, so that (18) holds, in view of (23). In order to establish (20), note that, for all $y \in U$,

$$\left\langle w_0, h_0(y) + \left(\sum_{i=1}^{p} \langle z_i, h_i(y) \rangle \right) z^0 \right\rangle = \langle w_0, h_0(y) \rangle + \sum_{i=1}^{p} \langle z_i, h_i(y) \rangle \langle w_0, z_0 \rangle =$$

$$\langle w_0, h_0(y) \rangle + \sum_{i=1}^{p} \langle z_i, h_i(y) \rangle \leq 0, \tag{28}$$

using (27) in the second equality and (21), (19) in the inequality. On the other hand, since $K_0 = K_0^{**}$ (e.g. Theorem 14.1 of [16]), it follows from Proposition 4(i) that $\mathrm{int}(K_0) = \mathrm{int}(K_0^{**}) = \{u \in V_0 : \langle u, v_0 \rangle > 0\ \forall 0 \neq v_0 \in K_0^*\}$. In view of (23) and (26), we conclude from (28) that $h_0(y) + \left(\sum_{i=1}^{p} \langle z_i, h_i(y) \rangle \right) z_0 \notin \mathrm{int}(K_0)$, establishing (20). □

Necessity of (8)–(11), under K_i-convexity of the constraint functions g_i and a Slater-type constraint qualification, is established in the next theorem.

Theorem 2. *Consider (P) and assume that*

i) f is differentiable.

ii) g_i is K_i-convex and differentiable ($1 \leq i \leq p$).

iii) (Slater's condition) There exists $\hat{x} \in U$ such that $g_i(\hat{x}) \prec 0$ ($1 \leq i \leq p$).

If $\bar{x} \in U$ solves (P) then there exist $z_i \in V_i$ ($0 \leq i \leq p$) such that (8)–(11) hold.

Proof. Define $\bar{y} = \bar{x} - \hat{x}$, with $\hat{x}$ as in (iii). By (iii) and Proposition 3(iii) applied to the K_i-convex function g_i, we have that $g_i(\bar{x}) - dg_i(\bar{x})\bar{y} \preceq g_i(\hat{x}) \prec 0$, implying

$$-g_i(\bar{x}) + dg_i(\bar{x})\bar{y} \succ 0 \qquad (1 \leq i \leq p). \tag{29}$$

We claim now that there exists no $y \in U$ such that

$$df(\bar{x})y \in \mathrm{int}(K_0), \tag{30}$$

$$-g_i(\bar{x}) + dg_i(\bar{x})y \in K_i \qquad (1 \leq i \leq p). \tag{31}$$

We proceed to establish the claim by contradiction. Suppose that there exists $y \in U$ such that (30) and (31) hold. By (30), there exists $\theta \in (0, 1]$ such that $df(\bar{x})[y + \theta(\bar{y} - y)] \succ 0$. Defining $\hat{y} = y + \theta(\bar{y} - y)$, we have that

$$df(\bar{x})\hat{y} \succ 0. \tag{32}$$

By (32), assumption (i) and Proposition 2, there exists $\bar{t} \in (0, 1]$ such that

$$f(\bar{x} - t\hat{y}) \prec f(\bar{x}) \tag{33}$$

for all $t \in (0, \bar{t}]$. From (29) and (31) we get

$$-g_i(\bar{x}) + dg_i(\bar{x})\hat{y} = \theta[-g_i(\bar{x}) + dg_i(\bar{x})\bar{y}] + (1-\theta)[-g_i(\bar{x}) + dg_i(\bar{x})y] \succ 0 \quad (1 \leq i \leq p). \tag{34}$$

By assumption (ii) and Proposition 3(iv), all the g_i's are continuously differentiable. Thus, it follows from (34) that there exists $\tilde{t} \in (0, \bar{t}] \subset (0, 1]$ such that

$$-g_i(\bar{x}) + dg_i(\bar{x} - \tilde{t}\hat{y})\hat{y} \succ 0 \quad (1 \leq i \leq p), \tag{35}$$

By Proposition 3(ii) applied to the K_i-convex function g_i,

$$g_i(\bar{x}) \succeq g_i(\bar{x} - \tilde{t}\hat{y}) + \tilde{t}dg_i(\bar{x} - \tilde{t}\hat{y})\hat{y}. \tag{36}$$

By (36), for all $i \in \{1, \dots p\}$,

$$g_i(\bar{x} - \tilde{t}\hat{y}) \preceq g_i(\bar{x}) - \tilde{t}dg_i(\bar{x} - \tilde{t}\hat{y})\hat{y} = (1 - \tilde{t})g_i(\bar{x}) - \tilde{t}[-g_i(\bar{x}) + dg_i(\bar{x} - \tilde{t}\hat{y})\hat{y}] \preceq 0, \tag{37}$$

using (9), (35) and the fact that $\tilde{t} \in (0, 1]$ in the second inequality. Since $\tilde{t} \in (0, \bar{t}\,]$, (33) and (37) contradict the fact that $\bar{x}$ solves (P), and thus the claim holds, i.e. no $y \in U$ satisfies (30) and (31). Now we apply Lemma 1(ii) with $h_0 = df(\bar{x})$, $h_i = dg_i(\bar{x})$ and $b_i = g_i(\bar{x})$ $(1 \leq i \leq p)$. Note that $-b_i \in K_i$ because $\bar{x}$, being a solution of (P), belongs to F. Conditions (15) and (16) in Lemma 1(ii) are satisfied by the claim just established, and (17) holds by (29). Now conclusions (18), (19) and (20) establish conditions (10), (11) and (8) respectively. Condition (9) holds because, as mentioned above, $\bar{x}$ belongs to F. Thus (8)–(11) hold and the thorem is proved. □

We conjecture that Theorem 2 holds without K_i-convexity and differentiability of the g_i's, after replacing assumption (iii) by an appropriate constraint qualification, like, e.g., those in [14], [15], suitably adapted to the case of cone constraints.

Next, we prove sufficiency of (8)–(11). Now we need no constraints qualifications, like assumption (iii) of Theorem 2, but, as could be expected in view of the similar result for the real-valued case, we do need K_0-convexity of the objective function f.

Theorem 3. *Consider problem (P) and assume that f is K_0-convex and differentiable and g_i is K_i-convex and differentiable $(1 \leq i \leq p)$. If there exist $\bar{x} \in U$ and $z_i \in V_i$ $(0 \leq i \leq p)$ such that* (8)–(11) *hold, then $\bar{x}$ solves (P).*

Proof. Assume that the result does not hold, i.e. $\bar{x}$ is not a solution of (P). Since $\bar{x} \in F$ by (9), it follows that there exists $x \in F$ such that

$$f(\bar{x}) - f(x) \succ 0. \tag{38}$$

Let $y = \bar{x} - x$. By (38) and Proposition 3(iii) applied to the K_0-convex function f, we get $df(\bar{x})y \succeq f(\bar{x}) - f(x) \succ 0$, implying that

$$df(\bar{x})y \succ 0. \tag{39}$$

By K_i-convexity of g_i and Proposition 3(iii),

$$g_i(x) - g_i(\bar{x}) + dg_i(\bar{x})(\bar{x} - x) \succeq 0. \tag{40}$$

Then,

$$\langle z_i, g_i(x)\rangle + \langle z_i, dg_i(\bar{x})y\rangle = \langle z_i, g_i(x)\rangle - \langle z_i, g_i(\bar{x})\rangle + \langle z_i, dg_i(\bar{x})(\bar{x} - x)\rangle =$$

$$\langle z_i, g_i(x) - g_i(\bar{x}) + dg_i(\bar{x})(\bar{x} - x)\rangle \geq 0, \tag{41}$$

using (11) in the first equality and (10), (40) in the inequality. Since $x \in F$, we have that $-g_i(x) \succeq 0$, and so, using (10),

$$-\langle z_i, g_i(x)\rangle \geq 0. \tag{42}$$

Summing (41) and (42) with i between 1 and p, we get

$$\sum_{i=1}^{p} \langle z_i, dg_i(\bar{x})y\rangle \geq 0. \tag{43}$$

Since $z_0 \in K_0$ by (10), we get from (43) that $\big(\sum_{i=1}^{p}\langle z_i, dg_i(\bar{x})y\rangle\big) z_0 \in K_0$, so that

$$\left(\sum_{i=1}^{p} \langle z_i, dg_i(\bar{x})y\rangle\right) z_0 \succeq 0. \tag{44}$$

From (39) and (44) we conclude that

$$df(\bar{x})y + \left(\sum_{i=1}^{p} \langle z_i, (dg_i(x)y)\rangle\right) z_0 \succ 0,$$

which implies immediately that

$$\mathrm{Im}\left[df(\bar{x}) + \left(\sum_{i=1}^{p} \langle z_i, dg_i(\bar{x})\rangle\right) z_0\right] \cap \mathrm{int}(K_0) \neq \emptyset,$$

in contradiction with (8). Thus, $\bar{x}$ solves (P). □

A problem similar to (P) was studied in [6] and [9]. It consists of finding the so called ideal (or full) minimizers of f instead of weak Pareto ones. A point $\bar{x} \in F$ is an *ideal minimizer* of f subject to the constraint F if $f(\bar{x}) \preceq f(x)$ for all $x \in F$. It is easy to check, using Proposition 1, that every ideal minimizer is a Pareto minimizer, but the converse, of course, does not hold.

A set of first order optimality conditions for the full minimization problem was presented for the case of $U = \mathbb{R}^n$, $V_i = S^{r_i}$ and $K_i = S^{r_i}_+$ $(0 \leq i \leq p)$. These conditions, which were shown in Theorem 3 of [9] to be necessary and sufficient for full optimization of (P) under the same hypotheses of Theorems 2 and 3, plus an additional constraint qualification for necessity, akin to linear independence of the gradients of the active constraints in the real-valued case, are the following.

$$df(\bar{x}) + \sum_{i=1}^{p} Z^i dg_i(\bar{x}) = 0 \qquad \text{Lagrangian condition,} \tag{45}$$

$$g_i(\bar{x}) \preceq 0 \quad (1 \leq i \leq p) \qquad \text{primal feasibility,} \tag{46}$$

$$Z^i \succeq 0 \quad (1 \leq i \leq p), \qquad \text{dual feasibility,} \tag{47}$$

$$Z^i(g_i(\bar{x})) = 0 \quad (1 \leq i \leq p) \qquad \text{complementarity,} \tag{48}$$

where Z^i belongs to $L(V_i, V_0)$, the space of linear transformations from V_i to V_0. The condition $Z^i \succeq 0$ means that $Z^i(v_i) \succeq 0$ for all $v_i \succeq 0$ in V_i, i.e. $Z^i(K_i) \subset K_0$. The expression $Z^i dg_i(\bar{x})$ in (45) is the composition of Z^i and $dg_i(\bar{x})$, i.e. the linear transformation from U to V_0 which maps $y \in U$ to $Z^i(dg_i(\bar{x})y)$.

A set of first order optimality conditions for problem (P), i.e. for constrained weak Pareto minimization, formally similar to (45)–(48), can also be developed. It consists of (46), (47), (48) plus the following condition which substitutes for (45),

$$\mathrm{Im}\left[df(\bar{x}) + \sum_{i=1}^{p} Z^i dg_i(\bar{x})\right] \cap \mathrm{int}(K_0) = \emptyset. \tag{49}$$

The fact that these conditions enjoy the same properties as (8)–(11) is established in the next proposition. We will refer to (49), (46), (47) and (48) as *LT conditions* (because they use linear transformations as multipliers, instead of vectors).

Proposition 8. *Consider problem (P). Then, under the hypotheses of Theorem 2, the existence of $Z^i \in L(V_i, V_0)$ $(1 \leq i \leq p)$ satisfying the LT conditions is necessary for weak Pareto optimality of $\bar{x}$, and it is sufficient under the hypotheses of Theorem 3.*

Proof. For necessity, consider the vectors $z_i \in V_i$ $(0 \leq i \leq p)$ whose existence is guaranteed by Theorem 2, and define $Z^i \in L(V_i, V_0)$ as

$$Z^i(v_i) = \langle z_i, v_i \rangle z_0. \tag{50}$$

It is a matter of routine to verify that (8)–(11) imply that the LT conditions hold, with the Z^i's defined as in (50).

For sufficiency, proceed as in the proof of Theorem 3 up to (41) and then, instead of scalar multiplication by z_i, apply Z^i to both sides of (41), and continue in a similar way, using (47) and (48) instead of (10) and (11), obtaining at the end a contradiction with (49). □

In fact, we can look at Theorem 2 as saying that the Z^i's which satisfy the LT conditions can be chosen so that they are all rank-one, as given by (50). We could say that the LT conditions are better that (8)–(11) for sufficiency, because there are more options for the multipliers (we can use linear transformations of rank larger than 1), and worse for necessity, for the same reason.

If we compare now the first order conditions in their LT version for weak Pareto minimization and "ideal" minimization, i.e. (49) and (45), we note that (45) is more demanding, because it requires that the image of a linear transformation be $\{0\}$, while for (49) it suffices that this image does not intersect the interior of K_0. This is reasonable, since ideal minimizers are scarcer than weak Pareto ones.

5. A duality scheme for convex vector optimization problems

Next we develop a duality scheme for (P). First we need to define weak Pareto optimality for point-to-set mappings. Consider finite dimensional real vector spaces V, V', a closed and convex cone $K \subset V$ inducing a partial order "$\preceq$" in V, a subset A of V' and a point-to-set mapping $\Phi : V' \to \mathcal{P}(V)$.

Definition 1. *$\bar{z} \in V'$ solves*

$$\max \Phi(z)$$

$$\text{s.t.} \quad z \in A$$

if and only if $\bar{z} \in A$ and there exists $\bar{v} \in \Phi(\bar{z})$ such that for any $z \in A$ and all $v \in \Phi(z)$ it does not hold that $\bar{v} \prec v$.

Now, in the setting of (P), we define the Lagrangian $\mathcal{L} : U \times V_0 \times \cdots \times V_p \to V_0$ as

$$\mathcal{L}(x, z_0, \ldots, z_p) = f(x) + \left(\sum_{i=1}^{p} \langle z_i, g_i(x) \rangle \right) z_0, \tag{51}$$

and the dual objective $\Psi : V_0 \times \cdots \times V_p \to \mathcal{P}(V_0)$ as

$$\Psi(z_0, \ldots, z_p) = \min_{x \in U} \mathcal{L}(x, z_0, \ldots, z_p), \tag{52}$$

i.e. $v \in \Psi(z_0, \ldots, z_p)$ if and only if there exists no $x \in U$ such that $\mathcal{L}(x, z_0, \ldots, z_p) \prec v$.

Next, we define problem (D) as

$$\max \Psi(z_0, \ldots, z_p)$$

$$\text{s.t.} \quad z_0 \in K_0 \cap K_0^*,$$

$$z_i \in K_i^* \quad (1 \leq i \leq p),$$

in the sense of Definition 1. Our duality result is the following.

Theorem 4. *i) If $x \in F$, $z_0 \in K_0 \cap K_0^*$ and $z_i \in K_i^*$ $(1 \leq i \leq p)$ then $\mathcal{L}(x, z_0, \ldots, z_p) \preceq f(x)$.*

ii) (weak duality) If $x \in F$, $z_0 \in K_0 \cap K_0^*$ and $z_i \in K_i^*$ $(1 \leq i \leq p)$ then there exists no $v \in \Psi(z_0, \ldots, z_p)$ such that $f(x) \prec v$.

iii) If $\bar{x} \in F$, $\bar{z}_0 \in K_0 \cap K_0^*$, $\bar{z}_i \in K_i^*$ $(1 \leq i \leq p)$ and $f(\bar{x}) \in \Psi(\bar{z}_0, \ldots, \bar{z}_p)$ then $\bar{x}$ solves (P) and $(\bar{z}_0, \ldots, \bar{z}_p)$ solves (D).

iv) (strong duality) If $\bar{x}$ solves (P), f is K_0-convex and the assumptions of Theorem 2 hold, then there exists $(\bar{z}_0, \ldots, \bar{z}_p) \in V_0 \times \cdots \times V_p$ such that $f(\bar{x}) \in \Psi(\bar{z}_0, \ldots, \bar{z}_p)$ and $(\bar{z}_0, \ldots, \bar{z}_p)$ solves (D).

Proof. i) Note that (51) can be rewritten as

$$\mathcal{L}(x, z_0, \ldots, z_p) = f(x) - \Big(\sum_{i=1}^{p} \langle z_i, -g_i(x) \rangle \Big) z_0. \tag{53}$$

Since $x \in F$, we have $-g_i(x) \in K_i$. Since $z_i \in K_i^*$ $(1 \leq i \leq p)$, it follows that $\langle z_i, -g_i(x) \rangle \geq 0$, and thus, since $z_0 \in K_0 \cap K_0^*$, we conclude that $\big(\sum_{i=1}^{p} \langle z_i, -g_i(x) \rangle \big) z_0 \in K_0$, which implies, in view of (53), that $\mathcal{L}(x, z_0, \ldots, z_p) \preceq f(x)$.

ii) Suppose that $f(x) \prec v$ with $v \in \Psi(z_0, \ldots, z_p)$. We get from (i) $\mathcal{L}(x, z_0, \ldots, z_p) \prec v$, in contradiction with (52).

iii) Applying (ii) with $v = f(\bar{x}) \in \Psi(\bar{z}_0, \ldots, \bar{z}_p)$, we get that there exists no $x \in F$ such that $f(x) \prec f(\bar{x})$. Since $\bar{x} \in F$, it follows that $\bar{x}$ solves (P).

For dual weak Pareto optimality of $(\bar{z}_0, \ldots, \bar{z}_p)$, consider Definition 1 with $V' = V_0 \times \cdots \times V_p$, $V = V_0$, $\Phi = \Psi$ and $A = (K_0 \cap K_0^*) \times K_1^* \times \cdots \times K_p^*$. Take $\bar{v} = f(\bar{x}) \in \Psi(\bar{z}_0, \ldots, \bar{z}_p)$. We must verify that for all $v \in \Psi(z_0, \ldots, z_p)$, with $(z_0, \ldots, z_p) \in (K_0 \cap K_0^*) \times K_1^* \times \cdots \times K_p^*$, it does not hold that $f(\bar{x}) \prec v$. Suppose that this is not true, i.e. $f(\bar{x}) \prec v \in \Psi(z_0, \ldots, z_p)$. By (i), we get

$$\mathcal{L}(\bar{x}, z_0, \ldots, z_p) \preceq f(\bar{x}) \prec v \in \Psi(z_0, \ldots, z_p),$$

in contradiction with (52). Thus $(\bar{z}_0, \ldots, \bar{z}_p)$ solves (D).

iv) Take $\bar{z}_0, \ldots, \bar{z}_p$ $(0 \leq i \leq p)$ as the z_i's whose existence is guaranteed by Theorem 2, i.e. such that $\bar{x}, z_0, \ldots, z_p$ satisfy (8)–(11). By (9), $\bar{z}_0 \in K_0 \cap K_0^*$, $(\bar{z}_1, \ldots, \bar{z}_p) \in K_1^* \times \cdots \times K_p^*$ and by (11) we have $\langle \bar{z}_i, g_i(\bar{x}) \rangle = 0$ $(1 \leq i \leq p)$, so that we obtain, in view of (51),

$$\mathcal{L}(\bar{x}, \bar{z}_0, \ldots, \bar{z}_p) = f(\bar{x}). \tag{54}$$

Observe that

$$d\mathcal{L}(\cdot, \bar{z}_0, \ldots, \bar{z}_p) = df(\cdot) + \Big(\sum_{i=1}^{p} \langle \bar{z}_i, dg_i(\cdot) \rangle \Big) \bar{z}_0. \tag{55}$$

By (8) and (55),

$$\mathrm{Im}[d_x \mathcal{L}(\bar{x}, \bar{z}_0, \ldots, \bar{z}_p)] \cap \mathrm{int}(K_0) = \emptyset. \tag{56}$$

Now we claim that $\mathcal{L}(\cdot, z_0, \ldots, z_p)$ is K_0-convex for $z_0 \in K_0 \cap K_0^*$, $z_i \in K_i^*$ $(1 \leq i \leq p)$. We proceed to prove the claim. Take x, $y \in U$, $\alpha \in (0,1)$, and define $\sigma_i(x) = \langle z_i, g_i(x)\rangle z_0$ $(1 \leq i \leq p)$. We have

$$\alpha\sigma_i(x) + (1-\alpha)\sigma_i(y) - \sigma_i(\alpha x + (1-\alpha)y) = \langle z_i, \rho(x,y)\rangle z_0, \tag{57}$$

with $\rho(x,y) = \alpha g_i(x) + (1-\alpha)g_i(y) - g_i(\alpha x + (1-\alpha)y)$. Since $\rho(x,y) \succeq 0$ by K_i-convexity of g_i, and $z_i \in K_i^*$, we conclude that $\langle z_i, \rho(x,y)\rangle \geq 0$, and so $\langle z_i, \rho(x,y)\rangle z_0 \in K_0$, because $z_0 \in K_0$. Therefore $\langle z_i, \rho(x,y)\rangle z_0 \succeq 0$, and then it follows from (57) that

$$\alpha\sigma_i(x) + (1-\alpha)\sigma_i(y) - \sigma_i(\alpha x + (1-\alpha)y) \succeq 0,$$

establishing K_0-convexity of σ_i $(1 \leq i \leq p)$. Since $\mathcal{L}(\cdot, z_0, \ldots, z_p) = f(\cdot) + \sum_{i=1}^p \sigma_i(\cdot)$, Proposition 3(ii) establishes the claim. In view of (56) and the K_0-convexity of $\mathcal{L}(\cdot, \bar{z}_0, \ldots, \bar{z}_p)$, we can apply Theorem 1(ii) and conclude that $\bar{x}$ is a weak Pareto minimizer of $\mathcal{L}(\cdot, \bar{z}_0, \ldots, \bar{z}_p)$. By (52) and (54), we have that $f(\bar{x}) \in \Psi(\bar{z}_0, \ldots, \bar{z}_p)$, proving the first statement of the result, and then $(\bar{z}_0, \ldots, \bar{z}_p)$ solves (D) by (iii). □

6. Examples

Next we discuss the problem, appearing in several application areas, which arises when a matrix known to be symmetric and positive semidefinite (e.g. a covariance matrix) is estimated from real data, and the resulting matrix $C \in \mathbb{R}^{n\times n}$ turns out not to be positive semidefinite, due to errors in the estimating process. It is necessary then to determine a matrix X which is "closest" to C, in some sense, among all symmetric positive semidefinite matrices, or perhaps among all symmetric positive definite matrices with eigenvalues bounded below by some ε.

A standard criterion for "closeness" is provided by the Frobenius norm $||\cdot||_F$, given by $||X||_F^2 = \sum_{i=1}^n \sum_{j=1}^n X_{ij}^2$, in which case the problem is formulated as

$$\min ||X - C||_F^2 \tag{58}$$

$$\text{s.t. } X \succeq \varepsilon I, \tag{59}$$

$$X = X^t, \tag{60}$$

where $I \in \mathbb{R}^{n\times n}$ is the identity matrix and $\varepsilon \in \mathbb{R}_+$, which is a quadratic problem with linear positive semidefinite constraints (see [4]).

The choice of the Frobenius norm is somewhat arbitrary, and a more encompassing approach consists of considering the vectorial problem

$$\min\ (X - C)^2 \tag{61}$$

$$\text{s.t. } X \succeq \varepsilon I, \tag{62}$$

which is a particular instance of problem (P) with $p = 1$, $U = V_0 = V_1 = S^n$, $K_0 = K_1 = S^n_+$, $\langle X, Y\rangle = \mathrm{tr}[XY]$, $f(X) = (X - C)^2$ and $g_1(X) = \varepsilon I - X$.

Regarding the issue of existence of solutions of (61)–(62), we use the K-coercivity criterion given by Proposition 7. Since $(S^n_+)^* = S^n_+$, and $\mathrm{int}(S^n_+) = S^n_{++}$, we can take as w any symmetric positive definite matrix, e.g. the identity $I \in \mathbb{R}^{n\times n}$. Observe that with this choice of w we have

$$\varphi(x) = \langle w, f(x)\rangle = \mathrm{tr}[I(X - C)^2] = \mathrm{tr}[(X - C)^2] = ||X - C||^2_F,$$

which is certainly coercive. Thus (61)–(62) always has solutions by Proposition 7. Note that when $w = I$, the problem min $\langle w, f(x)\rangle$ subject to $g_1(x) \preceq 0$ becomes precisely problem (58)–(60). By Proposition 5 (i), the solution of (58)–(60) is a weak Pareto minimizer, i.e. a solution of (61)–(62), but (61)–(62) has in general other solutions, e.g. those arising from using as w an arbitrary symmetric positive definite matrix B, i.e. the solution of min $\mathrm{tr}[B(X - C)^2]$ subject to $X \succeq \varepsilon I$.

Now we look at the first order optimality conditions for (61)–(62). Note that the Slater's condition (assumption (iii) in Theorem 2), is satisfied by taking as $\hat{x}$ the matrix $(1 + \varepsilon)I$. Since f and g_1 are S^n_+-convex and differentiable, as can be easily verified, Theorems 2 and 3 imply that $\bar{X} \in S^n$ is a solution of (61)–(62) if and only if there exist matrices Z^0, $Z^1 \in S^n$ such that

$$\bar{X} \succeq \varepsilon I, \tag{63}$$

$$Z^0 \succeq 0, \qquad Z^1 \succeq 0, \tag{64}$$

$$\mathrm{tr}[Z^1\bar{X}] = \varepsilon\,\mathrm{tr}[Z^1], \tag{65}$$

and for all $Y \in S^n$ the matrix $H(Y)$, with entries given by

$$H(Y)_{ij} = [(\bar{X} - C)Y]_{ii} + [(\bar{X} - C)Y]_{jj} - \mathrm{tr}[Z^1Y]Z^0_{ij},$$

is not positive definite.

Conditions (63)–(65) result from (9)–(11) respectively, and the condition on $H(Y)$ is just (8) for this particular instance of (P).

Finally, we consider the following example, taken from [9]. Take $p = 4$, $U = \mathbb{R}^3$, $V_0 = S^2$, $K_0 = S^2_+$, $V_i = \mathbb{R}$, $K_i = \mathbb{R}_+$ $(1 \le i \le 4)$, $f : \mathbb{R}^3 \to S^2$ defined as

$$f(x) = \begin{pmatrix} x_1 & x_2 \\ x_2 & x_3 \end{pmatrix}$$

and $g_i : \mathbb{R}^3 \to \mathbb{R}$ $(1 \le i \le 4)$ defined as

$$g_1(x) = -x_1 - x_2 + 1, \qquad g_2(x) = -x_2 - x_3 + 1, \qquad g_3(x) = -x_1 + x_2 + 1, \qquad g_4(x) = x_2 - x_3 + 1.$$

It is easy to verify directly that $\bar{x} = (1, 0, 1)^t$ is a solution of problem (P) with these data. Even more, this $\bar{x}$ is an ideal minimizer for the problem, i.e. $f(\bar{x}) \preceq f(x)$ for all $x \in F$, as shown in Example 2 of [9]. It is immediate that $\hat{x}^t = (2, 0, 2)$ satisfies the Slater's condition $g_i(\hat{x}) \prec 0$ $(1 \leq i \leq 4)$. This problem, nevertheless, fails to satisfy the stronger additional constraint qualification required in Theorem 3 of [9], where it is shown that in fact there exist no linear trasformations $Z^i : \mathbb{R} \to S^2$ $(1 \leq i \leq 4)$ such that $\bar{x}$ and the Z^i's satisfy (45)–(48). On the other hand, the situation is simpler for weak Pareto minimization, i.e. conditions (8)–(11). (9) and (11) are immediately satisfied because $g_i(\bar{x}) = 0$ $(1 \leq i \leq 4)$. (8) and (10) demand existence of $z_i \in \mathbb{R}_+$, $z_0 \in S^2_+$ such that for all $y \in \mathbb{R}^3$ the matrix

$$H(y) = \begin{pmatrix} y_1 & y_2 \\ y_2 & y_3 \end{pmatrix} + [-z_1(y_1 + y_2) - z_2(y_2 + y_3) + z_3(y_2 - y_1) + z_4(y_2 - y_3)]z_0$$

is not positive definite. A suitable choice for the multipliers is $z_0 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, $z_1 = z_3 = 1/2$, $z_2 = z_4 = 0$, in which case it holds that $H(y) = \begin{pmatrix} 0 & y_2 \\ y_2 & y_3 \end{pmatrix}$, which is not positive definite for any $y \in \mathbb{R}^3$. If we look now at the LT conditions for weak Pareto optimality, namely (46), (47), (48) and (49), they are satisfied with the linear transformations $Z^i : \mathbb{R} \to S^2$ $(1 \leq i \leq 4)$ constructed from the z_i's above through (50), and also with the Z^i's defined as $Z^1(t) = \begin{pmatrix} t & 0 \\ 0 & 0 \end{pmatrix}$, $Z^2(t) = \begin{pmatrix} 0 & 0 \\ 0 & t \end{pmatrix}$, $Z^3 = Z^4 = 0$, which are not of the form precribed in (50). This is in agreement with Theorem 2 and Proposition 8, which ensure the existence of the z_i's $(0 \leq i \leq 4)$ and the Z^i's $(1 \leq i \leq 4)$ respectively, without further constraint qualifications, besides the Slater's condition.

7. Final Remarks

Our main results, basically Theorems 2 and 3, can be useful to devise and analyze primal-dual algorithms for problem (P). In view of the scalarization scheme suggested by Proposition 5, such algorithms might seem unnecessary at first view, since, in rather general situations, it is possible to reduce problem (P) to another one, which we call (RV), with real-valued objective. In the case of $K_i = \mathbb{R}^{r_i}_+$, problem (RV) turns out to be just a nonlinear programming problem; for $K_i = S^{r_i}_+$ it becomes a nonlinear positive semidefinite problem, which has also been systematically studied in recent years from an algorithmic point of view (see, e.g., [21]). The difficulty lies in the fact that the scalarization approach is not actually constructive: even in the K-convex case, Proposition 5(ii) tells us that there exists some $w \in K^*$ such that (P) reduces to min $\langle w, f(x)\rangle$ subject to (7), but it does not indicate how to find the appropriate w. In general, it is easy to find elements of K^*, but if we take an arbitrary one, the resulting problem (RV) might very well lack solutions. This obstacle is highlighted by the definition of K-coercivity: how can we find the right $w \in K^*$ which verifies the K-coercivity of f? Thus, it seems pertinent to devise methods which work directly

with vector-valued objectives. For unconstrained vector optimization, a vector-valued version of the steepest descent method has been recently introduced in [5] and further developed in [10]; we hope that our results in this paper will allow similar extensions to the vector-valued case of the many primal-dual methods currently available for real-valued nonlinear programming. This issue is left as an open problem deserving future research.

References

[1] Avriel, M. *Nonlinear Programming: Analysis and Methods*. Prentice Hall, New Jersey (1976).

[2] Cambini, R. Some new classes of generalized concave vector valued functions. *Optimization* **36** (1996) 11–24.

[3] Clarke F.H. *Optimization and Nonsmooth Analysis*. John Wiley, New York (1983).

[4] Escalante, C., Raydan, M. Dijkstra's algorithm for a constrained least-squares matrix problem. *Numerical Linear Algebra with Applications* **3** (1996) 459-471.

[5] Fliege, J., Svaiter, B.F. Steepest descent methods for multicriteria optimimization. *Mathematical Methods of Operations Research*, **51** (2000) 479–494.

[6] Flores-Bazán, F. Ideal, weakly efficient solutions for convex vector optimization problems, Technical Report 99-26. Facultad de Ciencias Físicas y Matemáticas, Universidad de Concepción, Concepción, Chile (1999).

[7] Frenk, J.B.G., Kassay, G. On classes of generalized convex functions, Gordan-Farkas type theorems, and Lagrangian duality. *Journal of Optimization Theory and Applications* **102** (1999) 315-343.

[8] Geoffrion, A.M. Proper efficiency and the theory of vector maximization. *Journal of Mathematical Analysis and Applications* **22** (1968) 618-630.

[9] Graña Drummond, L.M., Iusem, A.N. On convexity of vector-valued mappings (to be published).

[10] Graña Drummond, L.M., Svaiter, B.F. A steepest descent method for vector optimization (to be published).

[11] Hiriart-Urruty, J.-B., Lemaréchal, C. *Convex Analysis and Minimization Algorithms*. Springer, Berlin (1993).

[12] Illes, T., Kassay, G. Theorems of the alternative and optimality conditions for convexlike and general convexlike programming. *Journal of Optimization Theory and Applications* **101** (1999) 243-257.

[13] Luc, T.D. *Theory of Vector Optimization*. Lecture Notes in Economics and Mathematical Systems **319**, Springer (1989).

[14] Maeda, T. Constraint qualifications in multiobjective optimization problems: differentiable case. *Journal of Optimization Theory and Applications* **80** (1994) 483-500.

[15] Preda, V., Chiţescu, I. On constraint qualifications in multiobjective optimization problems: semidifferentiable case. *Journal of Optimization Theory and Applications* **100** (1999) 417-434.

[16] Rockafellar, R.T. *Convex Analysis*. Princeton University Press, New Jersey (1970).

[17] Shapiro, A. First and second order analysis of nonlinear semidefinite programs. *Mathematical Programming* **77** (1997) 301–320.

[18] Shapiro, A. On uniqueness of Lagrange multipliers in optimization problems subject to cone constraints. *SIAM Journal on Optimization* **7** (1997) 508-518.

[19] Singh, C. Optimality conditions in multiobjective differentiable programming. *Journal of Optimization Theory and Applications* **53** (1987) 115-123.

[20] Sundaram, R.K. *A First Course in Optimization Theory*. Cambridge University Press, Cambridge (1996).

[21] Vanderberghe, L., Boyd, S. Positive-definite programmming. In *Mathematical Programming: state of the art* (J.R. Birge and K.G.M. Murty, editors). University of Michigan, Ann Arbor (1994) 276-308.

[22] van Rooyen, M., Zhou, X., Zlobec, S. A saddle-point characterization of Pareto optima. *Mathematical Programming* **67** (1994) 77-88.

[23] Zlobec, S. Characterizing optimality in mathematical programming models. *Acta Applicandae Mathematicae* **12** (1988) 113-180.