ON THE THEORY OF VECTOR OPTIMIZATION AND VARIATIONAL INEQUALITIES. IMAGE SPACE ANALYSIS AND SEPARATION[1]

Franco Giannessi, Giandomenico Mastroeni

Department of Mathematics
University of Pisa, Pisa, Italy

Letizia Pellegrini

Institute of Mathematics
University of Verona, Verona, Italy

ABSTRACT. By exploiting recent results, it is shown that the theories of Vector Optimization and of Vector Variational Inequalities can be based on the image space analysis and theorems of the alternative or separation theorems. It is shown that, starting from such a general scheme, several theoretical aspects can be developed – like optimality conditions, duality, penalization – as well as methods of solution – like scalarization.

KEY WORDS. Vector optimization, variational inequality, separation, optimality conditions, saddle point, duality, penalization, scalarization, image space, quasi–variational inequality.

AMS Classification: 90C, 49J, 65K

1. INTRODUCTION

The growing interest in vector problems, both from a theoretical

[1] Sections 1 and 2 are due to F. Giannessi; Sections 3,5,6,9,11, are due to G. Mastroeni; Sections 4,7,8,10 are due to L. Pellegrini.

point of view and as it concerns applications to real problems, asks for a general scheme which embraces several existing developments and stimulates new ones.

The present paper aims at contributing to set up such a scheme, by taking into account recent results. Indeed, a first proposal in this direction was made in [18]. It has produced some developments in the field of Vector Optimization Problems (for short, VOP): initially in [14]; subsequently, by exploiting the approach of [19], in [30,31]; recently, in [32,5,23]. The hints contained in [18,19] have been taken by Chen and other Authors in the field of Vector Variational Inequalities (for short, VVI); see References on VVI at the end of this Volume.

As far as VOP are concerned, we note that, several years after [18], some aspects of the image space analysis proposed in [18] have appeared in some papers independently from each other and from [18]; see [2,3,9,10,15–17].

Sects. 2–4 contain the general Image Space (for short, IS) analysis and separation theory for VOP as concerns both the generalized Pareto case and the so–called weak case. This analysis is extended to VVI in Sect. 9. Then, it is shown how to derive, from the general separation theory, necessary optimality conditions for VOP (Sect. 5), saddle point optimality conditions for VOP (Sect. 6), duality for VOP (Sect. 7), scalarization (Sect. 8 for VOP and Sect. 10 for VVI) and penalization (Sect. 11).

Since the main scope of the present paper is not the existence of extrema, in what follows the assumptions on their existence will be understood.

2. IMAGE SPACE AND SEPARATION FOR VOP

Let the positive integers ℓ, m, n and the cone $C \subset \mathbb{R}^\ell$ be given. In the following it will be assumed that C is convex, closed and pointed with apex at the origin and with $\operatorname{int} C \neq \emptyset$, namely with nonempty interior [2]. Inclusion (with possible coincidence) and strict

[2] Some of the propositions which will be established do not require all these assumptions on C.

inclusion (without coincidence) will be denoted by $\subseteq$ and $\subset$, respectively.

Let us consider the vector–valued functions $f : \mathbb{R}^n \to \mathbb{R}^\ell$, $g : \mathbb{R}^n \to \mathbb{R}^m$, and the subset $X \subseteq \mathbb{R}^n$. We will consider the following vector minimization problem, which is called *generalized Pareto problem:*

$$\text{(2.1)} \quad \min_{C\setminus\{0\}} f(x) \quad , \quad \text{subject to} \quad x \in K := \{x \in X : g(x) \geq 0\},$$

where $\min_{C\setminus\{0\}}$ marks vector minimum with respect to the cone $C\setminus\{0\}$: $y \in K$ is a (global) vector minimum point (for short, v.m.p.) of (2.1), iff

$$\text{(2.2)} \qquad f(y) \not\geq_{C\setminus\{0\}} f(x) \ , \quad \forall x \in K,$$

where the inequality means $f(y) - f(x) \notin C\setminus\{0\}$. At $C = \mathbb{R}^\ell_+$, (2.1) becomes the classic *Pareto vector problem.*

A vector minimization problem which is often associated to (2.1) is the following one, called *weak vector problem*:

$$\text{(2.3)} \qquad \min_{\operatorname{int} C} f(x) \quad , \quad \text{s.t.} \quad x \in K,$$

where $\min_{\operatorname{int} C}$ marks vector minimum with respect to the cone $\operatorname{int} C$: $y \in K$ is a (global) v.m.p. of (2.3), iff

$$\text{(2.4)} \qquad f(y) \not\geq_{\operatorname{int} C} f(x) \quad , \quad \forall x \in K,$$

where the inequality means $f(y) - f(x) \notin \operatorname{int} C$. At $C = \mathbb{R}^\ell_+$, (2.3) is called *weak vector Pareto problem.*

Problem (2.3) is obviously different from (2.1), since different cones identify different vector problems. The term “weak” comes from the following tradition. Notwithstanding the fact that (2.1) and (2.3) are distinct problems, since the solutions of (2.1) are solutions also of (2.3) (but not necessarily vice versa), then the solutions of (2.3) are often called “weak solutions” of (2.1). It would be better to say that being a solution to (2.3) is *necessary* in order to be a solution to (2.1). Indeed, the term “weak” is misleading and in contrast with its use in other branches of Mathematics; “relaxed” would be a more appropriate term. Instead of cutting off the entire[3] frt C we might subtract any part of it obtaining several “relaxations” of (2.1).

[3] frt denotes frontier.

It is trivial to note that (2.2) is satisfied iff the system (in the unknown x):

$$(2.5)\qquad f(y)-f(x)\geq_{C\backslash\{0\}}0\ \ ,\ \ g(x)\geq 0\ \ ,\ \ x\in X$$

is impossible. Consider the sets:

$$\mathcal{H}:=\{(u,v)\in\mathbb{R}^{\ell}\times\mathbb{R}^{m}:u\geq_{C\backslash\{0\}}0\ ,\ v\geq 0\}=(C\backslash\{0\})\times\mathbb{R}^{m}_{+}\ ,$$
$$\mathcal{K}(y):=\{(u,v)\in\mathbb{R}^{\ell}\times\mathbb{R}^{m}:u=f(y)-f(x)\ ,\ v=g(x),\ x\in X\}.$$

In what follows, when there is no fear of confusion, $\mathcal{K}(y)$ will be denoted merely by $\mathcal{K}$. $\mathcal{H}$ and $\mathcal{K}$ are subsets of $\mathbb{R}^{\ell+m}$, which is called *image space*; $\mathcal{K}$ is called the *image* of (2.1). We see that (2.1) is equivalent – through the map $\mathcal{M}_y(x):=(f(y)-f(x),g(x))$ – to the following vector maximization problem:

$$(2.6)\qquad \max_{C\backslash\{0\}}u\ \ ,\ \ \text{s.t.}\ \ (u,v)\in\mathcal{K}\cap(\mathbb{R}^{\ell}\times\mathbb{R}^{m}_{+})\ ,$$

where $\max_{C\backslash\{0\}}$ marks vector maximum with respect to the cone $C\backslash\{0\}$: $(\bar{u},\bar{v})\in\mathcal{K}\cap(\mathbb{R}^{\ell}\times\mathbb{R}^{m}_{+})$ is a vector maximum point of (2.6) iff

$$(2.7)\qquad \bar{u}\not\leq_{C\backslash\{0\}}u\ \ ,\ \ \forall(u,v)\in\mathcal{K}\cap(\mathbb{R}^{\ell}\times\mathbb{R}^{m}_{+}).$$

(2.6) is called *image problem* associated to (2.1).

Let K^0 and $\mathcal{K}^0$ denote the sets of solutions of (2.1) and (2.6), respectively; of course, $\mathcal{K}^0=\mathcal{M}_y(K^0)$. We observe that if K is compact, f and g continuous, then $\mathcal{K}$ is compact. If K is a polytope, f linear and g affine, then $\mathcal{K}$ is a polytope; in this case $K^0\subseteq \text{frt}\ K$ and $\mathcal{K}^0\subseteq\text{frt}\ \mathcal{K}$. If K,f and $-g$ are convex, then $\mathcal{K}$ is not necessarily convex; this might seem a drawback; it is easy to overcome it by replacing equivalently $\mathcal{K}$ with a wider set (see $(2.8)'$ below), which is convex.

The systematic investigation of the relations between the properties of (2.1) and those of (2.6) is substantially still to be done. This aspect is important, not only for the development of the theory, but also for the construction of algorithms. Indeed, the investigations done in the IS have occurred as an auxiliary step toward other achievements. For instance, in [2,3,15] the main scope is the definition of algorithms for solving (2.1); they take great advantage from the analysis in the IS, notwithstanding the fact that it is limited.

With regard to the theory, even existence theorems can receive improvements. In the scalar case, the analysis in the IS has led to extend substantially the known existence theorems [40] and, above all, to give a "source" for deriving theorems; it would be interesting to extend these results to the vector case.

The analysis carried out in the present paper is based on a generic cone C, instead of the classic $\mathbb{R}^\ell_+$. This has several advantages. Apart from the fact that the ordering criterion might be expressed by a cone $C \neq \mathbb{R}^\ell_+$, even if it is exactly $\mathbb{R}^\ell_+$, we might desire to cut off certain "irregular" cases. More precisely, consider the class where: (i) a nonzero element of the ordering cone C is limit point of $f(y) - f(K)$; or (ii) an element of $\mathcal{H}$ is limit point of a sequence of elements of $\mathcal{K}$; or (iii) the Bouligand tangent cone to $\mathcal{K}$ at the origin intersects $\mathcal{H}$. These cases (which would be extremely interesting to characterize in terms of C, X, f and g) might represent undesirable situations, in the sense that small perturbations in the data might delete the optimality of a v.m.p.. Drawback (i) can be overcome by replacing, in (2.1) and (2.2), the ordering cone with a cone C' with apex at the origin and such that $C \backslash (0) \subset \operatorname{int} C'$. In this case a v.m.p. of (2.1) is called *proper efficient* v.m.p. (see Definition 3.1.8 of [39]). If $C = \mathbb{R}^\ell_+$, the previous definition is equivalent to the following one [39]: $y \in K$ is a proper Pareto minimum point (or is a v.m.p. of (2.1) at $C' \supset \mathbb{R}^\ell_+$) iff $\exists M > 0$, such that

$$\left.\begin{array}{l} x \in K \\ f_i(x) < f_i(y) \end{array}\right\} \quad \Rightarrow \quad \left\{\begin{array}{l} \exists j \ \text{ such that } \ f_j(x) > f_j(y) \ \text{ and} \\ f_i(y) - f_i(x) \leq -M[f_j(y) - f_j(x)]. \end{array}\right.$$

When drawback (ii) happens, which is more general that (i), since the elements of the sequence in (ii) are not necessarily images of elements of $\mathcal{K}$, then replacing C with C' may not work; namely, a proper efficient v.m.p. may continue to suffer from the above–mentioned "instability".

Drawbacks (ii) and (iii) are illustrated by the two following examples.

Example 1. Let us set $\ell = 2, m = 1, n = 3, C = \mathbb{R}^2_+, X = \mathbb{R}^3_+$, $f_1(x) = x_1^2 x_3, f_2(x) = x_2^2 x_3, g(x) = x_3; y^0 := (0,0,0)$ is a v.m.p. of (2.1), since (2.2) becomes

$$(-x_1^2 x_3, -x_2^2 x_3) \notin \mathbb{R}^2_+ \backslash \{0\} \quad , \quad \forall (x_1, x_2, x_3) \ \text{ s.t. } \ x_3 \geq 0,$$

and it is trivially satisfied. No other feasible y is a v.m.p. We have:

$$\begin{aligned}\mathcal{K}(y^0) &= \{(u_1,u_2,v)\in \mathbb{R}^3 : u_1=-x_1^2 v,\ u_2=-x_2^2 v,\ x_1,x_2\in\mathbb{R}\} = \\ &= \{(u_1,u_2,v)\in\mathbb{R}^3 : u_1,u_2\geq 0, v<0 \ \text{ or } \ u_1,u_2\leq 0, v>0\}.\end{aligned}$$

Therefore every element of $(C\backslash\{0\})\times\{0\}$ is a limit point of elements of $\mathcal{K}(y^0)$. In fact, let $(\bar{u}_1,\bar{u}_2)\in C\backslash\{0\}$ and set $x_1=\sqrt{\bar{u}_1 n}$, $x_2=\sqrt{\bar{u}_2 n}$, $v=-1/n$, with $n\in\mathbb{N}\backslash\{0\}$. Then, the point $(u_1,u_2,v)=(-x_1^2 v,-x_2^2 v,v)=(\bar{u}_1,\bar{u}_2,-1/n)\in\mathcal{K}(y^0)$ defines a sequence of elements of $\mathcal{K}(y^0)$ converging to $(\bar{u}_1,\bar{u}_2,0)$ as $n\to+\infty$.

Example 2. Let us set $\ell=2, m=1, n=1, C=\mathbb{R}^2_+, X=\mathbb{R}_-$, $f_1(x)=x, f_2(x)=-\sqrt[4]{-x}, g(x)=-\sqrt{-x}; y^0=0$ is obviously (the only) v.m.p. of (2.1), since it is the only feasible point. We have:

$$\mathcal{K}(y^0)=\{(u_1,u_2,v)\in\mathbb{R}^3 : u_1\geq 0, u_2=\sqrt[4]{u_1}\, , v=-\sqrt{u_1}\}\ .$$

It is easy to see that the Bouligand tangent cone to $\mathcal{K}(y^0)$ at the origin intersects $(C\backslash\{0\})\times\{0\}$ (even if no element of it is limit point of elements of $\mathcal{K}(y^0)$). In fact, we will show that the following element of $(C\backslash\{0\})\times\{0\}$, namely $(\bar{u}_1=0,\bar{u}_2>0,\bar{v}=0)$, belongs also to the Bouligand tangent cone to $\mathcal{K}(y^0)$ at the origin. By setting $u_1=1/n$ with $n\in\mathbb{N}\backslash\{0\}$, we see that the point $(u_1,u_2,v)=(1/n,,1/\sqrt[4]{n},-1/\sqrt{n})\in\mathcal{K}(y^0)$ defines a sequence converging to $(0,0,0)$ as $n\to+\infty$, and that the same point, when multiplied by the positive scalar $\alpha_n:=\bar{u}_2\sqrt[4]{n}$, defines a sequence converging to $(0,\bar{u}_2,0)$ as $n\to+\infty$.

Now, observe that system (2.5) is impossible iff

$$\mathcal{H}\cap\mathcal{K}=\emptyset\ . \tag{2.8}$$

It is easy to see that (2.8) holds iff

$$\mathcal{H}\cap[\mathcal{K}-c\ell\mathcal{H}]=0\ , \tag{2.8$'$}$$

where the difference is meant in vector sense and $c\ell$ denotes closure. $\mathcal{E}(y):=\mathcal{K}-c\ell\mathcal{H}$ is called *conic extension*[4] of the image; it often

[4] Unless necessary, the depencence on y will be understood.

enjoys more properties than the image itself. The map $\mathcal{M}_y$, which leads from the image to its conic extension, implies a change of the functions f and g and hence a transformation of the VOP without losing the optimality of a global v.m.p. (the implications of this aspect are still to be investigated). As mentioned earlier, the conic extension $\mathcal{E}$ plays an important role in making (2.6) convex in some cases where $\mathcal{K}$ is not so. More generally, $\mathcal{E}$ "regularizes" (2.6) by deleting, for instance, some nonconvexities and discontinuities. It can be useful also in constructing algorithms. For instance, in [15] it has been introduced, independently of the existing literature (where the concept of conic extension is an old one; its! ! use in optimization has been done in [19]), under a neologism (set of "satisfactory vectors") and turned out to be of fundamental importance in the algorithmic developments.

The analysis in the IS must be viewed as a preliminary and auxiliary step – and not as a concurrent analysis – for studying a VOP. If this aspect is understood, then the IS analysis may be highly fruitful. In fact, in the IS we may have a sort of "regularization": the conic extension of the image may be convex or continuous or smooth when the VOP (and its image) do not enjoy the same property, so that convex or continuous or smooth analysis can be developed in the IS, but not in the given space. The image of a VOP is finite dimensional even if X is a subset of a Hilbert space (and not of $\mathbb{R}^n$ as it is here) provided that the image of g be finite dimensional (if not, the present approach is still valid, but must be substantially modified); hence, in the IS such infinite dimensional problems can be analysed by means of the same mathematical concepts which are used for the finite dimensional case. Since the definition of $\mathcal{K}$ depends on y – namely, the unknown – it ! ! may seem that the analysis of its properties depends on the knowledge of y; note that the change of y implies merely a translation of $\mathcal{K}$ and does not affect its properties.

Now, let us observe that obviously y is a v.m.p. of (2.1) iff (2.8) holds. Unfortunately, to prove (2.8) directly is, in general, a difficult task. A way of proving (2.8) indirectly consists in obtaining the existence of a function, such that two of its disjoint level sets contain $\mathcal{H}$ and $\mathcal{K}$, respectively. To this end, let k be a positive integer and $D \subset \mathbb{R}^k$ a cone. The *vector polar* of D with respect to C is given by

[5]

$$D_C^* := \{M \in \mathbb{R}^{\ell \times k} : Md \geq_C 0 \quad , \quad \forall d \in D\},$$

where $\mathbb{R}^{\ell \times k}$ denotes the set of matrices with real entries and of order $\ell \times k$, and where the inequality means $Md \in C$. At $\ell = 1$, D_C^* becomes either the *positive* or *negative polar cone* of D, according to $C = \mathbb{R}_+$ or $C = \mathbb{R}_-$, respectively. When $C = \mathbb{R}_+^\ell$ (the Pareto case), then we have:

$$(2.9) \quad D^*_{\mathbb{R}^\ell_+} = \left\{ M = \begin{pmatrix} d_1^* \\ \vdots \\ d_\ell^* \end{pmatrix} \in \mathbb{R}^{\ell \times k} : d_i^* \in D^* \quad , \quad i = 1, \ldots, \ell \right\} ,$$

where d_i^* is the i–th row of M and D^* denotes the (positive) polar cone of D.

Now, we can introduce the above–mentioned function for proving indirectly (2.8). To this end, let us set $U := C \backslash \{0\}$ and $V := \mathbb{R}^m_+$, and consider the function $w : \mathbb{R}^\ell \times \mathbb{R}^m \to \mathbb{R}^\ell$, given by

$$(2.10) \qquad w = w(u, v; \Theta, \Lambda) = \Theta u + \Lambda v \quad , \quad \Theta \in U^*_{C \backslash \{0\}}, \Lambda \in V^*_C \ ,$$

where Θ and Λ are parameters. Indeed, (2.10) is a family of functions within which we will look for one in order to achieve the above purpose. The following proposition shows that the class of functions (2.10) has the aforesaid property. Consider the "positive" level set of (2.10):

$$W_{C \backslash \{0\}}(u, v; \Theta, \Lambda) := \{(u, v) \in \mathbb{R}^\ell \times \mathbb{R}^m : w(u, v; \Theta, \Lambda) \geq_{C \backslash \{0\}} 0\} \ .$$

Proposition 1. If w is given by (2.10), then we have:

$$(2.11a) \qquad \mathcal{H} \subset W_{C \backslash \{0\}}(u, v; \Theta, \Lambda), \quad \forall \Theta \in U^*_{C \backslash \{0\}} \quad \forall \Lambda \in V^*_C,$$

$$(2.11b) \qquad \mathcal{H} = \bigcap_{\substack{\Theta \in U^*_{C \backslash \{0\}} \\ \Lambda \in V^*_C}} W_{C \backslash \{0\}}(u, v; \Theta, \Lambda) \ .$$

[5] Of course, such a definition does not require that C be convex, closed and pointed. However, since C expresses a partial order, the lack of these properties restricts the possible applications.

Proof. $(u,v) \in \mathcal{H} \;\Leftrightarrow\; u \in U, v \in V$. Therefore, $\forall \Theta \in U^*_{C\backslash\{0\}}$ and $\forall \Lambda \in V^*_C$, we have $\Theta u + \Lambda v \in C\backslash\{0\}$; hence $w(u,v;\Theta,\Lambda) \geq_{C\backslash\{0\}} 0$. (2.11a) follows. Because of (2.11a), to show (2.11b) it is enough to prove that no element of the complement of $\mathcal{H}$ belongs to the right–hand side of (2.11b); namely that, $\forall (\tilde{u},\tilde{v}) \notin \mathcal{H}, \exists \tilde{\Theta} \in U^*_{C\backslash\{0\}}$ and $\exists \tilde{\Lambda} \in V^*_C$, such that

$$(\tilde{u},\tilde{v}) \notin W_{C\backslash\{0\}}(u,v;\tilde{\Theta},\tilde{\Lambda}). \tag{2.12}$$

$(\tilde{u},\tilde{v}) \notin \mathcal{H}$ implies at least one of the following cases: (i) $\tilde{u} \notin C\backslash\{0\}$ or (ii) $\tilde{v} \notin \mathbb{R}^m_+$. If (i) holds, (2.12) is obtained with $\tilde{\Theta} = I_\ell$ (the identity matrix of order ℓ) and $\tilde{\Lambda} = 0$ (the null matrix of order $\ell \times m$), since we have $w(\tilde{u},\tilde{v};\tilde{\Theta},\tilde{\Lambda}) = \tilde{u} \notin C\backslash\{0\}$. If (ii) holds, then $\exists i_0 \in \{1,\ldots,m\}$, such that $\tilde{v}_{i_0} < 0$. Set

$$\tilde{\Theta} = \alpha I_\ell, \quad \tilde{\Lambda} = \begin{pmatrix} 0 & \ldots & 0 & \tilde{c}_{1i_0} & 0 & \ldots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ 0 & \ldots & 0 & \tilde{c}_{\ell i_0} & 0 & \ldots & 0 \end{pmatrix},$$

where $\tilde{c}^T := (\tilde{c}_{1i_0},\ldots,\tilde{c}_{\ell i_0}) \in C\backslash\{0\}$ and $\alpha > 0$. Note that $\tilde{\Lambda} \in V^*_C$, since, $\forall v \in \mathbb{R}^m_+$, $\tilde{\Lambda} v = v_{i_0}\tilde{c} \in C$. We have $w(\tilde{u},\tilde{v};\tilde{\Theta},\tilde{\Lambda}) = \alpha\tilde{u} + \tilde{v}_{i_0}\tilde{c}$. Since C is pointed and $\tilde{v}_{i_0} < 0$, then $\tilde{v}_{i_0}\tilde{c} \notin C$. Therefore, $\tilde{v}_{i_0}\tilde{c}$ belongs to the complement of C which is open. If α is small enough we obtain

$$w(\tilde{u},\tilde{v};\tilde{\Theta},\tilde{\Lambda}) = \tilde{\alpha}\tilde{u} + \tilde{v}_{i_0}\tilde{c} \notin C$$

which shows (2.12). ◻

Now, we are in the position to state an optimality condition for (2.1).

Theorem 1. Let $y \in K$. If there exist matrices $\Theta \in U^*_{C\backslash\{0\}}$ and $\Lambda \in V^*_C$, such that

$$\Theta[f(y) - f(x)] + \Lambda g(x) \not\geq_{C\backslash\{0\}} 0\ , \quad \forall x \in X, \tag{2.13}$$

then y is a (global) v.m.p. of (2.1).

Proof. From Proposition 1 and (2.13) we have[6]

$$\mathcal{H} \subset W_{C\setminus\{0\}}(u, v; \Theta, \Lambda) \quad \text{and} \quad \mathcal{K} \subseteq \ \sim W_{C\setminus\{0\}}(u, v; \Theta, \Lambda) \ ,$$

respectively. Therefore, (2.8) holds. □

At $\ell = 1$ the above theorem collapses to an existing one for scalar optimization (see Corollary 5.1 of [19]); Θ can now be replaced by 1, since here it is a positive scalar and (2.13) is homogeneous with respect to the parameters.

Theorem 1 shows that the functions of the class (2.10) are *weak separation functions* in the sense of [19]; namely, $\mathcal{H}$ is contained in the "positive level set" of each w of the class; indeed, (2.11b) shows something more. In fact, because of (2.11b), the intersection of the "positive level sets" of the separation functions (2.10) is not open, since $\mathcal{H}$ is not open. Therefore, an element of $\mathcal{H}$ may be limit point of elements of $\mathcal{K}$, or the Bouligand tangent cone to $\mathcal{K}$ at the origin may intersect $\mathcal{H}$ even when (2.8) holds[7], so that no element of the class (2.10) may exist which fulfils (2.13) and then (2.8). Hence, it is useful to develop an alternative approach, which consists in replacing the class (2.10) with the wider class:

$$w = w(u, v; \Theta, \Lambda) = \Theta u + \Lambda v, \quad \Theta \in U_C^*, \quad \Lambda \in V_C^* \ . \tag{2.14}$$

The intersection of the "positive level sets" of the separation functions (2.14) is open.

Of course, with a w of class (2.14) condition (2.13) no longer guarantees that y be a v.m.p. of (2.1). This drawback can be overcome by restricting the class of problems (2.1). Such a restriction is made by means of a condition, which is called *constraint qualification* if it deals only with the constraints, and *regularity condition* if it involves the objective functions too.

Before closing this section, let us pose a question which should be interesting to investigate in the IS. A VOP may be a dummy

[6] $\sim$ denotes complement.

[7] A characterization, in terms of X, f and g, of this case would be extremely interesting.

problem, in the sense that the set of its v.m.p. coincides with K or it is a singleton. With regard to the first aspect, we state a proposition, in the given space, in order to stimulate further investigation.

It may happen that every element of K be a v.m.p. of (2.1). In this case (2.1) is trivial; it has been investigated in the case of $C = \mathbb{R}^{\ell}_{+}$ and called *completely efficient* (see [3]). We will show that, when C is as in Sect. 2, a characterization of complete efficiency is easy:

let $p \in \operatorname{int} C^*$; every element of K is a v.m.p. of (2.1) iff

$$\max_{\substack{f(y)-f(x)\in C \\ x,y\in K}} \langle p, f(y) - f(x)\rangle = 0 \ . \tag{2.15}$$

Only if. By assumption, (2.2) holds, or

$$\forall y \in K : f(y) - f(x) \notin C \backslash \{0\} \quad \forall x \in K \ , \tag{2.16}$$

so that

$$\forall y \in K : \quad \left.\begin{array}{c} f(y) - f(x) \in C \\ x \in K \end{array}\right\} \quad \Rightarrow \quad f(y) - f(x) = 0 \quad \Rightarrow \quad (2.15) \ .$$

If. Ab absurdo, suppose that $\exists y \in K$ and $\exists x_y \in K$, such that

$$f(y) - f(x_y) \in C \backslash \{0\} \ ,$$

so that $f(y) - f(x_y) \neq 0$. It follows $\langle p, f(y) - f(x_y)\rangle > 0$, which contradicts (2.15). □

At $C = \mathbb{R}^{\ell}_{+}$, condition (2.15) becomes a known one; see [3]. Of course, by considering y fixed, (2.15) is a necessary and sufficient condition for y to be a v.m.p. of (2.1). Starting from (2.15) it is interesting to study existence conditions for (2.1). When the maximum of (2.15) is positive, its value might be used to define the "width" of the set of v.m.p. relative to that of K.

3. OTHER KINDS OF SEPARATION

With the class (2.10) we have adopted a linear vector separation function w having ℓ components. Of course, we might have adopted a nonlinear w or a w with any number of components, even with infinite ones. In particular, we can choose a scalar w, i.e. $w : \mathbb{R}^\ell \times \mathbb{R}^m \to \mathbb{R}$. It may be linear or nonlinear; in the linear case it is given by[8] [19,23]:

(3.1) $w = w(u, v; \theta, \lambda) = \langle \theta, u\rangle + \langle \lambda, v\rangle$, $\theta \in \operatorname{int} C^*$, $\lambda \in V^* = \mathbb{R}^m_+$,

where $\langle \cdot, \cdot \rangle$ denotes scalar product.

Proposition 2. If w is given by (3.1), then we have

(3.2a) $$\mathcal{H} \subset \operatorname{lev}_{>0} w := \{(u, v) \in \mathbb{R}^\ell \times \mathbb{R}^m : w(u, v; \theta, \lambda) > 0\}, \quad \forall \theta \in \operatorname{int} C^*, \ \forall \lambda \in V^* ,$$

(3.2b) $$\mathcal{H} = \bigcap_{\substack{\theta \in \operatorname{int} C^* \\ \lambda \in V^*}} \operatorname{lev}_{>0} w(u, v; \theta, \lambda) .$$

Proof. $(u, v) \in \mathcal{H} \iff u \in U,\ v \in V$. From this we have that, $\forall \theta \in \operatorname{int} C^*$ and $\forall \lambda \in V^*$, inequalities $\langle \theta, u\rangle > 0$ and $\langle \lambda, v\rangle \geq 0$ hold , so that $w(u, v; \theta, \lambda) > 0$, and (3.2a) follows. Because of (3.2a), to show (3.2b) it is enough to prove that no element of the complement of $\mathcal{H}$ belongs to the right–hand side of (3.2b); namely that, $\forall (\tilde{u}, \tilde{v}) \notin \mathcal{H}$, $\exists \tilde{\theta} \in \operatorname{int} C^*$ and $\exists \tilde{\lambda} \in V^*$, such that

(3.3) $$(\tilde{u}, \tilde{v}) \notin \operatorname{lev}_{>0} w(u, v; \tilde{\theta}, \tilde{\lambda}) .$$

$(\tilde{u}, \tilde{v}) \notin \mathcal{H}$ implies at least one of the following cases: (i) $\tilde{u} \notin C \setminus \{0\}$ or (ii) $\tilde{v} \notin \mathbb{R}^m_+$. If (i) holds and $\tilde{u} = 0$, then (3.3) is fulfilled by choosing any $\tilde{\theta} \in \operatorname{int} C^*$ and $\tilde{\lambda} = 0$. If (i) holds and $\tilde{u} \neq 0$, then $\tilde{u} \notin C$ and hence $\exists \tilde{\theta} \in \operatorname{int} C^*$ such that $\langle \tilde{\theta}, \tilde{u}\rangle < 0$; in fact, the inequality $\langle \theta, \tilde{u}\rangle \geq 0$, $\forall \theta \in \operatorname{int} C^*$ would imply $\tilde{u} \in (\operatorname{int} C^*)^* = (C^*)^* = C$ and lead to contradict the assumption. Then by setting, here too, $\tilde{\lambda} = 0$, the pair $(\tilde{\theta}, \tilde{\lambda})$ fulfils (3.3). If (ii) holds, then there exists (at least) an index j such that $v_j < 0$. Let us choose any $\tilde{\theta} \in \operatorname{int} C^*$, and $\tilde{\lambda} = (\tilde{\lambda}_1, \dots, \tilde{\lambda}_m)$ such that $\tilde{\lambda}_j = \alpha > 0$, $\tilde{\lambda}_i = 0\ \forall i \neq j$. Then $\langle \tilde{\theta}, \tilde{u}\rangle + \langle \tilde{\lambda}, \tilde{v}\rangle = \langle \tilde{\theta}, \tilde{u}\rangle + \alpha \tilde{v}_j$, and (3.3) holds with $\alpha \geq -\frac{1}{\tilde{v}_j} \langle \tilde{\theta}, \tilde{u}\rangle$. $\square$

[8] U and V are as in Sect.2.

From the class of separation functions (3.1) we can derive a sufficient condition, in the same way Theorem 1 has been obtained from the class (2.10).

Theorem 2. Let $y \in K$. If there exist vectors $\theta \in \operatorname{int} C^*$ and $\lambda \in V^*$, such that

$$\langle \theta, f(y) - f(x) \rangle + \langle \lambda, g(x) \rangle \leq 0 \ , \ \forall x \in X \ , \tag{3.4}$$

then y is a (global) v.m.p. of (2.1).

Proof. Because of Proposition 2 and of (3.4) we have

$$\mathcal{H} \subset \operatorname{lev}_{>0} w \quad \text{and} \quad \mathcal{K} \subseteq \operatorname{lev}_{\leq 0} w \ ,$$

respectively. Therefore, (2.8) holds. □

At $\ell = 1$ the above theorem and Theorem 1 coincide with the corresponding one of [19].

When the kind of separation (linear, nonlinear,...) has been fixed, then the outcome obviously depends on the format of the system we are referred to. For instance, system (2.5) can be equivalently split into the ℓ systems (in the unknown x):

$$\begin{aligned} &f_r(y) - f_r(x) \neq 0 \ , \ f(y) - f(x) \in C \ , g(x) \geq 0, \ x \in X \ , \\ &r \in \mathcal{I} := \{1, \ldots, \ell\} \ . \end{aligned} \tag{3.5}$$

Hence (2.2) is satisfied iff all the ℓ systems (3.5) are impossible. System (3.5) can be further split into

$$\begin{aligned} &f_r(y) - f_r(x) > 0 \ , \ f(y) - f(x) \in C, \ g(x) \geq 0 \ , \ x \in X \ , \\ &r \in \mathcal{I} \ . \end{aligned} \tag{3.6a}$$

and

$$\begin{aligned} &f_r(x) - f_r(y) > 0 \ , \ f(y) - f(x) \in C, \ g(x) \geq 0 \ , x \in X \ , \\ &r \in \mathcal{I} \ . \end{aligned} \tag{3.6b}$$

Obviously, (2.2) is satisfied iff all the 2ℓ systems (3.6) are impossible. To each of the systems (3.6) we can apply the separation scheme adopted for scalar optimization [19]. Indeed, by setting

$$K(y) := \{x \in X : f(y) - f(x) \in C,\ g(x) \geq 0\},$$

the impossibility of (3.6a) is obviously a necessary and sufficient condition for y to be a scalar global minimum point of the parametric problem:

$$\min f_r(x) \quad , \quad \text{s.t. } x \in K(y) \ , \tag{3.7a}$$

where the feasible region – which will play an important role later in the scalarization – depends (parametrically) on the unknown; we may call (3.7a) *Quasi–minimum Problem* (following the variational terminology). Analogously, the impossibility of (3.6b) is obviously a necessary and sufficient condition for y to be a scalar global maximum point of the parametric problem:

$$\max f_r(x) \quad , \quad \text{s.t. } x \in K(y), \tag{3.7b}$$

which may be called *Quasi–maximum Problem.* Consider the sets:

$$\begin{aligned}
\hat{\mathcal{H}} \ &:= \{(t,u,v) \in \mathbb{R} \times \mathbb{R}^\ell \times \mathbb{R}^m : t > 0, u \in C, v \in \mathbb{R}^m_+\} = \\
&= (\mathbb{R}_+ \backslash \{0\}) \times C \times \mathbb{R}^m_+; \\
\underline{\mathcal{K}}_r &:= \{(t,u,v) \in \mathbb{R} \times \mathbb{R}^\ell \times \mathbb{R}^m : t = f_r(y) - f_r(x), u = f(y) - f(x), \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad v = g(x), x \in X\}; \\
\bar{\mathcal{K}}_r \ &:= \{(t,u,v) \in \mathbb{R} \times \mathbb{R}^\ell \times \mathbb{R}^m : t = f_r(x) - f_r(y), u = f(y) - f(x), \\
&\qquad\qquad\qquad\qquad\qquad\qquad\qquad v = g(x), x \in X\} \ ;
\end{aligned}$$

$r \in \mathcal{I}$.

Now, observe that the systems (3.6) are all impossible iff

$$\hat{\mathcal{H}} \cap \underline{\mathcal{K}}_r = \emptyset \quad , \quad \hat{\mathcal{H}} \cap \bar{\mathcal{K}}_r = \emptyset \quad , \quad r \in \mathcal{I} \ . \tag{3.8}$$

$\forall r \in \mathcal{I}$, for each of the pairs $(\hat{\mathcal{H}}, \underline{\mathcal{K}}_r), (\hat{\mathcal{H}}, \bar{\mathcal{K}}_r)$, we can introduce a weak separation function, respectively,

$$\begin{aligned}
\underline{w}_r = \underline{w}_r(t,u,v;\underline{\theta}^r,\underline{\lambda}^r) = t + \langle \underline{\theta}^r, u\rangle + \langle \underline{\lambda}^r, v\rangle \quad , \\
\underline{\theta}^r \in C^* \ , \ \underline{\lambda}^r \in \mathbb{R}^m_+,
\end{aligned} \tag{3.9a}$$

and

$$\bar{w}_r = \bar{w}_r(t,u,v;\bar{\theta}^r,\bar{\lambda}^r) = t + \langle\bar{\theta}^r,u\rangle + \langle\bar{\lambda}^r,v\rangle \ , \quad \bar{\theta}^r \in C^* \ , \ \bar{\lambda}^r \in \mathbb{R}^m_+ \ . \tag{3.9b}$$

It can be shown that (see [19]):

Proposition 3. If $\forall r \in \mathcal{I}$, $\underline{w}_r$ and $\bar{w}_r$ are given by (3.9), then we have:

$$\hat{\mathcal{H}} \subset \mathrm{lev}_{>0}\ \underline{w}_r(t,u,v;\underline{\theta}^r,\underline{\lambda}^r) \quad , \quad \hat{\mathcal{H}} = \bigcap_{\substack{\underline{\theta}^r \in C^* \\ \underline{\lambda}^r \in \mathbb{R}^m_+}} \mathrm{lev}_{>0}\ \underline{w}_r(t,u,v;\underline{\theta}^r,\underline{\lambda}^r),$$

$$\hat{\mathcal{H}} \subset \mathrm{lev}_{>0}\ \bar{w}_r(t,u,v;\bar{\theta}^r,\bar{\lambda}^r) \quad , \quad \hat{\mathcal{H}} = \bigcap_{\substack{\bar{\theta}^r \in C^* \\ \bar{\lambda}^r \in \mathbb{R}^m_+}} \mathrm{lev}_{>0}\ \bar{w}_r(t,u,v;\bar{\theta}^r,\bar{\lambda}^r),$$

where all the level sets are meant with respect to t,u,v.

Now, we can state an optimality condition. Take into account that polarity and Cartesian product are permutable.

Theorem 3. Let $y \in K$. If $\forall r \in \mathcal{I}$, there exist $\underline{\theta}^r, \bar{\theta}^r \in C^*$ and $\underline{\lambda}^r, \bar{\lambda}^r \in \mathbb{R}^m_+$, such that:

$$f_r(y) - f_r(x) + \langle\underline{\theta}^r, f(y)-f(x)\rangle + \langle\underline{\lambda}^r, g(x)\rangle \leq 0 \ , \ \ \forall x \in X \ , \tag{3.10a}$$

$$f_r(x) - f_r(y) + \langle\bar{\theta}^r, f(y)-f(x)\rangle + \langle\bar{\lambda}^r, g(x)\rangle \leq 0, \ \forall x \in X, \tag{3.10b}$$

then y is a (global) v.m.p. of (2.1).

Proof. Because of Proposition 3 and of (3.10), $\forall r \in \mathcal{I}$ we have:

$$\hat{\mathcal{H}} \subset \mathrm{lev}_{>0}\ \underline{w}_r \quad , \quad \underline{\mathcal{K}} \subseteq \mathrm{lev}_{\leq 0}\ \underline{w}_r \ ,$$
$$\hat{\mathcal{H}} \subset \mathrm{lev}_{>0}\ \bar{w}_r \quad , \quad \bar{\mathcal{K}} \subseteq \mathrm{lev}_{\leq 0}\ \bar{w}_r \ .$$

Therefore, (3.8) hold. $\square$

It is interesting to note that, $\forall r \in \mathcal{I}$, summing up side by side (3.10a) and (3.10b) leads to

$$\langle \underline{\theta}^r + \bar{\theta}^r \, , \, f(y) - f(x) \rangle + \langle \underline{\lambda}^r + \bar{\lambda}^r, g(x) \rangle \leq 0 \ , \ \forall x \in X \ ,$$

which is equivalent to (3.4), provided that $\underline{\theta}^r + \bar{\theta}^r \in \operatorname{int} C^*$. Moreover, the matrices $\bar{\Theta}$ and $\bar{\Lambda}$, whose $r-$th rows are given, respectively, by $\underline{\theta}^r + \bar{\theta}^r$ and $\underline{\lambda}^r + \bar{\lambda}^r$, allow us to find, within the class (2.14), a vector separation function $w(u, v; \bar{\Theta}, \bar{\Lambda})$ which fulfils (2.13).

When $C = \mathbb{R}^\ell_+$, then the ℓ systems (3.6b) are impossible $\forall y$, so that the impossibility of all the 2ℓ systems (3.6) is equivalent to that of the following ℓ systems (in the unknown x):

$$(3.6)' \qquad \begin{array}{l} f_r(y) - f_r(x) > 0 \ , \ f(y) - f(x) \geq 0 \ , \ g(x) \geq 0 \ , \ x \in X \ , \\ r \in \mathcal{I} \ . \end{array}$$

Consequently, (3.7b), $\bar{\mathcal{K}}_r$ and $\bar{w}_r$ disappear. Since in the present case the condition $f(y) - f(x) \in C$ is split into the system $f_s(y) - f_s(x) \geq 0$, $s \in \mathcal{I}$, then the $r-$ th of these inequalities becomes redundant and can be deleted. Therefore, $\underline{\mathcal{K}}^r$ can be replaced by $\mathcal{K}$ itself and, with the introduction of $\mathcal{I}_r := \mathcal{I} \backslash \{r\}$ $\forall r \in \mathcal{I}$, the set $\hat{\mathcal{H}}$, conditions (3.8) and functions (3.9a) can be replaced, respectively, by[9]

$$\mathcal{H}_r := \{(u, v) \in \mathbb{R}^\ell \times \mathbb{R}^m : u_r > 0, \ u_s \geq 0 \quad s \in \mathcal{I}_r, \ v \geq 0\},$$

$$(3.8)' \qquad \mathcal{H}_r \cap \mathcal{K} = \emptyset \quad , \quad r \in \mathcal{I}$$

and

$$(3.9)' \qquad \begin{array}{r} w_r = w_r(u, v; \theta^r_s \ s \in \mathcal{I}_r, \lambda^r) := u_r + \displaystyle\sum_{s \in \mathcal{I}_r} \theta^r_s u_s + \langle \lambda^r, v \rangle, \\ \theta^r_s \geq 0 \quad s \in \mathcal{I}_r, \ \lambda^r \in \mathbb{R}^m_+ \ , \end{array}$$

for each $r \in \mathcal{I}$. Hence, Theorem 3 becomes:

Corollary 1. Let $y \in K$. Assume that, $\forall r \in \mathcal{I}$, there exist $\theta^r_s \geq 0 \ \ s \in \mathcal{I}_r$ and $\lambda^r \in \mathbb{R}^m_+$, such that

$$(3.10)' \qquad \begin{array}{l} f_r(y) - f_r(x) + \displaystyle\sum_{s \in \mathcal{I}_r} \theta^r_s [f_s(y) - f_s(x)] + \\ \qquad\qquad + \langle \lambda^r, g(x) \rangle \leq 0 \ , \ \forall x \in X \ , \end{array}$$

[9] Recall that u_r, and u_s $s \in \mathcal{I}_r$ are the ℓ elements of u.

for each $r \in \mathcal{I}$. Then, y is a (global) v.m.p. of (2.1) with $C = \mathbb{R}^{\ell}_{+}$.

Remark 1. For every fixed $r \in \mathcal{I}$, (3.10)$'$ ensures a scalar separation with the property that the $r-$th component of the multiplier θ^r is strictly positive. Therefore, summing up side by side the ℓ inequalities (3.10)$'$, we obtain a (scalar) separation function of the class (3.1).

Of course, Theorem 3 can be simplified in cases other than $C = \mathbb{R}^{\ell}_{+}$. If $C \subseteq \mathbb{R}^{\ell}_{+}$ (resp., $C \subseteq \mathbb{R}^{\ell}_{-}$), then system (3.6b) (resp., (3.6a)) is impossible and disappears.

In Sect. 2 the image space associated with (2.1) was $(\ell + m)$ – dimensional; the same happened with the approach which led to Theorem 2. While, having turned (2.5) into (3.5), the associated image space became $(1 + \ell + m)$ – dimensional. Such an increase of dimensionality can be avoided by replacing the cone C with the cone $\hat{C}_r := C \cap \{u \in \mathbb{R}^{\ell} : u_r \neq 0\}$. However, the advantage due to the decrease of dimensionality is balanced by the fact that we would be faced with $\mathcal{H}_r := \hat{C}_r \times \mathbb{R}^m_{+}$ (while $\mathcal{K}$ would remain as in Sect. 2) which now might be not convex, because of the possible lack of convexity of $\hat{C}_r$. Then the introduction of a weak separation function would be difficult.

4. SEPARATION IN THE WEAK CASE

The separation approach, which has been developed in Sect. 2 for problem (2.1), can be defined for other kinds of problems, in particular for problem (2.3). This will be now briefly outlined.

Obviously, (2.4) is satisfied iff the system (in the unknown x):

$$f(y) - f(x) \geq_{\operatorname{int} C} 0 \quad , \quad g(x) \geq 0 \quad , \quad x \in X \tag{4.1}$$

is impossible. The sets $\mathcal{H}$ and U of Sect. 2 are, in the present section, replaced by

$$\mathcal{H} := \{(u, v) \in \mathbb{R}^{\ell} \times \mathbb{R}^m : u \geq_{\operatorname{int} C} 0, v \geq 0\} = (\operatorname{int} C) \times \mathbb{R}^m_{+}$$

and $U := \operatorname{int} C$, respectively; while $\mathcal{K}$ and V are the same as in Sect. 2. In place of (2.10), we now consider the function $w : \mathbb{R}^{\ell} \times \mathbb{R}^m \to$

$\mathbb{R}^\ell$, given by[10]:

$$w = w(u,v;\Theta,\Lambda) = \Theta u + \Lambda v \quad , \quad \Theta \in U^*_{\mathrm{int}\, C} \ , \ \Lambda \in V^*_C \ , \tag{4.2}$$

where Θ and Λ are the same parameters as in Sect.2. Consider the "positive level set" of (4.2):

$$W_{\mathrm{int}\, C}(u,v;\Theta,\Lambda) := \{(u,v) \in \mathbb{R}^\ell \times \mathbb{R}^m : w(u,v;\Theta,\Lambda) \geq_{\mathrm{int}\, C} 0\}.$$

Proposition 4. If w is given by (4.2), then we have

$$\mathcal{H} \subset W_{\mathrm{int}\, C}(u,v;\Theta,\Lambda), \quad \forall \Theta \in U^*_{\mathrm{int}\, C}, \quad \forall \Lambda \in V^*_C \ ; \tag{4.3a}$$

$$\mathcal{H} = \bigcap_{\substack{\Theta \in U^*_{\mathrm{int}\, C} \\ \Lambda \in V^*_C}} W_{\mathrm{int}\, C}(u,v;\Theta,\Lambda) \ . \tag{4.3b}$$

Proof. $(u,v) \in \mathcal{H} \Leftrightarrow u \in U, v \in V$. This implies that, $\forall \Theta \in U^*_{\mathrm{int}\, C}$ and $\forall \Lambda \in V^*_C$, we have $\Theta u + \Lambda v \in \mathrm{int}\, C$, and thus (4.3a) follows. Because of (4.3a), to show (4.3b) it is enough to prove that no element of the complement of $\mathcal{H}$ belongs to the right–hand side of (4.3b); namely that, $\forall (\tilde{u},\tilde{v}) \notin \mathcal{H}$, $\exists \tilde{\Theta} \in U^*_{\mathrm{int}\, C}$ and $\exists \tilde{\Lambda} \in V^*_C$, such that

$$(\tilde{u},\tilde{v}) \notin W_{\mathrm{int}\, C}(u,v;\tilde{\Theta},\tilde{\Lambda}) \ . \tag{4.4}$$

$(\tilde{u},\tilde{v}) \notin \mathcal{H}$ implies at least one of the following cases: (i) $\tilde{u} \notin \mathrm{int}\, C$ or (ii) $\tilde{v} \notin \mathbb{R}^m_+$. Since now the proof is the same as that of Proposition 1, by merely changing $C \backslash \{0\}$ into $\mathrm{int}\, C$. □

Theorem 4. Let $y \in K$. If there exist matrices $\Theta \in U^*_{\mathrm{int}\, C}$ and $\Lambda \in V^*_C$, such that

$$\Theta[f(y) - f(x)] + \Lambda g(x) \not\geq_{\mathrm{int}\, C} 0 \ , \quad \forall x \in X \ , \tag{4.5}$$

then y is a (global) v.m.p. of (2.3).

Proof. Because of Proposition 4 and of (4.5) we have

$$\mathcal{H} \subset W_{\mathrm{int}\, C}(u,v;\Theta,\Lambda) \ \text{ and } \ \mathcal{K} \subseteq \sim W_{\mathrm{int}\, C}(u,v;\Theta,\Lambda) \ ,$$

respectively. Therefore, $\mathcal{H} \cap \mathcal{K} = \emptyset$ which is (here too) a necessary and sufficient condition for (2.4) to hold. □

---

[10] Without any fear of confusion, we use here the notation w as in Sect. 2.

At $\ell = 1$ Theorems 4 and 1 coincide. At $C = \mathbb{R}^\ell_+$, (4.5) is a sufficient optimality condition for y to be a so–called weak Pareto v.m.p.; in this case, the class of separation functions (4.2) can be "scalarized" by choosing the scalar function within the class of the so–called min–functions [11]; we restrict ourselves to the linear case. To this end, consider the function $w^0 : \mathbb{R}^\ell \times \mathbb{R}^m \to \mathbb{R}$, given by

$$w^0 = w^0(u, v; \Theta, \Lambda) := \min\{\langle \theta^r, u\rangle + \langle \lambda^r, v\rangle, r \in \mathcal{I}\}, \qquad \Theta \in U^*_{\mathrm{int}\ \mathbb{R}^\ell_+}\ ,\ \Lambda \in V^*_{\mathbb{R}^\ell_+}\ , \tag{4.6}$$

where θ^r and λ^r are the r–th rows of Θ and Λ, respectively. It is easy to see that (w is given by (4.2)):

$$\sim W_{\mathrm{int}\, C}(u, v; \Theta, \Lambda) = \mathrm{lev}_{\le 0}\ w^0(u, v; \Theta, \Lambda) := \{(u, v) \in \mathbb{R}^\ell \times \mathbb{R}^m : w^0(u, v; \Theta, \Lambda) \le 0\}\ ,$$

which shows that the sufficient condition (4.5) can be scalarized by replacing w with w^0 and $\not\geq_{\mathrm{int}\ \mathbb{R}^\ell_+}$ with merely $\le$.

5. NECESSARY OPTIMALITY CONDITIONS

In this section it will be shown that the separation approach in the image space can be considered as a source for deriving also necessary optimality conditions for the several kinds of VOP (sufficient ones have been obtained in Sects.2–4). More precisely, following the approach developed in [20] for the scalar case, necessary conditions will be obtained by separating $\mathcal{H}$ from a suitable approximation of the image $\mathcal{K}$. Here, such an approximation will be obtained by assuming G–semidifferentiability[11] for f and g and then by replacing them with their G–semiderivatives.

Assume that[12] $X = \mathbb{R}^n$ and denote by G a given subset of the set, say $\mathcal{G}$, of positively homogeneous functions of degree 1 on $X - y$.

Definition 1. A function $\varphi : X \to \mathbb{R}$ is said *lower G–semidifferentiable at $y \in X$* iff there exist functions $\underline{D}_G\varphi : X \times X \to \mathbb{R}$ and $\epsilon_\varphi : X \times X \to \mathbb{R}$, such that:

[11] It will be briefly recalled; see [20] for details.

[12] X can be merely a truncated cone with apex at y.

(i) $\underline{\mathcal{D}}_G\varphi(y;\cdot) \in G$;

(ii) $\varphi(x) - \varphi(y) = \underline{\mathcal{D}}_G\varphi(y; x-y) + \epsilon_\varphi(y; x-y)$, $\forall x \in X$;

(3i) $\liminf\limits_{x\to y} \frac{\epsilon_\varphi(y;x-y)}{\|x-y\|} \geq 0$;

(4i) for every pair (h, ϵ) of functions (of the same kind, respectively, of $\underline{\mathcal{D}}_G\varphi$ and ϵ_φ), which satisfy (i)–(3i), we have $\operatorname{epi} h \supseteq \operatorname{epi} \underline{\mathcal{D}}_G\varphi$.

$\underline{\mathcal{D}}_G\varphi(y; \frac{x-y}{\|x-y\|})$ is called the *lower G–semiderivative of* φ *at* y.
An analogous definition is given for an *upper G–semidifferentiable* function, whose *upper G–semiderivative at* y is denoted by

$$\bar{\mathcal{D}}_G\varphi(y; \frac{x-y}{\|x-y\|}).$$

Let us introduce the index sets $J := \{1, \ldots, m\}$ and $J^0(y) := \{i \in J : g_i(y) = 0,\ \epsilon_{g_i}(y; x-y) \not\equiv 0\}$. In the present section we will consider the particular case $C = \mathbb{R}^\ell_+$, namely the Pareto case. Next lemma is a generalization of the classic Linearization Lemma of Abadie [1].

Lemma 1. Let the functions f_i $i \in \mathcal{I}$ be upper Φ–semidifferentiable at y and g_i $i \in J$ lower Γ–semidifferentiable at y, where $\Phi, \Gamma \subseteq \mathcal{G}$. If y is a v.m.p. of (2.1), then the system (in the unknown x):

$$\text{(5.1)} \qquad \begin{cases} \bar{\mathcal{D}}_\Phi f_i(y; x-y) < 0 \ , & i \in \mathcal{I}, \\ \underline{\mathcal{D}}_\Gamma g_i(y; x-y) > 0 \ , & i \in J^0(y), \\ \underline{\mathcal{D}}_\Gamma g_i(y; x-y) \geq 0 \ , & i \in J \backslash J^0(y), \\ x \in X \end{cases}$$

is impossible.

Proof. The assumption that y be v.m.p. is equivalent to the impossibility of (2.5) which, due to the semidifferentiability of f and g, becomes:

$$\text{(5.2)} \qquad \begin{cases} -(\bar{\mathcal{D}}_\Phi f_i(y; x-y) + \epsilon_{f_i}(y; x-y), i \in \mathcal{I}) \in C \backslash \{0\}, \\ g_i(y) + \underline{\mathcal{D}}_\Gamma g_i(y; x-y) + \epsilon_{g_i}(y; x-y) \geq 0, \quad i \in J, \\ x \in X \ . \end{cases}$$

Ab absurdo, suppose that (5.1) be possible, and let $\hat{x}$ be a solution of (5.1). Then $x(\alpha) := (1-\alpha)y+\alpha\hat{x}$ is a solution of (5.1) $\forall \alpha \in]0,1]$, due to the positive homogeneity of the semiderivatives. The remainders satisfy the inequalities:

$$\limsup_{x\to y} \frac{\epsilon_{f_i}(y; x-y)}{\|x-y\|} \le 0,\ i \in \mathcal{I};\quad \liminf_{x\to y} \frac{\epsilon_{g_i}(y; x-y)}{\|x-y\|} \ge 0,\ i \in J\ .$$

From the definitions of lim sup and lim inf, for every fixed $\delta > 0$, $\exists \alpha_\delta > 0$ such that $\forall \alpha \in]0, \alpha_\delta]$, we have:

$$\text{(5.3a)}\qquad \frac{\epsilon_{f_i}(y; \alpha(\hat{x}-y))}{\|\alpha(\hat{x}-y)\|} \le \delta \quad , \quad i \in \mathcal{I},$$

$$\text{(5.3b)}\qquad \frac{\epsilon_{g_i}(y; \alpha(\hat{x}-y))}{\|\alpha(\hat{x}-y)\|} \ge -\delta \quad , \quad i \in J^0(y).$$

Because of the positive homogeneity of $\bar{\mathcal{D}}_\Phi f_i$ and $\underline{\mathcal{D}}_\Gamma g_i$, we have:

$$\text{(5.4a)}\qquad \bar{\mathcal{D}}_\Phi f_i(y; \alpha(\hat{x}-y)) + \epsilon_{f_i}(y; \alpha(\hat{x}-y)) =$$
$$= \left[\frac{\bar{\mathcal{D}}_\Phi f_i(y; \hat{x}-y)}{\|\hat{x}-y\|} + \frac{\epsilon_{f_i}(y; \alpha(\hat{x}-y))}{\alpha\|\hat{x}-y\|} \right] \cdot \alpha\|\hat{x}-y\|,\ i \in \mathcal{I}$$

$$\text{(5.4b)}\qquad \underline{\mathcal{D}}_\Gamma g_i(y; \alpha(\hat{x}-y)) + \epsilon_{g_i}(y; \alpha(\hat{x}-y)) =$$
$$= \left[\frac{\underline{\mathcal{D}}_\Gamma g_i(y; \hat{x}-y)}{\|\hat{x}-y\|} + \frac{\epsilon_{g_i}(y; \alpha(\hat{x}-y))}{\alpha\|\hat{x}-y\|} \right] \cdot \alpha\|\hat{x}-y\|,\ i \in J^0(y).$$

From (5.4), taking into account (5.3), we draw the existence of $\delta > 0$ and $\alpha_\delta > 0$, such that $\forall \alpha \in]0, \alpha_\delta]$ we have:

$$\bar{\mathcal{D}}_\Phi f_i(y; \alpha(\hat{x}-y)) + \epsilon_{f_i}(y; \alpha(\hat{x}-y)) < 0,\ i \in \mathcal{I},$$
$$\underline{\mathcal{D}}_\Gamma g_i(y; \alpha(\hat{x}-y)) + \epsilon_{g_i}(y; \alpha(\hat{x}-y)) > 0,\ i \in J^0(y).$$

When $i \in J \backslash J^0(y)$, there are two cases: $g_i(y) = 0$ and $\epsilon_{g_i} \equiv 0$; or $g_i(y) > 0$. In the latter, choosing α small enough, we have

$$\text{(5.5)}\qquad \underline{\mathcal{D}}_\Gamma g_i(y; \alpha(\hat{x}-y)) + g_i(y) + \epsilon_{g_i}(y; \alpha(\hat{x}-y)) \ge 0;$$

while in the former the inequality (5.5) is fulfilled $\forall \alpha \ge 0$. Therefore $\exists \bar{\alpha} \in]0, \alpha_\delta]$ such that $x(\bar{\alpha})$ is a solution of (5.2), and this fact contradicts the hypothesis. □

The impossibility of system (5.1) is a 1st order optimality condition for (2.1), which is expressed in terms of the sign of semiderivatives. Now, we will give it a “multiplier form”. To this end, carrying on the image space analysis of Sect.2, let us introduce the sets:

$$\mathcal{H}_G := \{(u,v) \in \mathbb{R}^\ell \times \mathbb{R}^m : u \in \text{ int } \mathbb{R}^\ell_+; v_i > 0,\ i \in J^0(y); \\ v_i \geq 0,\ i \in J \backslash J^0(y)\}$$

$$\mathcal{K}_G := \{(u,v) \in \mathbb{R}^\ell \times \mathbb{R}^m : u_i = -\bar{\mathcal{D}}_\Phi f_i(y; x-y),\ i \in \mathcal{I}; \\ v_i = \underline{\mathcal{D}}_\Gamma g_i(y; x-y),\ i \in J;\ x \in X\}$$

so that (5.1) is impossible iff $\mathcal{H}_G \cap \mathcal{K}_G = \emptyset$. Let $\mathcal{C}$ be the set of sublinear functions[13] on $X - y$.

Theorem 5. Suppose that f_i $i \in \mathcal{I}$ are upper Φ–semidifferentiable functions at y, with $\Phi \subseteq \mathcal{C}$; suppose that g_i $i \in J$ are lower Γ–semidifferentiable functions at y, with $\Gamma \subseteq (-\mathcal{C})$. If y is an optimal solution of (2.1), then there exists a vector $(\theta, \lambda) \in \mathbb{R}^\ell_+ \times \mathbb{R}^m_+$, with $(\theta, \lambda) \neq 0$, such that:

(i) $$\sum_{i \in \mathcal{I}} \theta_i \bar{\mathcal{D}}_\Phi f_i(y; x-y) - \sum_{i \in J} \lambda_i \underline{\mathcal{D}}_\Gamma g_i(y; x-y) \geq 0,\ \forall x \in X;$$

(ii) $$\sum_{i \in J} \lambda_i g_i(y) = 0\ .$$

Proof. By the assumption the functions $-\bar{\mathcal{D}}_\Phi f_i$ $i \in \mathcal{I}$, and $\underline{\mathcal{D}}_\Gamma g_i$ $i \in J$, are concave so that the set $\mathcal{E}_G := \mathcal{K}_G - c\ell \mathcal{H}_G$ is convex (see [19]). Moreover, it has been proved (see [6]) that $\mathcal{H}_G \cap \mathcal{K}_G = \emptyset$ iff $\mathcal{H}_G \cap \mathcal{E}_G = \emptyset$. Therefore, by Lemma 1, the optimality of y implies that $\mathcal{H}_G \cap \mathcal{E}_G = \emptyset$. Since $\mathcal{H}_G$ and $\mathcal{E}_G$ are convex sets, there exists a vector $(\theta, \lambda) \in \mathbb{R}^\ell \times \mathbb{R}^m$, with $(\theta, \lambda) \neq 0$, such that:

$$\sum_{i \in \mathcal{I}} \theta_i(-\bar{\mathcal{D}}_\Phi f_i(y; x-y)) + \sum_{i \in J^0(y)} \lambda_i \underline{\mathcal{D}}_\Gamma g_i(y; x-y) \\ + \sum_{i \in J \backslash J^0(y)} \lambda_i(\underline{\mathcal{D}}_\Gamma g_i(y; x-y) + g_i(y)) \leq 0,\ \forall x \in X. \tag{5.6}$$

---

[13] i.e., convex and positively homogeneous of degree one.

It is trivial to prove that $(\theta, \lambda) \in \mathbb{R}^{\ell}_{+} \times \mathbb{R}^{m}_{+}$, since $\langle \theta, u\rangle + \langle \lambda, v\rangle \geq 0$ $\forall (u,v) \in \mathcal{H}_G$. From (5.6) we obtain:

$$\begin{aligned} \sum_{i\in\mathcal{I}} \theta_i \bar{\mathcal{D}}_\Phi f_i(y; x-y) - \sum_{i\in J} \lambda_i \underline{\mathcal{D}}_\Gamma g_i(y; x-y) \geq \\ \sum_{i\in J\setminus J^0(y)} \lambda_i g_i(y),\ \forall x \in X. \end{aligned} \tag{5.7}$$

Computing (5.7) for $x = y$, we obtain $\sum_{i\in J\setminus J^0(y)} \lambda_i g_i(y) = 0$ and therefore (i). Taking into account that $g_i(y) = 0\ \forall i \in J^0(y)$, we have (ii) and hence the thesis. □

We recall (see [20]) that, when $G \subseteq \mathcal{C}$, the generalized subdifferential of a lower (or upper) G–semidifferentiable function φ at y, denoted by $\partial_G \varphi(y)$, is defined as the subdifferential of the convex function $\underline{\mathcal{D}}_G \varphi(y; x-y)$ (or $\bar{\mathcal{D}}_G \varphi(y; x-y)$).

Theorem 6. Let $\Phi \subseteq \mathcal{C}$ and $\Gamma = -\Phi$. Suppose that f_i $i \in \mathcal{I}$ are upper Φ–semidifferentiable functions and g_i $i \in J$ are lower Γ–semidifferentiable functions at $y \in X$. If y is a v.m.p. of (2.1), then there exists $(\theta, \lambda) \in \mathbb{R}^{\ell}_{+} \times \mathbb{R}^{m}_{+}$, with $(\theta, \lambda) \neq 0$, such that:

$$\begin{cases} 0 \in \partial_\Phi L(x; \theta, \lambda),\ \langle \lambda, g(x)\rangle = 0 \\ \lambda \geq 0,\ \theta \geq 0,\ g(x) \geq 0,\ x \in X. \end{cases} \tag{5.8}$$

Proof. We recall that $-\underline{\mathcal{D}}_\Gamma g = \bar{\mathcal{D}}_\Phi(-g)$, so that the function $\langle \theta, f\rangle - \langle \lambda, g\rangle$ is upper Φ–semidifferentiable at y. The following inequalities hold:

$$\begin{aligned} &\bar{\mathcal{D}}_\Phi(\langle \theta, f\rangle - \langle \lambda, g\rangle)(y; x-y) \geq \\ &\langle \theta, \bar{\mathcal{D}}_\Phi f(y; x-y)\rangle + \langle \lambda, \bar{\mathcal{D}}_\Phi(-g)(y; x-y)\rangle \geq 0,\ \forall x \in X. \end{aligned}$$

The first inequality follows from Theorem 3.3 of [33], while the second from Theorem 5. The previous relation implies that $0 \in \partial_\Phi L(y; \theta, \lambda)$; from Theorem 5 we have the complementarity relation $\langle \lambda, g(y)\rangle = 0$, and the proof is complete. □

The system (5.8) is a generalization to nondifferentiable VOP of the well–known John condition [1,29].

Theorem 7. Let f_i $i \in \mathcal{I}$, and g_i $i \in J$ be differentiable functions at y. If y is a v.m.p. of (2.1), then there exists $(\theta, \lambda) \in \mathbb{R}^\ell_+ \times \mathbb{R}^m_+$, with $(\theta, \lambda) \neq 0$, such that:

$$\begin{cases} \sum\limits_{i\in\mathcal{I}} \theta_i \nabla f_i(y) - \sum_{i\in J} \lambda_i \nabla g_i(y) = 0, \ \langle \lambda, g(y) \rangle = 0, \\ \theta \geq 0, \ \lambda \geq 0, \ g(y) \geq 0, \ y \in X. \end{cases}$$

Proof. Let G be the set of the linear functions. With $\Phi = \Gamma = G$, we have that the hypotheses of Theorem 5 are fulfilled, and it is known [19] that

$$\bar{\mathcal{D}}_G f_i(y; x-y) = \langle \nabla f_i(y), \frac{x-y}{\|x-y\|} \rangle, \ i \in \mathcal{I}$$

$$\underline{\mathcal{D}}_G g_i(y; x-y) = \langle \nabla g_i(y), \frac{x-y}{\|x-y\|} \rangle, \ i \in J.$$

Therefore, there exists $(\theta, \lambda) \in \mathbb{R}^\ell_+ \times \mathbb{R}^m_+$, with $(\theta, \lambda) \neq 0$, such that

$$\sum_{i\in\mathcal{I}} \theta_i \langle \nabla f_i(y), \frac{x-y}{\|x-y\|} \rangle - \sum_{i\in J} \lambda_i \langle \nabla g_i(y), \frac{x-y}{\|x-y\|} \rangle \geq 0, \quad \forall x \in X$$

.
The previous condition is equivalent to the following:

$$\langle \sum_{i\in\mathcal{I}} \theta_i \nabla f_i(y) - \sum_{i\in J} \lambda_i \nabla g_i(y) \ , \ z \rangle \geq 0,$$

$\forall z \in \mathbb{R}^n$, with $\|z\| = 1$. Since z is arbitrary, we obtain:

$$\sum_{i\in\mathcal{I}} \theta_i \nabla f_i(y) - \sum_{i\in J} \lambda_i \nabla g_i(y) = 0.$$

From (ii) of Theorem 5 we have that $\langle \lambda, g(y) \rangle = 0$ and this completes the proof. $\square$

Under suitable regularity assumptions on (2.1), which ensure that the vector of multipliers is different from the zero vector in $\mathbb{R}^\ell$, the system (5.8) is a generalization to nondifferentiable VOP of the classic Kuhn–Tucker conditions. Similar results have been obtained in [8] using the Clarke subdifferential but under the more general assumptions that the feasible region be defined by means of equality and inequality constraints and the ordering cone be not necessarily the positive orthant.

6. SADDLE POINT CONDITIONS

Like in the scalar case [19], also in the vector case the saddle point type conditions can be derived from the separation scheme. More precisely, we will show that the sufficient condition (2.13) can be equivalently put in a saddle point format. At first we recall the definition of a saddle point for a vector function.

Definition 2. Let $F : X \times Y \to \mathbb{R}^\ell$. $(\bar{x}, \bar{y})$ is a *saddle point* for F on $X \times Y$ iff

$$F(x, \bar{y}) \not\leq_{C\setminus\{0\}} F(\bar{x}, \bar{y}) \not\leq_{C\setminus\{0\}} F(\bar{x}, y) \ , \quad \forall x \in X, \quad \forall y \in Y.$$

Let us introduce the (generalized) *vector Lagrangian function* $\mathcal{L} : \mathbb{R}^n \times \mathbb{R}^{\ell\times\ell} \times \mathbb{R}^{\ell\times m} \to \mathbb{R}^\ell$, defined by

$$\mathcal{L}(x; \Theta, \Lambda) := \Theta f(x) - \Lambda g(x) \ . \tag{6.1}$$

Definition 3. A pair $(\bar{x}, \bar{\Lambda}) \in X \times V_C^*$ is a *generalized vector saddle point* of (6.1) on $X \times V_C^*$ iff there exists $\bar{\Theta} \in U_{C\setminus\{0\}}^*$ such that:

$$\mathcal{L}(x; \bar{\Theta}, \bar{\Lambda}) \not\leq_{C\setminus\{0\}} \mathcal{L}(\bar{x}; \bar{\Theta}, \bar{\Lambda}) \not\leq_{C\setminus\{0\}} \mathcal{L}(\bar{x}; \bar{\Theta}, \Lambda), \tag{6.2}$$

$\forall x \in X, \quad \forall \Lambda \in V_C^*$.

Now let us consider the class of separation functions (2.10), and state the following:

Theorem 8. Let $y \in K$. The following statements are equivalent:
(i) There exists $(\bar{\Theta}, \bar{\Lambda}) \in U^*_{C\setminus\{0\}} \times V^*_C$ such that:

$$w(u, v; \bar{\Theta}, \bar{\Lambda}) \ngeq_{C\setminus\{0\}} 0 \ , \quad \forall (u, v) \in \mathcal{K}(y);$$

(ii) $(y, \bar{\Lambda}) \in X \times V^*_C$ is a generalized vector saddle point for $\mathcal{L}(x; \bar{\Theta}, \Lambda)$ on $X \times V^*_C$.

Proof. (i) $\Rightarrow$ (ii). (i) is equivalent to the condition

$$\bar{\Theta}[f(y) - f(x)] + \bar{\Lambda} g(x) \ngeq_{C\setminus\{0\}} 0 \ , \quad \forall x \in X \ . \tag{6.3}$$

By setting $x = y$ in (6.3), we obtain $\bar{\Lambda} g(y) \ngeq_{C\setminus\{0\}} 0$. Since $\bar{\Lambda} \in V^*_C$ we have that $\bar{\Lambda} g(y) \geq_C 0$ and therefore

$$\bar{\Lambda} g(y) = 0 \ . \tag{6.4}$$

Taking into account (6.4), condition (6.3) is equivalent to

$$\mathcal{L}(y; \bar{\Theta}, \bar{\Lambda}) \ngeq_{C\setminus\{0\}} \mathcal{L}(x; \bar{\Theta}, \bar{\Lambda}) \ , \quad \forall x \in X,$$

that is y is a v.m.p. of $\mathcal{L}(x; \bar{\Theta}, \bar{\Lambda})$ on X. We have to show that $\bar{\Lambda}$ is a v.m.p. for $-\mathcal{L}(y; \bar{\Theta}, \Lambda)$ on V^*_C. We see that:

$$\mathcal{L}(y; \bar{\Theta}, \Lambda) = \bar{\Theta} f(y) - \Lambda g(y) \ \text{ so that } \ -\mathcal{L}(y; \bar{\Theta}, \bar{\Lambda}) = -\bar{\Theta} f(y) \ .$$

For every $\Lambda \in V^*_C$ it is $\Lambda g(y) \geq_C 0$ and therefore

$$-\mathcal{L}(y; \bar{\Theta}, \bar{\Lambda}) + \mathcal{L}(y; \bar{\Theta}, \Lambda) = -\Lambda g(y) \ngeq_{C\setminus\{0\}} 0 \ , \quad \forall \Lambda \in V^*_C,$$

since C is a pointed cone.
(ii) $\Rightarrow$ (i). From the condition

$$-\mathcal{L}(y; \bar{\Theta}, \bar{\Lambda}) + \mathcal{L}(y; \bar{\Theta}, \Lambda) \ngeq_{C\setminus\{0\}} 0 \ , \quad \forall \Lambda \in V_C$$

computed for Λ equal to the null matrix, we obtain $\bar{\Lambda} g(y) \ngeq_{C\setminus\{0\}} 0$ and, since $\bar{\Lambda} \in V^*_C$, we have (6.4). As in the proof of the reverse implication, exploiting the complementarity relation (6.4), we have that the condition

$$\mathcal{L}(y; \bar{\Theta}, \bar{\Lambda}) \ngeq_{C\setminus\{0\}} \mathcal{L}(x; \bar{\Theta}, \bar{\Lambda}) \ , \quad \forall x \in X$$

is equivalent to (6.3), that is (i). □

Remark 2. We observe that, in the statement of Theorem 8, the set $U^*_{C\setminus\{0\}}$ can be replaced by any subset of the $s \times \ell$ matrices, provided that C be a closed and convex cone in $\mathbb{R}^s$.

Example 3. Let $f_1(x) = 2x_1 + x_2$, $f_2(x) = x_1 + 2x_2$, $g(x) = x_1 + x_2 - 2$, $X = \mathbb{R}^2_+$, $C = \mathbb{R}^2_+$. It is easy to prove that $\hat{K} := \{x \in X : x_1 + x_2 = 2\}$ is the set of v.m.p. of (2.1). Let $\bar{x} = (a, b)$ with $a + b = 2$, $a, b \geq 0$. Since the present VOP is linear, there exist $\bar{\Theta} \in U^*_{C\setminus\{0\}}$ and $\bar{\Lambda} \in V^*_C$ such that $(\bar{x}; \bar{\Lambda})$ is a vector saddle point of $\mathcal{L}(x; \bar{\Theta}, \Lambda)$ on $X \times V^*_C$. Put $\bar{\Theta} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ and $\Lambda = \begin{pmatrix} 3/2 \\ 3/2 \end{pmatrix}$. The first inequality in (6.2) becomes:

$$(6.5) \quad \begin{pmatrix} 2a + b \\ a + 2b \end{pmatrix} \not\geq_{C\setminus\{0\}} \begin{pmatrix} 2x_1 + x_2 - \frac{3}{2}(x_1 + x_2 - 2) \\ x_1 + 2x_2 - \frac{3}{2}(x_1 + x_2 - 2) \end{pmatrix}, \forall x \in X.$$

Consider the system

$$(6.6) \qquad \begin{cases} 2a + b \geq \frac{1}{2}(x_1 - x_2) + 3 \\ a + 2b \geq \frac{1}{2}(x_2 - x_1) + 3 \\ x_1, x_2 \geq 0 \ . \end{cases}$$

Since $b = 2 - a$, (6.6) is equivalent to

$$\left. \begin{array}{l} a \geq \frac{1}{2}(x_1 - x_2) + 1 \\ \\ -a \geq \frac{1}{2}(x_2 - x_1) - 1 \end{array} \right\} \Rightarrow a = \frac{1}{2}(x_1 - x_2) + 1 \text{ and } b = \frac{1}{2}(x_2 - x_1) + 1.$$

Therefore any solution x of (6.6) fulfils the relation $\mathcal{L}(\bar{x}; \bar{\Theta}, \bar{\Lambda}) - \mathcal{L}(x; \bar{\Theta}, \bar{\Lambda}) = 0$ and this implies that (6.5) holds. With regard to the second inequality in (6.2), it is immediate to see that it becomes:

$$-\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} (a + b - 2) \not\geq_{C\setminus\{0\}} -\begin{pmatrix} 3/2 \\ 3/2 \end{pmatrix} (a + b - 2)$$

which is fulfilled by any $\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} \in \mathbb{R}^2_+$.

In the last part of this section, like in Sect. 3, we will consider the case of a scalar separation function $w(u,v;\theta,\lambda) := \langle\theta,u\rangle + \langle\lambda,v\rangle$, where $(\theta,\lambda) \in C^* \times \mathbb{R}^m_+$, $(\theta,\lambda) \neq (0,0)$. If we choose $\theta \in \operatorname{int} C^*$, then we have precisely (3.1), so that

$$w(u,v;\theta,\lambda)\rangle > 0 \ , \quad \forall (u,v) \in \mathcal{H} \ ,$$

and (3.4) is a sufficient optimality condition for the point $y \in K$. More generally, if $(\theta,\lambda) \in C^* \times \mathbb{R}^m_+$ exists such that (3.4) holds, then we will say that $\mathcal{K}$ and $\mathcal{H}$ admit a linear separation.

Consider the Lagrangian function associated with (2.1), namely $L(x;\theta,\lambda) := \langle\theta, f(x)\rangle - \langle\lambda, g(x)\rangle$. Under suitable convexity assumptions and regularity conditions, the optimality of a point y is equivalent to the existence of a saddle point for $L(x;\bar{\theta},\lambda)$ on $X \times \mathbb{R}^m_+$. We recall the following preliminary result.

Proposition 5. $\mathcal{H}$ and $\mathcal{K}(y)$ admit a linear separation iff $\exists(\bar{\theta},\bar{\lambda}) \in C^* \times \mathbb{R}^m_+$, with $(\bar{\theta},\bar{\lambda}) \neq 0$, such that the point $(y,\bar{\lambda})$ is a (scalar) saddle point of $L(x;\bar{\theta},\lambda)$ on $X \times \mathbb{R}^m_+$.

Proof. It follows from Proposition 3.1 of [25] putting $f(x,y) := f(y) - f(x)$, or readapting the proof of Theorem 8, taking into account Remark 2. $\square$

Moreover, let us recall the following regularity condition [28], which is a generalization of the Slater condition for scalar optimization problems.

Condition 1. Let $C = \mathbb{R}^\ell_+$ and $y \in K$; assume that $\forall i \in \mathcal{I}$ the following system is possible:

$$f_j(y) - f_j(x) > 0, \ j \in \mathcal{I} \setminus \{i\}; \ g(x) > 0; \ x \in X \ .$$

Proposition 6. Suppose that f and $-g$ are convex functions on the convex set X and that:

(i) Condition 1;
or
(ii) $C = \operatorname{int} \mathbb{R}^{\ell}_{+}$ and $\exists \bar{x} \in X$ such that $g(\bar{x}) > 0$;
hold. Then y is a v.m.p. of (2.1) iff there exists $(\bar{\theta}, \bar{\lambda}) \in \mathbb{R}^{\ell}_{+} \times \mathbb{R}^{m}_{+}$, with $(\bar{\theta}, \bar{\lambda}) \neq 0$, such that $(y, \bar{\lambda})$ is a saddle point for $L(x; \bar{\theta}, \lambda)$ on $X \times \mathbb{R}^{m}_{+}$.

Proof. It is proved [19] that, if f and $-g$ are convex, then the conic extension $\mathcal{E}$ (see Sect. 2) is a convex set. Recall the equivalence between (2.8) and (2.8)$'$.
Necessity. Let y be a v.m.p. of (2.1). Noting that $\mathcal{K} \subseteq \mathcal{E}$, we have that $\mathcal{H}$ and $\mathcal{K}$ are linearly separable and, by Proposition 5, a saddle point for L exists.
Sufficiency. It is well–known that, if $(y, \lambda) \in X \times \mathbb{R}^{m}_{+}$ is a saddle point for $L(x; \bar{\theta}, \lambda)$ on $X \times \mathbb{R}^{m}_{+}$, then $y \in K$. If (i) holds, then it is proved (see Theorem 2.1 of [33]) that $\bar{\theta} > 0$. Applying Proposition 2.4 of [33], and recalling the equivalence between (2.8) and (2.8)$'$, we obtain that $\mathcal{H} \cap \mathcal{E} = \emptyset$ and therefore y is a v.m.p. of (2.1). If (ii) holds then $\bar{\theta} \neq 0$ (see Theorem 2.1 of [33]) and, by Proposition 2.4 of [33], $\mathcal{H} \cap \mathcal{E} = \emptyset$, which implies that y is v.m.p. of (2.1). $\square$

7. DUALITY

In trying to satisfy the sufficient condition expressed by Theorem 1, it is natural to study, for each fixed (Θ, Λ), the following vector maximum problem:

$$(7.1) \qquad \max_{C \setminus \{0\}} w(u, v; \Theta, \Lambda)\ ,\ \text{s.t. } (u, v) \in \mathcal{K}\ ,$$

where w is given by (2.10), and, like in (2.6), $\max_{C \setminus \{0\}}$ marks vector maximum with respect to $C \setminus \{0\}$: $(\tilde{u}, \tilde{v}) \in \mathcal{K}$ is a vector maximum point of (7.1), iff

$$(7.2) \qquad w(\tilde{u}, \tilde{v}; \Theta, \Lambda) \not\leq_{C \setminus \{0\}} w(u, v; \Theta, \Lambda), \quad \forall (u, v) \in \mathcal{K}\ .$$

Lemma 2. If a maximum point in (7.1) exists, then we have:

$$(7.3) \quad \max_{(u,v) \in \mathcal{K}} {}_{C \setminus \{0\}}\, w(u, v; \Theta, \Lambda) \not\leq_{C \setminus \{0\}} 0, \quad \forall \Theta \in U^{*}_{C \setminus \{0\}},\ \forall \Lambda \in V^{*}_{C},$$

where the inequality is meant to be satisfied by every element of the "set max".

Proof. The pair $(\tilde{u} := f(y) - f(y), v := g(y))$ belongs to $\mathcal{K}$ and is such that $w(\tilde{u}, \tilde{v}; \Theta, \Lambda) = \Theta[f(y) - f(y)] + \Lambda g(y) = \Lambda g(y) \geq_C 0$. Since C is pointed we have $-\Lambda g(y) \notin C \backslash \{0\}$. Hence, if $(\tilde{u}, \tilde{v})$ is a vector maximum point of (7.1), then (7.3) holds. If not, let $(u^0, v^0) \in \mathcal{K}$ be a maximum point of (7.1), then by (7.2) we have:

$$w(u^0, v^0; \Theta, \Lambda) \not\leq_{C \backslash \{0\}} w(\tilde{u}, \tilde{v}; \Theta, \Lambda)$$

or, equivalently,

$$w(\tilde{u}, \tilde{v}; \Theta, \Lambda) - w(u^0, v^0; \Theta, \Lambda) \notin C \backslash \{0\}.$$

Since $w(\tilde{u}, \tilde{v}; \Theta, \Lambda) \in C$, we have $-w(u^0, v^0; \Theta, \Lambda) \notin C \backslash \{0\}$ and the thesis follows. □

Theorem 9. For any $y \in K$ and $\Lambda \in V_C^*$ it results

$$f(y) \not\leq_{C \backslash \{0\}} \min_{x \in X} {}_{C \backslash \{0\}} [f(x) - \Lambda g(x)] \ . \tag{7.4}$$

Proof. From (7.3), since $I_\ell \in U_{C \backslash \{0\}}^*$, we have that $\forall y \in K$ and $\forall \Lambda \in V_C^*$ it results

$$\max_{x \in X} {}_{C \backslash \{0\}} [f(y) - f(x) + \Lambda g(x)] \not\leq_{C \backslash \{0\}} 0$$

or, equivalently,

$$f(y) \not\leq_{C \backslash \{0\}} - \max_{x \in X} [-f(x) + \Lambda g(x)]$$

and therefore (7.4) holds. □

Consider the set–valued function $\Phi : U_{C \backslash \{0\}}^* \times V_C^* \rightrightarrows \mathbb{R}^\ell$, where $\Phi(\Theta, \Lambda)$ is the set of the optimal values of (7.1).

Let us recall the definition of vector Maximum of the above set–valued map [38]:

Definition 4. $(\hat{\Theta}, \hat{\Lambda}) \in U^*_{C\setminus\{0\}} \times V^*_C$ is a *vector Maximum*[14], with respect to the cone $C\setminus\{0\}$, of the set–valued map $\Phi(\Theta, \Lambda)$ iff

$$\exists \hat{z} \in \Phi(\hat{\Theta}, \hat{\Lambda}) \quad \text{s.t.} \quad \hat{z} \not\leq_{C\setminus\{0\}} z, \quad \forall z \in \Phi(\Theta, \Lambda), \quad \forall (\Theta, \Lambda) \in U^*_{C\setminus\{0\}} \times V^*_C \ . \tag{7.5}$$

The definition of vector Minimum is quite analogous.

Let us define the following vector optimization problem:

$$\operatorname*{Max}_{\Lambda \in V^*_C} {}_{C\setminus\{0\}} \ \min_{x \in X} {}_{C\setminus\{0\}} \ \mathcal{L}(x; I_\ell, \Lambda) \ , \tag{7.6}$$

where $\mathcal{L}(x; I_\ell, \Lambda)$ has been defined in (6.1).

Problem (7.6) is called the *vector dual problem* of (2.1). Observe that, when $\ell = 1$ and $C = \mathbb{R}_+$, (7.6) collapses to the well–known Lagrangian dual.

Theorem 9 states that the vector of the objectives of the primal (2.1) calculated at any feasible solution y is not less or equal, with respect to $C\setminus\{0\}$, to the vector of the objectives of the dual (7.6) calculated at any $\Lambda \in V^*_C$; hence Theorem 9 is a Weak Duality Theorem, in the vector case.

Now, the aim is to establish a Strong Duality Theorem. To this end, let us observe that, taking into account Definition 4, from (7.3) we have:

$$\operatorname*{Min}_{\substack{\Theta \in U^*_{C\setminus\{0\}} \\ \Lambda \in V^*_C}} {}_{C\setminus\{0\}} \ \max_{(u,v) \in \mathcal{K}} {}_{C\setminus\{0\}} \ w(u, v; \Theta, \Lambda) \not\leq_{C\setminus\{0\}} 0 \ . \tag{7.7}$$

Let Ω be the set of the optimal values of (7.7); Ω is called the *image gap.* The following result holds.

Lemma 3. There exist $\bar{\Theta} \in U^*_{C\setminus\{0\}}$ and $\bar{\Lambda} \in V^*_C$, such that

$$w(u, v; \bar{\Theta}, \bar{\Lambda}) \not\geq_{C\setminus\{0\}} 0 \ , \quad \forall (u, v) \in \mathcal{K} \tag{7.8}$$

[14] Capital letters in Max or in Min denote that they deal with a set–valued map. In this definition Φ plays the role of a generic set–valued function.

iff $0 \in \Omega$.

Proof. Suppose that (7.8) holds. Let us consider the pair $(\bar{u}, \bar{v}) = (f(y) - f(y), g(y)) \in \mathcal{K}$; since $g(y) \geq 0$, it results $\bar{\Lambda} g(y) \in C$, moreover by (7.8) we have $\bar{\Lambda} g(y) \notin C \backslash \{0\}$. Hence $\bar{\Lambda} g(y) = 0$ or, equivalently, $w(\bar{u}, \bar{v}; \bar{\Theta}, \bar{\Lambda}) = 0$. This equality and (7.8) imply that $(\bar{u}, \bar{v})$ is a vector maximum point of $w(u, v; \bar{\Theta}, \bar{\Lambda})$ and that the null vector of $\mathbb{R}^{\ell}$ is the corresponding optimal value; recalling the definition of $\Phi(\Theta, \Lambda)$, this result can be expressed by the condition $\bar{z} = 0 \in \Phi(\bar{\Theta}, \bar{\Lambda})$. Now, rewrite (7.3) as follows:

$$\bar{z} - \max_{(u,v) \in \mathcal{K}} {}_{C \backslash \{0\}} w(u, v; \Theta, \Lambda) \not\geq_{C \backslash \{0\}} 0 \quad \forall \Theta \in U^*_{C \backslash \{0\}}, \quad \forall \Lambda \in V^*_C,$$

and observe that it is equivalent to

$$\bar{z} - z \not\geq_{C \backslash \{0\}} 0 \ \ \forall z \in \Phi(\Theta, \Lambda), \ \ \forall (\Theta, \Lambda) \in U^*_{C \backslash \{0\}} \times V^*_C \ .$$

This proves that $\bar{z} = 0$ is a Minimum value of $\Phi(\Theta, \Lambda)$ and completes the first part of the proof. Vice versa suppose that $0 \in \Omega$. Therefore there exists $(\bar{\Theta}, \bar{\Lambda}) \in U^*_{C \backslash \{0\}} \times V^*_C$ such that $0 \in \Phi(\bar{\Theta}, \bar{\Lambda})$. By definition of vector maximum we have that $\exists (\bar{u}, \bar{v}) \in \mathcal{K}$ such that $w(\bar{u}, \bar{v}; \bar{\Theta}, \bar{\Lambda}) = 0$ and $w(u, v; \bar{\Theta}, \bar{\Lambda}) - w(\bar{u}, \bar{v}; \bar{\Theta}, \bar{\Lambda}) \not\geq_{C \backslash \{0\}} 0$ $\forall (u, v) \in \mathcal{K}$, that is $w(u, v; \bar{\Theta}, \bar{\Lambda}) \not\geq_{C \backslash \{0\}} 0 \ \ \forall (u, v) \in \mathcal{K}$. □

Let Δ_1 be the set of the optimal values of (2.1) and Δ_2 the set of the optimal values of (7.6). Define $\Delta := \Delta_1 - \Delta_2$; Δ is called the *duality gap*.

Lemma 4. There exist $y \in K$ and $\bar{\Lambda} \in V^*_C$ such that

$$[f(y) - f(x)] + \bar{\Lambda} g(x) \not\geq_{C \backslash \{0\}} 0 \ \ \forall x \in X \tag{7.9}$$

iff $0 \in \Delta$.

Proof. Suppose that (7.9) holds. This hypothesis is equivalent to the existence of $\bar{\Theta} = I_{\ell}$ and $\bar{\Lambda} \in V^*_C$ such that (7.8) holds. Hence, following the proof of Lemma 3, it results that $\bar{z} = 0 \in \Phi(I_{\ell}, \Lambda)$ and that $\bar{z} - z \not\geq_{C \backslash \{0\}} 0 \ \forall z \in \Phi(I_{\ell}, \Lambda), \ \ \forall \Lambda \in V^*_C$. Therefore we have the following equivalences:

$$\bar{z} = 0 \in \operatorname*{Min}_{\Lambda \in V^*_C} {}_{C \backslash \{0\}} \ \max_{(u,v) \in \mathcal{K}} {}_{C \backslash \{0\}} \ w(u, v; I_{\ell}, \Lambda) \ \Leftrightarrow$$

$$0 \in \underset{\Lambda \in V_C^*}{\mathrm{Min}}{}_{C\setminus\{0\}} \max_{x \in X}{}_{C\setminus\{0\}}[f(y) - f(x) + \Lambda g(x)] \;\Leftrightarrow$$

$$0 \in f(y) + \underset{\Lambda \in V_C^*}{\mathrm{Min}}{}_{C\setminus\{0\}} \max_{x \in X}{}_{C\setminus\{0\}}[-f(x) + \Lambda g(x)] \;\Leftrightarrow$$

$$0 \in f(y) - \underset{\Lambda \in V_C^*}{\mathrm{Max}}{}_{C\setminus\{0\}} \min_{x \in X}{}_{C\setminus\{0\}}[f(x) - \Lambda g(x)] \;.$$

Taking into account that, by Theorem 1, (7.9) implies the optimality of y for (2.1), the thesis follows.
Vice versa, suppose that $0 \in \Delta$. If y is an element of the "set min" of (2.1), by the previous equivalences it results that $\exists \bar{\Lambda} \in V_C^*$ such that $0 \in \max_{x \in X}{}_{C\setminus\{0\}}[f(y) - f(x) + \bar{\Lambda} g(x)]$. Hence, by definition of vector maximum, we have $f(y) - f(x) + \bar{\Lambda} g(x) \not\geq_{C\setminus\{0\}} 0, \ \forall x \in X$. $\square$

Observe that, when $\ell = 1$ and $C = \mathbb{R}_+$, the condition $0 \in \Delta$ becomes $\Delta = \{0\}$ or, equivalently,

$$\min_{x \in K} f(x) = \max_{\lambda \in \mathbb{R}^m_+} \min_{x \in X}[f(x) - \langle \lambda, g(x) \rangle]$$

which means that the duality gap is equal to 0 in the scalar case.

Now, in order to obtain a Strong Duality Theorem in the vector case, we have to find classes of vector optimization problems for which (7.9) is satisfied. This happens if the functions involved in (2.1) fulfil a regularity condition and (2.1) is "image convex" (i.e. $\mathcal{H}$ and $\mathcal{K}(y)$ are linearly separable in the image space, when y is a v.m.p. of (2.1)).

Definition 5. Let Z be a nonempty set, A be a convex cone in $\mathbb{R}^k$ with $\operatorname{int} A \neq \emptyset$ and $F : Z \to \mathbb{R}^k$. F is said to be *A–subconvexlike* iff $\exists a \in \operatorname{int} A$, such that $\forall \epsilon > 0$ we have:

$$(1-\alpha)F(Z) + \alpha F(Z) + \epsilon a \subseteq F(Z) + A, \quad \forall \alpha \in [0,1] \;.$$

If the condition "$\exists \alpha \in \operatorname{int} A$" is replaced by "$\forall a \in A$", then the above class collapses to that of *A–convexlike* functions.

In [33] (see Theorem 5.1) it is proved that, if $(-f, g)$ is $c\ell\,\mathcal{H}$–subconvexlike, then (2.1) is image convex.

The following theorem is a consequence of Lemma 4, Theorem 8 and of some results of [4] and [33].

Theorem 10 (Strong Duality Theorem). Consider problem (2.1) with $C = \mathbb{R}^\ell_+$; let y be a v.m.p. of (2.1). If $(-f, g)$ is $c\ell\,\mathcal{H}$–subconvexlike and Condition 1 holds, then $0 \in \Delta$.

Proof. Since $(-f, g)$ is $c\ell\,\mathcal{H}$–subconvexlike and y is a v.m.p. of (2.1), then $\mathcal{H}$ and $\mathcal{K}(y)$ admit linear separation. This fact and Condition 1 imply (see Theorem 2.1 and Proposition 3.1 of [33]) that $\bar\theta \in \operatorname{int} \mathbb{R}^\ell_+$ and $\bar\lambda \in \mathbb{R}^m_+$ exist such that $(y, \bar\theta, \bar\lambda)$ is a saddle point of the (scalar) Lagrangian function $L(x; \theta, \lambda)$ (defined in Sect. 6). This implies (see Theorem 3.1 of [4]) that $\bar\Lambda \in V^*_C$ exists such that $(y, \bar\Lambda)$ is a vector saddle point of the vector Lagrangian function $\mathcal{L}(x; I_\ell, \Lambda)$. Finally, Theorem 8 affirms that this is equivalent to the condition

$$w(u, v; I_\ell, \bar\Lambda) \not\geq_{C\setminus\{0\}} 0 \quad \forall (u, v) \in \mathcal{K}(y) \ ,$$

that is (7.9). By means of Lemma 4 the proof is complete. □

The next result is a straightforward consequence of Theorem 8 and of Lemma 4.

Corollary 2. $0 \in \Delta$ iff there exists $\bar\Lambda \in V^*_C$, such that $(y; I_\ell, \bar\Lambda)$ is a saddle point of $\mathcal{L}(x; I_\ell, \Lambda)$ on $X \times V^*_C$.

Example 3 (continuation). Consider again the problem of Example 3. Since $\mathcal{L}(x; I_\ell, \Lambda)$ admits a saddle point, taking into account Corollary 2, we have that $0 \in \Delta$.

Example 4. Let $f_1(x) = -\frac{1}{x+1}, f_2(x) = x, g(x) = x^3,\ X = \mathbb{R},\ C = \mathbb{R}^2_+$. $x = 0$ is the unique v.m.p. since the components of (f_1, f_2) are increasing. We will show that $\not\exists \bar\Theta \in U^*_{C\setminus\{0\}}, \bar\Lambda \in V^*_C$ such that $(0, \bar\Lambda)$ is a vector saddle point for $\mathcal{L}(x; \bar\Theta, \Lambda)$ on $X \times V^*_C$. In this case, we have:

$$\begin{aligned}\mathcal{L}(x; \bar\Theta, \Lambda) &= \begin{pmatrix} \bar\theta_{11}\bar\theta_{12} \\ \bar\theta_{21}\bar\theta_{22} \end{pmatrix} \begin{pmatrix} \frac{-1}{x+1} \\ x \end{pmatrix} - \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} x^3 = \\ &= (\, -\bar\theta_{11}\tfrac{1}{x+1} + \bar\theta_{12}x - \lambda_1 x^3, -\bar\theta_{21}\tfrac{1}{x+1} + \bar\theta_{22}x - \lambda_2 x^3 \,)\,,\end{aligned}$$

$\forall i \in \mathcal{I} = \{1,2\}$ $\exists j(i) \in \mathcal{I}$ such that $\bar{\theta}_{ij(i)} \neq 0$. If $(0,\bar{\Lambda})$ were a vector saddle point for $\mathcal{L}(x;\bar{\Theta},\Lambda)$ on $X \times V_C^*$, then there would exist $\mu_1,\mu_2 \in \mathbb{R}_+$, with $(\mu_1,\mu_2) \neq (0,0)$, s.t. $\nabla_x[\mu_1\mathcal{L}_1(0;\bar{\Theta},\bar{\Lambda}) + \mu_2\mathcal{L}_2(0,\bar{\Theta},\bar{\Lambda})] = 0$ or, equivalently, $\mu_1(\bar{\theta}_{11}+\bar{\theta}_{12})+\mu_2(\bar{\theta}_{21}+\bar{\theta}_{22}) = 0$, which implies that $\mu_1 = \mu_2 = 0$; therefore $x = 0$ is not a v.m.p. for $\mathcal{L}(x;\Theta,\Lambda)$, $\forall(\Theta,\Lambda) \in U^*_{C\setminus\{0\}} \times V_C^*$. Hence from Corollary 2 we have that $0 \notin \Delta$.

8. SCALARIZATION OF VECTOR OPTIMIZATION

Now, let us consider one of the most analysed topics in Vector Optimization: scalarization of (2.1), namely how to set up a scalar minimization problem, which leads to detecting all the solutions to (2.1) or at least one.

Assume that X be convex. Let us recall that f is called C–*function*, iff $\forall x', x'' \in X$ we have [18]:

$$(8.1)\quad (1-\alpha)f(x') + \alpha f(x'') - f((1-\alpha)x' + \alpha x'') \in C\ ,\ \forall \alpha \in [0,1].$$

When $C \supseteq \mathbb{R}^\ell_+$ or $C \subseteq \mathbb{R}^\ell_+$, then f is called C–*convex*. At $\ell = 1$ and $C = \mathbb{R}_+$ we find the classic definition of convexity. A C–function is also C–convexlike (see Definition 5); the vice versa is not true as the following example shows.

Example 5. Let $X = \mathbb{R}$, $C = \mathbb{R}^2_+$, and $f = (f_1, f_2)$ with $f_1(x) = f_2(x) = x^3$. We have $f(X) = \{(x_1,x_2) \in \mathbb{R}^2 : x_1 = x_2\}$, so that $f(X) + C$ is a convex set. Recalling that, if f is a C–function, then $\forall c^* \in C^*$ the function $\langle c^*, f(x)\rangle$ is convex (see Proposition 8), it is immediate to see that f is not a C–function.

Now, $\forall y \in X$, consider the sets[15]:

$$S(y) := \{x \in X : f(x) \in f(y) - C\} \ ,$$
$$S_p(y) := \{x \in X : \langle p, f(x)\rangle \leq \langle p, f(y)\rangle\} \ ,$$

where $p \in C^*$. When $X = \mathbb{R}^n$ and $C = \mathbb{R}^\ell_+$, then the above sets are the lower level sets of f and $\langle p, f\rangle$, respectively. If f is linear, then $S(y)$ is a cone with apex at y, and $S_p(y)$ a supporting halfspace of $S(y)$ at its apex.

Proposition 7. If f is a C–function, then $S(y)$ is convex $\forall y \in X$.

Proof. $x', x'' \in S(y) \Rightarrow \exists c', c'' \in C$ such that $f(x') = f(y) - c'$ and $f(x'') = f(y) - c''$. From these equalities, since the convexity of C implies $\tilde{c} := (1-\alpha)c' + \alpha c'' \in C$, $\forall \alpha \in [0,1]$, we find:

$$(1-\alpha)f(x') + \alpha f(x'') = f(y) - \tilde{c} \quad , \quad \forall \alpha \in [0,1]. \tag{8.2}$$

¿From (8.1) we have that $\exists \hat{c} \in C$, such that:

$$f((1-\alpha)x' + \alpha x'') = (1-\alpha)f(x') + \alpha f(x'') - \hat{c} = f(y) - \tilde{c} - \hat{c} = f(y) - c,$$

where $(1-\alpha)x' + \alpha x'' \in X$ (since X is convex), $c := \tilde{c} + \hat{c} \in C$ because C is a convex cone, and the last but one equality comes from (8.2). It follows that

$$(1-\alpha)x' + \alpha x'' \in S(y) \ , \quad \forall \alpha \in [0,1], \quad \forall x', x'' \in S(y).$$

□

Now, consider any fixed $p \in C^*$, and introduce the (scalar) Quasi–minimum Problem (in the unknown x):

$$\min \ \langle p, f(x)\rangle \quad , \quad \text{s.t.} \ \ x \in K \cap S(y), \tag{8.3}$$

[15] In what follows, p will not play the role of a parameter and will be consider fixed. Note that, now y is introduced as a parameter – notwithstanding the fact that in the preceding sections it has always denoted the unknown – since in subsequent development it will play both roles.

which depends on the parameter y.

Note that the feasible region of (8.3) is precisely $K(y)$ defined in (3.7a).

Remark 3. Under suitable assumptions the 1st order necessary condition of (8.3) is

$$\langle f'(y), x - y\rangle \geq 0 \ , \ \forall x \in K(y) \ ,$$

which is a particular case of a *Quasi–variational Inequality.*

Proposition 8. Let X be convex. If f is a C–function, g is concave and $p \in C^*$, then (8.3) is convex.

Proof. We have to show that $\langle p, f\rangle$ and $K \cap S(y)$ are convex. $p \in C^*$ and (8.1) imply, $\forall x', x'' \in X$,

$$\langle p, (1-\alpha)f(x') + \alpha f(x'') - f((1-\alpha)x' + \alpha x'')\rangle \geq 0 \ , \quad \forall \alpha \in [0,1],$$

or

$$\langle p, f((1-\alpha)x' + \alpha x'')\rangle \leq (1-\alpha)\langle p, f(x')\rangle + \alpha\langle p, f(x'')\rangle, \quad \forall \alpha \in [0,1],$$

which expresses the convexity of $\langle p, f(x)\rangle$. The convexity of X and the concavity of g give the convexity of K. Because of Proposition 7 we obtain the convexity of $S(y)$ and hence that of $K \cap S(y)$. $\square$

Proposition 9. If $p \in C^*$, then

$$S(y) \subseteq S_p(y) \quad , \quad y \in S(y) \cap S_p(y) \ , \quad \forall y \in X. \tag{8.4}$$

Proof. $x \in S(y) \Rightarrow \exists c \in C$ such that $f(x) = f(y) - c$. From this equality, taking into account that $p \in C^*$ and $c \in C$ imply $\langle p, c\rangle \geq 0$, we find:

$$\langle p, f(x)\rangle = \langle p, f(y)\rangle - \langle p, c\rangle \leq \langle p, f(y)\rangle \ , \ \forall y \in X.$$

The 1st of (8.4) follows. $0 \in C \Leftrightarrow y \in S(y)$; $y \in S_p(y)$ is trivial; hence the 2nd of (8.4) holds. $\square$

Now, let us state some properties; they might be useful in defining a method for finding one or all the solutions to (2.1) by solving (8.3).

Proposition 10. Let $p \in \operatorname{int} C^*$ be fixed. Then, (2.5) is impossible – and hence y is a solution to (1.1) – iff the system (in the unknown x):

(8.5) $\quad \langle p, f(y) - f(x) \rangle > 0 \;, \quad f(y) - f(x) \in C \;, \; g(x) \geq 0, \quad x \in X$

is impossible. Furthermore, the impossibility of (8.5) is a necessary and sufficient condition for y to be a (scalar) minimum point of (8.3).

Proof. The 1st of (8.5)$\Rightarrow f(y) - f(x) \neq 0$, so that the possibility of (8.5) implies that of (2.5). The 1st of (2.5) and $p \in \operatorname{int} C^*$ imply the 1st of (8.5), so that the possibility of (2.5) implies that of (8.5). By replacing the 1st of (8.5) equivalently with $\langle p, f(x) \rangle < \langle p, f(y) \rangle$, we immediately obtain the 2nd part of the statement. $\square$

Proposition 11. We have:

(8.6) $$x^0 \in S(y^0) \;\Rightarrow\; S(x^0) \subseteq S(y^0),$$

whatever $y^0 \in X$ may be.

Proof. $x^0 \in S(y^0) \Rightarrow \exists c^0 \in C$ such that $f(x^0) = f(y^0) - c^0$. $\hat{x} \in S(x^0) \Rightarrow \exists \hat{c} \in C$ such that $f(\hat{x}) = f(x^0) - \hat{c}$. Summing up side by side the two equalities we obtain $f(\hat{x}) = f(y^0) - c$, where $c := c^0 + \hat{c} \in C$ since C is a convex cone. It follows that $\hat{x} \in S(y^0)$ and hence $S(x^0) \subseteq S(y^0)$. $\square$

Proposition 12. If x^0 is a (global) minimum point of (8.3) at $y = y^0$, then x^0 is a (global) minimum point of (8.3) also at $y = x^0$.

Proof. Ab absurdo, suppose that x^0 be not a (global) minimum point of (8.3) at $y = x^0$. Then

(8.7) $$\exists \hat{x} \in K \cap S(x^0) \;\text{ s.t. }\; \langle p, f(\hat{x}) \rangle < \langle p, f(x^0) \rangle.$$

Because of Proposition 11, $x^0 \in S(y^0) \Rightarrow S(x^0) \subseteq S(y^0)$. This inclusion and (8.7) imply

$$\hat{x} \in K \cap S(y^0) \quad \text{and} \quad \langle p, f(\hat{x})\rangle < \langle p, f(x^0)\rangle,$$

which contradict the assumption. □

Proposition 12 suggests a method for finding a v.m.p. of (2.1). Let us choose any $p \in \operatorname{int} C^*$; p will remain fixed in the sequel. Then, we choose any $y^0 \in K$ and solve the (scalar) problem (8.3) at $y = y^0$. We find a solution x^0 (if any). According to Proposition 12, x^0 is a v.m.p. of (2.1). If we want to find all the solutions to (2.1) – this happens, for instance, when a given function must be optimized over the set of v.m.p. of (2.1) – , we must look at (8.3) as a parametric problem with respect to y; Propositions 10 and 12 guarantee that all the solutions to (2.1) will be reached. Note that such a scalarization method does not require any assumption on (2.1).

To find x^0 may not be, in the general case, an easy task. If this is due to the presence, in $K \cap S(y^0)$, of a difficult constraint, then a penalization method [19] can be obviously adopted for (8.3). Apart from computational aspects, such a penalization might be an alternative approach to the one that will be outlined in Sect. 11.

In order to stress the differences between the classic scalarization of a Vector Optimization Problem and the present one, let us consider the following examples.

Example 6. Let us set $\ell = 2$, $m = 2$, $n = 1$, $X = \mathbb{R}$, $C = \mathbb{R}^2_+$, and $f_1(x) = x$, $f_2(x) = x^2, g_1(x) = x + 1 \geq 0, g_2(x) = -x$. Obviously, $K = [-1, 0]$, and all the elements of K are v.m.p. of (2.1). Set $y = 0$. Consider the classic scalarized problem

$$\min_{x \in K} \, (p_1 x + p_2 x^2) \quad \text{with} \quad p_1, p_2 > 0 \; .$$

Note that $x = 0$ is not a (global) minimum point of the classic scalarized problem whatever $p_1, p_2 > 0$ may be.

Example 7. Let us set $\ell = 2$, $m = 2$, $n = 1$, $X = \mathbb{R}$, $C = \mathbb{R}^2_+$, and

$$f_1(x) = 2x - x^2, \; f_2(x) = 1 - x^2, \; g_1(x) = x,$$

$$g_2(x) = 1 - x, \; f = (f_1, f_2), \; g = (g_1, g_2).$$

We find $S(y) = \{y\} \; \forall y \in [0,1]$. Hence, the unique solution of (8.3) is y itself. By varying y, (8.3) gives, with its solutions, the interval $\sigma := [0,1]$, which is the set of v.m.p. of (2.1), as is obvious to check. Now, let us use the classic scalarization [12,15,41] outside the classic assumption of convexity, i.e. the scalar parametric problem which, here, becomes:

$$\min_{x \in \sigma} \left[c_1 f_1(x) + c_2 f_2(x) = -(c_1 + c_2)x^2 + 2c_1 x + c_2 \right] , \tag{8.8}$$

where $(c_1, c_2) \in \operatorname{int} C^* = \operatorname{int} \mathbb{R}^2_+$ are parameters. Every minimum point of (8.8) is a v.m.p. of (2.1). In the present example it is easy to see that the only solutions of (8.8) are $x = 0$, or $x = 0$ and $x = 1$, or $x = 1$, according to respectively $c_2 < c_1$, or $c_2 = c_1$, or $c_2 > c_1$. Hence, the scalarized problem (8.8) does not detect all the solutions to (2.1) (the same happens obviously to (8.3), if $S(y)$ is deleted).

Example 8. Let us set $\ell = 2, m = 1, n = 2, X = \mathbb{R}^2, C = \mathbb{R}^2_+$, $x = (x_1, x_2), y = (y_1, y_2)$, and

$$f_1(x) = x_1 + 2x_2 \; , \; f_2(x) = 4x_1 + 2x_2 \; , \; g(x) = -|x_1| + x_2.$$

Choose $p = (1,1)$ and $y^0 = (0,1)$. Then (8.3) becomes:

$$\min(5x_1 + 4x_2) \; , \; \text{s.t.} \; -|x_1| + x_2 \geq 0, x_1 + 2x_2 \leq 2, 2x_1 + x_2 \leq 1. \tag{8.9}$$

The (unique) solution to (8.9) is easily found to be $x^0 = (-2, 2)$. Because of Proposition 12, x^0 is a v.m.p. of (2.1) in the present case. Furthermore, we have $K \cap S(x^0) = \{x^0\}$, namely the parametric system (in the unknown x):

$$-|x_1| + x_2 \geq 0 \; , \; x_1 + 2x_2 \leq y_1 + 2y_2 \; , \; 2x_1 + x_2 \leq 2y_1 + y_2 \tag{8.10}$$

has the (unique) solution x^0. In order to find all the v.m.p. of (2.1) we have to search for all $y \in K$ such that (8.10) has y itself as the (unique) solution. (8.10) is equivalent to

$$\begin{cases} |x_1| \leq x_2 \leq -\dfrac{1}{2}x_1 + \dfrac{1}{2}(y_1 + 2y_2) \\ \qquad x_2 \leq -2x_1 + 2y_1 + y_2. \end{cases} \tag{8.11}$$

With $x_1 > 0$, (8.11) cannot have y as (unique) solution. Hence, we consider the case $x_1 \le 0$; by using the Motzkin elimination method and by requiring a unique solution, (8.11) becomes:

$$-\frac{1}{2}x_1 = \frac{1}{2}(y_1 + 2y_2) \quad , \quad x_1 = 2y_1 + y_2 \quad , \quad x_1 \le 0$$

and leads us to $y_1 + y_2 = 0$, $y_1 \le 0$ or

$$y = (y_1 = -t,\ y_2 = t) \quad , \quad t \in [0, +\infty[\ ,$$

which gives us all the v.m.p. of (2.1). Now, let us use the classic scalarization [12,15,41], i.e. the scalar parametric problem which, here, becomes:

$$\text{(8.12)} \qquad \begin{array}{l} \min[c_1 f_1(x) + c_2 f_2(x) = (c_1 + 4c_2)x_1 + (2c_1 + 2c_2)x_2], \\ \text{subject to} \quad -|x_1| + x_2 \ge 0, \end{array}$$

where $(c_1, c_2) \in C^* \backslash \{0\} = \mathbb{R}^2_+ \backslash \{0\}$ are parameters. Such a scalarization detects all the v.m.p. of (2.1) by solving (8.12) with respect to all possible pairs of parameters (c_1, c_2), even if, in general, it finds more points than the ones desired. In the present example, it is obvious to see that the minimum of (8.12) exists iff $-\frac{1}{2}c_1 \le c_2 \le \frac{1}{2}c_1$, and that the minimum points of (8.12) are all the v.m.p. of (2.1) at $0 \le c_2 \le \frac{1}{2}c_1$.

In classic scalarization – which works under convexity assumptions – the number of parameters is ℓ (as many as the objective functions), while in the present one it is n (as many as the elements of the unknown). Therefore, when we search for all the v.m.p., if $\ell < n$, the former is, in the convex case, advantageous with respect to the latter. However, the latter can be turned into a more suitable form. Instead of (8.3), consider the parametric problem, say $P_p(\xi)$, defined by

$$\text{(8.3a)}' \qquad \min \ \langle p, f(x) \rangle \quad , \quad \text{s.t.} \quad x \in K,\ f(x) \in \{\xi\} - C \ ,$$

where $\xi \in f(X)$ is a parameter. Moreover, consider the equality:

$$\text{(8.3b)}' \qquad f(x) = \xi \ .$$

In order to find all the v.m.p. of (2.1), a first v.m.p., say x^0, can be found by using (8.3) as previously described. To find all the others, we should parametrically move $y \in K$, starting with $y = x^0$ and mantaining y itself as a solution to (8.3). Alternatively, set $\xi^0 := f(x^0)$ and parametrically move $\xi \in f(X)$ in (8.3a)$'$, starting with $\xi = \xi^0$, mantaining a solution, say $x^0(\xi)$ to (8.3a)$'$, and guaranteeing that $x^0(\xi)$ fulfils (8.3b)$'$ or $f(x^0(\xi)) = \xi$. In this way, we have to handle a parameter with ℓ components (instead of n) and the additional condition (8.3b)$'$ (which makes up for the absence of assumptions).

If X, f and $-g$ are convex, then the last constraint of (8.3a)$'$ and condition (8.3b)$'$ are superfluous, as Proposition 13 will show. In such a case, in particular in the linear or quadratic ones, it would be interesting to compare a method based on (8.3a)$'$ for finding all the v.m.p. with those existing in the literature; see for instance [12,15,41].

Proposition 13. (i) Let $p \in \operatorname{int} C^*$. If y is a (scalar) minimum point of problem:

$$\min \ \langle p, f(x) \rangle \ , \quad x \in K, \tag{8.13}$$

then y is a v.m.p. of (2.1) and hence also of (2.3). (ii) Let f be C–convexlike. If $y \in K$ is a v.m.p. of (2.1), then $\exists p_y \in C^* \backslash \{0\}$, such that y is a solution of (8.13) at $p = p_y$.

Proof. (i) Ab absurdo, suppose that $\exists \hat{x} \in K$ such that (2.4) be not fulfilled or

$$f(y) - f(\hat{x}) \in C \backslash \{0\} \ .$$

This relation and $p \in \operatorname{int} C^*$ imply $\langle p, f(y) - f(\hat{x}) \rangle > 0$, or $\langle p, f(\hat{x}) \rangle < \langle p, f(y) \rangle$ which contradicts the assumption. The last part of the claim is obvious. (ii) The optimality of y means that $(f(y) - f(K)) \cap (C \backslash \{0\}) = \emptyset$. Since

$$\begin{aligned}[(f(y) - f(K)] \cap (C \backslash \{0\}) = (f(y) - f(K)) \cap [C + (C \backslash \{0\})] = \\ = (f(y) - [f(K) + C]) \cap (C \backslash \{0\}),\end{aligned}$$

the previous condition is equivalent to the following:

$$[(f(y) - (f(K) + C)] \cap (C \backslash \{0\}) = \emptyset \ .$$

The assumption on f is equivalent to the convexity of $f(K)+C$ [40], so that $f(y)-[f(K)+C]$ is convex too. Hence $\exists p_y \in C^*\backslash\{0\}$ such that

$$\langle p_y, f(y)-f(x)+c\rangle \leq 0, \quad \forall x \in K, \quad \forall c \in C \ .$$

At $c=0$, we obtain:

$$\langle p_y, f(y)-f(x)\rangle \leq 0, \quad \forall x \in K,$$

which completes the proof. □

Example 9. Let us set $\ell=2, n=2, X=\mathbb{R}^2$, $K=\{x \in \mathbb{R}^2_+ : x_1+x_2 \geq 2\}$, $f=(f_1,f_2)$ with $f_1(x)=x_1, f_2(x)=x_2, C=\mathbb{R}^2_+$. We have $\operatorname{int} C^* = \operatorname{int} \mathbb{R}^2_+$. Choose $p=(1,0)$. A minimum point of (8.13) is $y=(0,3)$ which, however, is not v.m.p. of (2.1). We come to the conclusion that, if $p \in C^*\backslash\{0\}$ (instead of $\operatorname{int} C^*$), then a minimum point of (8.13) is not necessarily a v.m.p. of (2.1).

Now, let us discuss briefly another way of finding all the solutions to (2.1) in the special case where $C=\mathbb{R}^\ell_+, X=\mathbb{R}^n$, f and $-g$ are convex and differentiable, and (8.3) satisfies a regularity condition [29]. In such a case, it is well–known that the so–called Karush–Kuhn–Tucker condition is necessary and sufficient for a stationary point to be also a minimum point. Hence, in the present case, y is a minimum point of (8.3), iff $\exists \lambda \in \mathbb{R}^m$ and $\exists \mu \in \mathbb{R}^\ell$, such that

$$\text{(8.14)} \qquad \begin{cases} (p+\mu)^T f'(y) - \lambda^T g'(y) = 0 \ , \ \langle \lambda, g(y)\rangle = 0 \ , \\ g(y) \geq 0 \ , \ \lambda \geq 0 \ , \ \mu \geq 0 \ . \end{cases}$$

The set of solutions to this system equals the set of v.m.p. of (2.1) under the above assumptions. This result can be extended to a wider class of VOP; for instance, to invex ones.

Without the 1st equation and last inequality, (8.14) would be a classic nonlinear complementarity system. It would be interesting to investigate the properties of (8.14) (and of its special cases, in particular that of f linear and g affine) by exploiting the existing theory of complementarity systems. Indeed, when we have to optimize a given function over the set of v.m.p. of (2.1), then (8.14) becomes the “feasible region” for such a problem.

Example 8 (continuation). Same data as in Example 8, with the sole exception of g, which is now splitted into $g_1(x) = x_1 + x_2$ and $g_2(x) = -x_1 + x_2$. With $p = (1, 1)$, (8.14) becomes:

$$\begin{aligned}&\lambda_1 - \lambda_2 - \mu_1 - 4\mu_2 = 5 \ , \quad y_1 + y_2 \geq 0 \ ,\\ &\lambda_1 + \lambda_2 - 2\mu_1 - 2\mu_2 = 4 \ , -y_1 + y_2 \geq 0 \ ,\\ &\lambda_1(y_1 + y_2) + \lambda_2(-y_1 + y_2) = 0, \ \lambda_1, \lambda_2, \mu_1, \mu_2 \geq 0 \ .\end{aligned}$$

We deduce

$$\mu_1 = -1 + \frac{1}{3}\lambda_1 + \lambda_2 \geq 0 \ , \ \mu_2 = -1 + \frac{1}{6}\lambda_1 - \frac{1}{2}\lambda_2 \geq 0$$

which require $\lambda_1 > 0$, (indeed, ≥ 6) and hence $y_1+y_2 = 0$. Therefore, we find that the only solutions to the above system are the set of (y_1, y_2) such that $y_1 + y_2 = \ 0, y_1 \leq 0$, as before.

Now, let us consider another particular case, where C is polyhedral and $X = \mathbb{R}^n$, so that it is not restrictive to set $C = \{z \in \mathbb{R}^\ell : Az \geq 0\}$, A being a $k \times \ell$ matrix with real entries. (8.3) becomes:

$$(8.13)'' \quad \min \langle p, f(x)\rangle \ , \ \text{s.t.} \ A[f(y) - f(x)] \geq 0 \ , \ g(x) \geq 0 \ , \ x \in X \ .$$

Let f and $-g$ be convex and differentiable, and $(8.13)''$ be regular [29]. A necessary and sufficient condition for y to be minimum point of $(8.13)''$ is that $\exists \lambda \in \mathbb{R}^m$ and $\exists \mu \in \mathbb{R}^k$, such that:

$$(8.15) \qquad \begin{cases} (p^T + \mu^T A)f'(y) - \lambda^T g'(y) = 0 \ , \ \langle \lambda, g(y)\rangle = 0 \ ,\\ g(y) \geq 0 \ , \ \lambda \geq 0 \ , \ \mu \geq 0 \ .\end{cases}$$

At $k = \ell$ and $A = I$, (8.15) is reduced to (8.14). The set of solutions of (8.15) equals the set of v.m.p. of (2.1).

Assume that the equations of (8.15) define y as implicit function of λ and μ, say $y = y(\lambda, \mu)$; this happens, for instance, if they fulfil the assumption of Dini Implicit Function Theorem. Then, (8.15) becomes:

$$(8.16) \qquad g(y(\lambda, \mu)) \geq 0 \ , \ \lambda \geq 0 \ , \ \mu \geq 0 \ , \ \langle \lambda, g(y(\lambda, \mu))\rangle = 0 \ ,$$

which is, for any fixed μ, a standard *nonlinear complementarity system* (see Example 10).

When a VOP is the mathematical model of an optimal control problem, then (8.16) can be considered as "restricted" feasible region – in literature, named set of efficient or nondominated points – and a scalar objective function can be defined on it and minimized or maximized; (see [3] for the linear case). If the format (8.16) is kept, such an objective function is expressed in terms of dual variables.

It would be interesting to find classes of problems for which $y(\lambda, \mu)$ can be determined explicitly. This happens when f is a quadratic form, g affine and $X = \mathbb{R}^n$. In general, it is important to detect properties of (8.16).

Example 10. Let us set $\ell = 2, m = 1, n = 2, X = \mathbb{R}^2_+, C = \mathbb{R}^2_+, x = (x_1, x_2), y = (y_1, y_2)$, and

$$f(x) = \begin{pmatrix} (x_1 - 2)^2 + (x_2 + 1)^2 \\ (x_1 + 1)^2 + (x_2 - 2)^2 \end{pmatrix} \quad , \quad g(x) = x_1 + x_2 - 2 \ .$$

Choose $p = (1, 1)$. The first $n = 2$ equations in (8.15) become:

$$2(1 + \mu_1)(y_1 - 2) + 2(1 + \mu_2)(y_1 + 1) - \lambda = 0 \ ,$$

$$2(1 + \mu_1)(y_2 + 1) + 2(1 + \mu_2)(y_2 - 1) - \lambda = 0 \ ,$$

and allow us to explicitly obtain

$$\begin{aligned} y_1(\lambda, \mu) &= \frac{\lambda + 4\mu_1 - 2\mu_2 + 2}{2(\mu_1 + \mu_2) + 4} \ , \\ y_2(\lambda, \mu) &= \frac{\lambda - 2\mu_1 + 4\mu_2 + 2}{2(\mu_1 + \mu_2) + 4} \ . \end{aligned} \tag{8.17}$$

By means of (8.17), system (8.16) is obtained explicitly. However, due to the particular case, we can simplify it. The 1st of (8.16) becomes

$$\frac{1}{\mu_1 + \mu_2 + 2} \lambda \geq 1 \ ,$$

so that $\lambda > 0$ necessarily and, consequently, (8.16) becomes:

$$y_1 + y_2 = 2 \ , \ y_1, y_2 \geq 0 \ .$$

Remark 4. The condition for complete efficiency established at the end of Sect. 2 can be viewed also as an obvious consequence of

Proposition 10. In fact, in such a circumstance, from Proposition 10 we get that, $\forall y \in K$

$$\langle p, f(y) - f(x) \rangle \leq 0, \quad \forall x \in K \text{ and s.t. } f(y) - f(x) \in C$$

and then we are led to condition (2.15).

9. IMAGE SPACE AND SEPARATION FOR VVI

The approach described in Sects.2 and 3 with reference to (2.1), and in Sect. 4 with reference to (2.3) can be adopted also in fields other than Optimization. Indeed, the starting point is the impossibility of a system; (2.5) is a special case. Now, let $F : \mathbb{R}^n \to \mathbb{R}^{\ell \times n}$ be a matrix–valued function, and consider the following VVI: find $y \in K$, such that

$$F(y)(x - y) \not\leq_{C \setminus \{0\}} 0 \ , \ \forall x \in K \ , \tag{9.1}$$

where C and K are as in Sect.2. At $\ell = 1$ and $C = \mathbb{R}_+$, (9.1) becomes the classic Stampacchia Variational Inequality [25]. At $\ell \geq 1$ and $C = \mathbb{R}^\ell_+$, the study of (9.1) was proposed in [18]. Obviously, y is a solution of (9.1) iff the system (in the unknown x):

$$F(y)(y - x) \geq_{C \setminus \{0\}} 0 \ , \ g(x) \geq 0 \ , \ \forall x \in X \ , \tag{9.2}$$

is impossible. Consider the set[16]

$$\mathcal{K}(y) := \{(u, v) \in \mathbb{R}^\ell \times \mathbb{R}^m : u = F(y)(y - x), \ g(x) \geq 0, \ x \in X\} \ ,$$

which replaces $\mathcal{K}(y)$ of Sect.2; while $\mathcal{H}$ is the same as in Sect.2. $\mathcal{K}(y)$ is the *image* of (9.1). Unlike what happens for VOP, a change of y does not necessarily imply a translation. To (9.1) we can associate the following *image problem*:

$$\max_{C \setminus \{0\}} u \ , \quad \text{s.t.} \ (u, v) \in \mathcal{K}(y) \cap (\mathbb{R}^\ell \times \mathbb{R}^m_+) \ ; \tag{9.3}$$

---

[16] Without any fear of confusion, for the sake of simplicity, we use here the same symbols as in Sect. 2.

like in Sects. 2 and 7, $\max_{C\setminus\{0\}}$ marks vector maximum with respect to the cone $C\setminus\{0\}$: $(\tilde{u},\tilde{v}) \in \mathcal{K}(y) \cap (\mathbb{R}^\ell \times \mathbb{R}^m_+)$ is a vector maximum point of (9.3) iff

$$u \not\geq_{C\setminus\{0\}} \tilde{u} \quad , \quad \forall (u,v) \in \mathcal{K}(y) \cap (\mathbb{R}^\ell \times \mathbb{R}^m_+) \ . \tag{9.4}$$

In the image space we define an extremum problem also when in the initial space we are not given an extremum problem; this fact is not surprising. Indeed, in all cases the starting point for introducing the image space is a system: (2.5), (3.5), (4.1) and (9.2) in the above cases.

Now, observe that system (9.2) is impossible iff

$$\mathcal{H} \cap \mathcal{K}(y) = \emptyset \ . \tag{9.5}$$

Hence, y is a solution of (9.1) iff (9.5) holds. Since (9.2) has the same kind of inequalities as (2.5), and $\mathcal{H}$ is the same as in Sect. 2, then the class of separation functions (2.10) works here too. Hence, with obvious changes, Theorem 1 becomes here:

Corollary 3. Let $y \in K$. If there exist matrices $\Theta \in U^*_{C\setminus\{0\}}$ and $\Lambda \in V^*_C$, such that

$$\Theta F(y)(y-x) + \Lambda g(x) \not\geq_{C\setminus\{0\}} 0 \quad , \quad \forall x \in X \ , \tag{9.6}$$

then y is a solution of (9.1).

The comments made at the end of Sect. 2 extend to the present case. Furthemore, note that here too the class of scalar separation functions (3.1) can be adopted for finding a theorem quite analogous to Theorem 2. Also the approach (3.5) can be followed; (3.5) becomes:

$$F(y)(y-x) \in C \ , F(y)(y-x) \neq 0 \ , \ g(x) \geq 0 \ , \ \forall x \in X \ . \tag{9.7}$$

When $C = \mathbb{R}^\ell_+$, $(3.6)'$ is replaced by

$$\begin{aligned} &F_r(y)(y-x) > 0 \ , F(y)(y-x) \geq 0 \ , \ g(x) \geq 0 \ , \ \forall x \in X \ , \\ &r \in \mathcal{I} \ . \end{aligned} \tag{9.7$'$}$$

From $(9.7)'$ we derive the analogous proposition of Corollary 1:

Corollary 4. Let $y \in K$. Assume that, $\forall r \in \mathcal{I}$, there exist $\theta_s^r \geq 0$ $s \in \mathcal{I}\backslash\{r\}$, and $\lambda^r \in \mathbb{R}_+^m$, such that:

$$
(9.8) \qquad \begin{aligned} & F_r(y)(y-x) + \sum_{s \in I\backslash\{r\}} \theta_s^r F_s(y)(y-x) + \langle \lambda^r, g(x) \rangle \leq 0 \ , \\ & \forall x \in X. \end{aligned}
$$

Then, y is a solution to (9.1).

In the same light as (2.3) is associated to (2.1), to the VVI (9.1) we can associate the following weak VVI (for short, WVVI), which consists in finding $y \in K$, such that

$$
(9.9) \qquad F(y)(x-y) \not\leq_{\operatorname{int} C} 0 \ \ , \ \ \forall x \in K \ .
$$

In this case, the sets U and $\mathcal{H}$ are as in Sect.4. This being done, then the class of separation functions (4.2) works here too, and Theorem 4 becomes here:

Corollary 5. Let $y \in K$. If there exist matrices $\Theta \in U_{\operatorname{int} C}^*$ and $\Lambda \in V_c^*$, such that

$$
(9.10) \qquad \Theta F(y)(y-x) + \Lambda g(x) \not\geq_{\operatorname{int} C} 0 \ , \ \forall x \in X \ ,
$$

then y is a solution to (9.9).

Even if (9.1) and (9.9) are distinct mathematical models, which represent distinct real equilibrium problems, the solutions to (9.9) are called *weak solutions* to (9.1), since they are strictly related to those to (9.1).

The VVI (9.1) can be called Stampacchia VVI, since it is the natural extension, to the vector case, of the simplest form of Stampacchia (scalar) Variational Inequality in Euclidean spaces. Analogously to what happens in the scalar case, to (9.1) we can associate another inequality [21,25,26], which we call Minty VVI; it consists in finding $y \in K$, such that

$$
(9.11) \qquad F(x)(y-x) \not\geq_{C\backslash\{0\}} 0 \ , \quad \forall x \in K \ .
$$

Of course, to (9.11) we can associate the Minty WVVI, which consists in finding $y \in K$, such that

$$F(x)(y-x) \not\geq_{\operatorname{int} C} 0 \;, \quad \forall x \in K \;. \tag{9.12}$$

In the scalar case the Minty Variational Inequality has been shown to be important for studying the stability of a dynamical system. It would be interesting to extend such a study to vector dynamical systems.

Remark 5. Taking the line adopted for developing the study of the Stampacchia VVI, it is possible to obtain, in the IS, analogous results for the Minty VVI.

10. SCALARIZATION OF VVI

The development of Sect. 8 can be extended to VVI. Let us start with (9.1) and set $X = \mathbb{R}^n$. Consider the sets:

$$\Sigma(y) := \{x \in \mathbb{R}^n : F(y)x \in F(y)y - C\} \;,$$

$$\Sigma_p(y) := \{x \in \mathbb{R}^n : \langle pF(y), x\rangle \leq \langle pF(y), y\rangle\},$$

where $p \in C^*$ is considered a row–vector. When $C = \mathbb{R}^\ell_+$, then the above sets are the lower level sets of the vector function $F(y)x$ and of the scalar function $\langle pF(y), x\rangle$, respectively. $\Sigma_p(y)$ is a supporting halfspace of $\Sigma(y)$ at y, as Proposition 16 will show. If $F(y)$ is a constant matrix and C is polyhedral, then $\Sigma(y)$ is a polyhedron.

Proposition 14. $\Sigma(y)$ is convex $\forall y \in \mathbb{R}^n$.

Proof. $x', x'' \in \Sigma(y) \Rightarrow \exists c', c'' \in C$ such that $F(y)x' = F(y)y - c'$, $F(y)x'' = F(y)y - c''$. From these equalities, $\forall \alpha \in [0,1]$ we have:

$$F(y)[(1-\alpha)x' + \alpha x''] = F(y)y - c(\alpha),$$

where $c(\alpha) := (1-\alpha)c' + \alpha c'' \in C$ since C is convex. Hence $(1-\alpha)x' + \alpha x'' \in \Sigma(y)$ $\forall \alpha \in [0,1]$. □

Now, let us introduce the (scalar) Quasi–Variational Inequality which consists in finding[17] $y \in K(y) := K \cap \Sigma(y)$, such that:

$$\langle F_p(y), x - y \rangle \geq 0 \ , \ \forall x \in K(y), \tag{10.1}$$

where $F_p(y) := pF(y)$ and p has to be considered fixed; (10.1) is a *scalarization* of (10.1).

F will be called *C–operator* iff

$$[F(x') - F(x'')](x' - x'') \in C \ , \ \forall x', x'' \in \mathbb{R}^n \, . \tag{10.2}$$

When $C \supseteq \mathbb{R}^\ell_+$ or $C \subseteq \mathbb{R}^\ell_+$, then F will be called *C–monotone*; when $\ell = 1$, the notion of C–operator is reduced to classic ones: F becomes monotone or antitone, according to $C = \mathbb{R}_+$ or $C = \mathbb{R}_-$, respectively.

Proposition 15. If X is convex, F is a C–operator, g is concave, and $p \in C^*$, then (10.1) is monotone.

Proof. We have to show that $K(y)$ is convex and F_p monotone. The assumptions on X and g imply the convexity of K. Because of Proposition 14, $\Sigma(y)$ is convex $\forall y \in \mathbb{R}^n$. Hence $K(y)$ is convex $\forall y \in \mathbb{R}^n$. $p \in C^*$ and (10.2) imply

$$\langle p, [F(x') - F(x'')](x' - x'') \rangle \geq 0 \quad \forall x', x'' \in \mathbb{R}^n,$$

or

$$\langle F_p(x') - F_p(x''), x' - x'' \rangle \geq 0 \quad \forall x', x'' \in \mathbb{R}^n \, .$$

□

Proposition 16. If $p \in C^*$, then

$$\Sigma(y) \subseteq \Sigma_p(y) \quad , \quad y \in \Sigma(y) \cap \Sigma_p(y), \quad \forall y \in \mathbb{R}^n \ . \tag{10.3}$$

---

[17] Without any fear of confusion, we use the same symbol $K(y)$ as in Sect. 8.

Proof. $x \in \Sigma(y) \Rightarrow \exists c \in C$ such that $F(y)x = F(y)y - c$. From this equality, having taken into account that $p \in C^*$ and $c \in C$ imply $\langle p, c \rangle \geq 0$, we have:

$$\langle pF(y), x \rangle = \langle pF(y), y \rangle - \langle p, c \rangle \leq \langle pF(y), y \rangle.$$

The 1st of (10.3) follows. The 2nd of (10.3) is a consequence of the obvious relations $y \in \Sigma(y)$ (due to the closure of C) and $y \in \Sigma_p(y)$. □

Now, let us state some preliminary properties, which might help in finding methods for solving (9.1) through (10.1).

Proposition 17. Let $p \in \operatorname{int} C^*$ be fixed. Then (9.7) is impossible – and hence y is a solution of (9.1) – iff the system (in the unknown x):

$$\text{(10.4)} \quad \langle pF(y), y - x \rangle > 0 \ , \ F(y)(y - x) \in C \ , \ g(x) \geq 0 \ , \ x \in X$$

is impossible. Furthermore, the impossibility of (10.4) is a necessary and sufficient condition for y to be a solution of (10.1).

Proof. The 1st of (10.4)$\Rightarrow F(y)(y - x) \neq 0$, so that the possibility of (10.4) implies that of (9.7). The 1st of (9.7) and $p \in \operatorname{int} C^*$ imply the 1st of (10.4). The last part of the statement is obvious. □

Note that Proposition 17 shows that the VVI (9.1) can be equivalently replaced by a scalar Quasi–Variational one, namely (10.1).

Consider, now, the function $\varphi : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, given by

$$\varphi(x; y) := \langle F_p(y), x - y \rangle,$$

and the problem (in the unknown x; $y \in K(y)$ is a parameter):

$$\text{(10.5)} \qquad \min \varphi(x; y) \ , \ \text{ s.t. } \ x \in K(y).$$

Since $\varphi(y; y) = 0$ for any y, the minimum in (10.5) is ≤ 0.

Proposition 18. $y \in K$ is a solution of (9.1) iff it is a global minimum point of (10.5).

Proof. Let y be a global minimum point of (10.5), so that $\varphi(y;y) = 0$ and $\varphi(x;y) \geq 0\ \forall x \in K(y)$. Hence, because of Proposition 17, y is a solution of (9.1). Now, let y be a solution of (9.1). Because of Proposition 17, system (10.4) is impossible, so that y is a solution of (10.1); this implies $\varphi(x;y) \geq 0\ \forall x \in K(y)$. Since $\varphi(y;y) = 0$, y is a global minimum point of (10.5). □

Example 11. Let us set $\ell = 2,\ m = 1,\ n = 2,\ C = \mathbb{R}^2_+ = C^*,\ X = \mathbb{R}^2_+,\ y = (y_1, y_2),\ x = (x_1, x_2),\ g(x) = x_1 + x_2 - 2$, and

$$F(y) = \begin{pmatrix} 2y_1 - 2 & 2y_2 + 2 \\ 2y_1 + 2 & 2y_2 - 2 \end{pmatrix}.$$

We choose $p = (p_1, p_2) = (1, 1)$, so that

$$\varphi(x;y) = 4(y_1x_1 + y_2x_2 - y_1^2 - y_2^2),$$

$$K(y) : \begin{cases} (1 - y_1)x_1 - (1 + y_2)x_2 + y_1^2 + y_2^2 - y_1 + y_2 \geq 0 \\ -(1 + y_1)x_1 + (1 - y_2)x_2 + y_1^2 + y_2^2 + y_1 - y_2 \geq 0 \\ x_1 + x_2 \geq 2\ ,\ x_1 \geq 0\ ,\ x_2 \geq 0. \end{cases}$$

We observe that every element y^0 of the segment $](2,0),(0,2)[\subset \mathbb{R}^2$ is a solution to (10.5) at $y = y^0$ and hence to (9.1). At $y = \hat{y} = (2, 0)$, problem (10.5) becomes:

$$\min\ 8(x_1 - 2)\quad ,\quad \text{s.t.}\ \ x_1 + x_2 = 2,\ 3x_1 - x_2 \leq 6,\ x_1 \geq 0,\ x_2 \geq 0,$$

whose unique global minimum point is $\hat{x} = (0, 2) \neq \hat{y}$. Hence, because of Proposition 18, $\hat{y}$ is not a solution to (9.1). A quite analogous conclusion can be drawn at $y = (0, 2)$.

Note that in Example 10 the operator $F(y)$ is the Jacobian matrix of the vector function $f : \mathbb{R}^2 \to \mathbb{R}^2$, given by

$$f(y) = \begin{pmatrix} f_1(y) \\ f_2(y) \end{pmatrix} = \begin{pmatrix} (y_1 - 1)^2 + (y_2 + 1)^2 \\ (y_1 + 1)^2 + (y_2 - 1)^2 \end{pmatrix}.$$

Consider (2.1) with the above function and with K as in Example 10, i.e. $K = \{x \in \mathbb{R}^2_+ : x_1 + x_2 \geq 2\}$. It is easy to check that all the v.m.p. of (2.1) are given by the segment $[(2,0),(0,2)]$, whose extrema are not solutions of the related VVI (9.1). Hence, such an inequality is a sufficient (but not necessary) condition for y to be v.m.p. of (2.1).

Now, let us consider Minty VVI (9.11). Instead of $\Sigma(y)$ and $\Sigma_p(y)$, we introduce the sets:

$$\Sigma^{\square}(y) := \{x \in \mathbb{R}^n : F(x)(y-x) \in C\},$$
$$\Sigma_p^{\square}(y) := \{x \in \mathbb{R}^n : \langle pF(x), y-x \rangle \geq 0\},$$

where again $p \in C^*$ is considered a row–vector. When $C = \mathbb{R}_+^{\ell}$, then the above sets are the upper level sets of the vector function $F(x)(y-x)$ and of the (scalar) function $\langle pF(x), y-x \rangle$, respectively. If $F(x)$ is a constant matrix and C is polyhedral, then $\Sigma^{\square}(y)$ is a polyhedron and $\Sigma_p^{\square}(y)$ a supporting halfspace of it.

Proposition 19. If X is convex and F a C–operator, then $\Sigma^{\square}(y)$ is convex $\forall y \in X$.

Proof. Let $x', x'' \in \Sigma^{\square}(y)$, so that

$$F(x')(y-x') \in C \quad , \quad F(x'')(y-x'') \in C \ . \tag{10.6}$$

From the assumption we have

$$F(x')(y-x')+F(y)(x'-y) \in -C \ , \ F(x'')(y-x'')+F(y)(x''-y) \in -C \ .$$

These relations together with, respectively, (10.6) imply

$$F(y)(x'-y) \in -C \quad , \quad F(y)(x''-y) \in -C \ ,$$

and hence (because of the convexity of C):

$$F(y)[x(\alpha)-y] \in -C \quad , \ \forall \alpha \in [0,1] \ , \tag{10.7}$$

where $x(\alpha) := (1-\alpha)x' + \alpha x''$. Exploiting again the assumption, (10.7) implies

$$F(x(\alpha))[y-x(\alpha)] \in C \quad , \quad \forall \alpha \in [0,1] \ ,$$

or $x(\alpha) \in \Sigma^{\square}(y),\ \forall \alpha \in [0,1]$, which completes the proof. $\square$

Now, let us introduce the (scalar) Quasi–Variational Inequality of Minty type, which consists in finding $y \in K^{\square}(y) := K \cap \Sigma^{\square}(y)$, such that:

$$\langle F_p^{\square}(x), y - x \rangle \leq 0 \quad , \quad \forall x \in K^{\square}(y) \ , \tag{10.8}$$

where $F_p^{\square}(x) := pF(x)$ and p has to be considered fixed; (10.8) is a *scalarization* of (9.11).

Proposition 20. If X is convex, F is a C–operator, g is concave, and $p \in C^*$, then (10.8) is monotone and equivalent to (10.1).

Proof. First of all, we have to show that $K^{\square}(y)$ is convex and $F_p^{\square}$ monotone. The assumptions on X and g imply the convexity of K. Because of Proposition 19, $\Sigma^{\square}(y)$ is convex $\forall y \in \mathbb{R}^n$. $p \in C^*$ and (10.2) imply

$$\langle p, [F(x') - F(x'')](x' - x'') \rangle \geq 0 \qquad \forall x', x'' \in \mathbb{R}^n \ ,$$

or

$$\langle p, F(x') - pF(x''), x' - x'' \rangle \geq 0 \ .$$

The equivalence is a classic fact [25]. $\square$

Proposition 21. If $p \in C^*$, then

$$\Sigma^{\square}(y) \subseteq \Sigma_p^{\square}(y) \quad , \quad y \in \Sigma^{\square}(y) \cap \Sigma_p^{\square}(y), \quad \forall y \in \mathbb{R}^n \quad . \tag{10.9}$$

Proof. $x \in \Sigma^{\square}(y) \ \Rightarrow \ \exists c \in C$ such that $F(x)(y - x) = c$. From this equality, having taken into account that $p \in C^*$ and $c \in C$ imply $\langle p, c \rangle \geq 0$, we have

$$\langle pF(x), y - x \rangle = \langle p, F(x)(y - x) \rangle = \langle p, c \rangle \geq 0 \ .$$

The 1st of (10.9) follows. The 2nd is a consequence of the obvious relations $y \in \Sigma^{\square}(y)$ (due to the closure of C) and $y \in \Sigma_p^{\square}(y)$. $\square$

Proposition 22. If $p \in \operatorname{int} C^*$, then (9.7) is impossible – and hence y is a solution to (9.11) – iff the system (in the unknown x):

$$(10.10) \quad \langle pF(x), y-x \rangle > 0 \ , \ F(x)(y-x) \in C \ , \ g(x) \geq 0 \ , \ x \in X$$

is impossible. Furthermore, the impossibility of (10.10) is a necessary and sufficient condition for y to be a solution to (10.8).

Proof. The 1st of (10.10) $\Rightarrow$ $\langle F(x), y-x \rangle \neq 0$, so that the possibility of (10.10) implies that of (9.7). The 1st of (9.7) and $p \in \operatorname{int} C^*$ imply the 1st of (10.10). The last part of the statement is obvious. $\square$

Note that Proposition 22 shows that Minty VVI (9.11) can be equivalently replaced by a scalar Quasi–Variational one, namely (10.8).

Consider, now, the function $\varphi^{\square} : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, defined by $\varphi^{\square}(x;y) := \langle F_p^{\square}(x), y-x \rangle$, and the problem (in the unknown x; $y \in K^{\square}(y)$ is a parameter):

$$(10.11) \qquad \max \varphi^{\square}(x;y) \quad , \quad \text{s.t. } x \in K^{\square}(y) \ .$$

Since $\varphi^{\square}(y;y) = 0 \ \forall y$, the maximum in (10.11) is ≥ 0.

Proposition 23. $y \in K$ is a solution to (9.11) iff it is a global maximum point of (10.11).

Proof. Let y be a maximum point of (10.11), so that $\varphi^{\square}(y;y) = 0$ and $\varphi^{\square}(x;y) \leq 0 \ \forall x \in K^{\square}(y)$. Hence, because of Proposition 22, y is a solution to (9.11). Now, let y be a solution to (9.11). Because of Proposition 22, system (10.10) is impossible, so that y is a solution to (10.8); this implies $\varphi^{\square}(x;y) \leq 0 \ \forall x \in K^{\square}(y)$. Since $\varphi^{\square}(y;y) = 0$, y is a global maximum point of (10.11). $\square$

11. SOME REMARKS ON PENALIZATION

In Sects. 2,3 and 9 we adopted mainly linear or piece–wise linear separation functions, but pointed out that any nonlinear one could be used equally well. In particular, if we had introduced a family of separation functions, whose positive level sets were ordered (in the sense of inclusion), then we might have hoped to define a *penalization approach*. It would have consisted in replacing each of (2.1), (2.3), (9.1), (9.9), (9.11), (9.12) with a sequence of VOP or VVI which had merely X as feasible region or domain and whose sets of solutions converge to that of above VOP or VVI. In this section we will outline a way for investigating such a topic, making some remarks for (2.1) and (9.1).

To this end, in place of (2.10), consider the class of functions

$$w(u,v;\alpha) = u - \Lambda(v;\alpha), \tag{11.1}$$

where $\alpha > 0$, $\Lambda(v;\alpha) = 0 \ \forall v \in \mathbb{R}^m_+$, $\Lambda(v;\alpha) \in C\backslash\{0\} \ \forall v \notin \mathbb{R}^m_+$. $\Lambda(v;\alpha)$ must fulfil the conditions:

$$\mathcal{H} = \bigcap_{\alpha>0} W_{C\backslash\{0\}}(u,v;\alpha) \ , \tag{11.2a}$$

$$\alpha' \leq \alpha'' \ \Rightarrow \ W_{C\backslash\{0\}}(u,v;\alpha'') \subseteq W_{C\backslash\{0\}}(u,v;\alpha') \ . \tag{11.2b}$$

where $W_{C\backslash\{0\}}(u,v;\alpha) := \{(u,v) \in \mathbb{R}^\ell \times \mathbb{R}^m : w(u,v;\alpha) \geq_{C\backslash\{0\}} 0\}$.

Let us consider the special case, where $C = \mathbb{R}^\ell_+$. A function of type (11.1) is obtained by setting $\tilde{\Lambda}(v;\mu) = \mu\Lambda(v)$ where $\mu > 0$ and

$$\Lambda(v) = (\Lambda_r(v), r \in \mathcal{I})^T, \tag{11.3a}$$

with

$$\Lambda_r(v) := \begin{cases} 0 & , \text{ if } v \in \mathbb{R}^m_+, \\ \sum\limits_{i=1}^m v_i^2 & , \text{ if } v \notin \mathbb{R}^m_+ . \end{cases} \tag{11.3b}$$

Proposition 24. Let $\Lambda(v)$ be defined by (11.3). The class of functions

$$w(u,v;\mu) := u - \mu\Lambda(v) \ , \ \mu > 0 \ ,$$

fulfils the conditions (11.2).

Proof. Let us prove (11.2a). Following the proof of Proposition 1 it is sufficient to prove that $\forall(\tilde{u},\tilde{v}) \notin \mathcal{H}\ \exists\, \tilde{\mu} > 0$ such that $(\tilde{u},\tilde{v}) \notin W_{C\setminus\{0\}}(u,v;\tilde{\mu})$. Consider any $(\tilde{u},\tilde{v}) \notin \mathcal{H}$. If $\tilde{v} \in \mathbb{R}^m_+$ then $\tilde{u} \notin \mathbb{R}^\ell_+ \;\Rightarrow\; w(\tilde{u},\tilde{v};\mu) = \tilde{u} \notin \mathbb{R}^\ell_+\setminus\{0\}$. If $\tilde{v} \notin \mathbb{R}^m_+$ then $\sum\limits_{i=1}^{m} \tilde{v}_i^2 > 0$. $\forall r \in \mathcal{I}$ we have:

$$w_r(\tilde{u},\tilde{v};\tilde{\mu}_r) = 0 \quad \text{if} \quad \tilde{\mu}_r := \frac{\tilde{u}_r}{\sum\limits_{i=1}^{m} \tilde{v}_i^2}\ ,$$

where w_r is the r–th component of w. Hence, if $\mu > \min\{\tilde{\mu}_1,\dots,\tilde{\mu}_\ell\}$, then $w(\tilde{u},\tilde{v};\mu) \notin \mathbb{R}^\ell_+$. (11.2a) follows. Consider any $\mu',\mu'' > 0$ such that $\mu' \le \mu''$; we have to prove that $\forall(u,v) \in \mathbb{R}^\ell \times \mathbb{R}^m$ $w(u,v;\mu'') \in \mathbb{R}^\ell_+\setminus\{0\}$ implies that $w(u,v;\mu') \in \mathbb{R}^\ell_+\setminus\{0\}$. If $v \in \mathbb{R}^m_+$, this implication is trivial. If $v \notin \mathbb{R}^m_+$, then $w(u,v;\mu'') \in \mathbb{R}^\ell_+\setminus\{0\}$ implies

$$w(u,v;\mu'') = u - \mu''\sum_{i=1}^{m} v_i^2 = \left(u - \mu'\sum_{i=1}^{m} v_i^2\right) - (\mu''-\mu')\sum_{i=1}^{m} v_i^2 \in \mathbb{R}^\ell_+\setminus\{0\}\ ,$$

so that

$$w(u,v;\mu') = u - \mu'\sum_{i=1}^{m} v_i^2 \in \{(\mu''-\mu')\sum_{i=1}^{m} v_i^2 + \mathbb{R}^\ell_+\setminus\{0\} \subseteq \mathbb{R}^\ell_+\setminus\{0\}\ .$$

Thus, (11.2b) follows. □

The properties (11.2) suggest the introduction of the following VOP (in the unknown x):

(11.4) $$\min_{C\setminus\{0\}}[f(x) + \mu\Lambda(g(x))]\ ,\ \text{s.t.}\ \ x \in X\ .$$

Consider a sequence $\{\mu_j\}_1^\infty$ of positive reals with $\lim\limits_{j\to+\infty} \mu_j = +\infty$. Call P_j the problem (11.4) at $\mu = \mu_j$. We aim to solve (2.1) by solving the sequence $\{P_j\}_1^\infty$.

Example 12. Let us set $\ell = 2, m = 1, X = \mathbb{R}_+, C = \mathbb{R}^2_+$, and

$$f_1(x) = 2x - x^2, f_2(x) = 1 - x^2, g(x) = 1 - x, f = (f_1, f_2) \ .$$

Problem (11.4) becomes:

$$\min_{\mathbb{R}^2_+ \setminus \{0\}} \Big(f_1(x) + \mu\Lambda(g(x)), f_2(x) + \mu\Lambda(g(x)) \Big) \ , \ \text{s.t.} \ \ x \geq 0 \ ,$$

where

$$f_1(x) + \mu\Lambda(g(x)) = \begin{cases} 2x - x^2 \ , & \text{if } 0 \leq x \leq 1, \\ (\mu - 1)x^2 + (2 - 2\mu)x + \mu \ , & \text{if } x > 1, \end{cases}$$

$$f_2(x) + \mu\Lambda(g(x)) = \begin{cases} 1 - x^2 \ , & \text{if } 0 \leq x \leq 1, \\ (\mu - 1)x^2 - 2\mu x + 1 + \mu \ , & \text{if } x > 1. \end{cases}$$

It is easy to find that, if $\mu > 1$, then the set of its v.m.p. is $\left[0, \frac{\mu}{\mu - 1}\right]$, which converges to $[0, 1]$, the set of v.m.p. of (2.1), as $\mu \to +\infty$.

Of course, (11.3) is not the only possibility for satisfying (11.2). Furthermore, we can start from the scalar separation (3.1) – instead of (2.10) – and associate with (2.1) a sequence of scalar minimum problems – instead of vector ones. Similar analysis can be done in the weak case.

Taking the above line, considering nonlinear separation functions, it is possible to extend the penalty methods also to a VVI. Similarly to VOP we can define a sequence of VVI on X, whose sets of solutions converge to that of (9.1). Consider the class of functions (11.1):

$$\omega = \omega(u, v; \alpha) = u - \varphi(v; \alpha),$$

where $\varphi(\cdot, \alpha) : \mathbb{R}^m \to \mathbb{R}^\ell, \ \forall \alpha > 0$ and $\varphi(v; \alpha) = 0 \ \forall v \in \mathbb{R}^m_+$, $\varphi(v; \alpha) \in C \backslash \{0\} \ \forall v \notin \mathbb{R}^m_+$. Suppose that ω fulfils the conditions (11.2). As shown by Proposition 24, a class of functions which fulfil the previous conditions is obtained by defining $\varphi(v; \alpha)$ as in (11.3). Because of the properties (11.2) we can replace (9.1) with the sequence of VVI defined by the following problems depending on the parameter $\alpha > 0$: find $y \in X$, such that

$$F(y)(y - x) - \varphi(g(x); \alpha) \not\geq_{C \backslash \{0\}} 0, \quad \forall x \in X \ . \tag{11.4}$$

We observe that, if a solution y to (11.4) belongs to the feasible set K, then it is also a solution to (9.1); in fact, by definition, it is $\varphi(g(x);\alpha) = 0, \ \ \forall x \in K$. If it were $F(y)(y - \bar{x}) \geq_{C\setminus\{0\}} 0$ for $\bar{x} \in K$, then

$$F(y)(y - \bar{x}) - \varphi(g(\bar{x});\alpha) \geq_{C\setminus\{0\}} 0 \ ,$$

which is against (11.4). In this case we have an exact penalization for a VVI.

REFERENCES

[1] ABADIE J., "On the Kuhn–Tucker Theorem". In "Nonlinear programming". North–Holland, Amsterdam, 1967, pp. 19–36.

[2] BENSON H.P., "Hybrid approach for solving multiple–objective linear programs in outcome space". Jou. of Optimiz. Theory Appls., Vol. 98, No. 1, 1998, pp. 17–35.

[3] BENSON H.P. and LEE D., "Outcome–based algorithm for optimizing over the efficient set of a bicriteria linear programming problem". Jou. of Optimiz. Theory Appls., Vol. 88, No. 1, 1996, pp. 77–105.

[4] BIGI G., "Lagrangian Functions and Saddle Points in Vector Optimization". Submitted to Optimization.

[5] BIGI G. and PAPPALARDO M., "Regularity conditions in Vector Optimization". Jou. of Optimiz. Theory Appls., Plenum, New York, Vol. 102, No.1, 1999, pp. 83–96.

[6] CASTELLANI M., MASTROENI G. and PAPPALARDO M., "On regularity for generalized systems and applications". In "Nonlinear Optimization and Applications", G. Di Pillo et al. (Eds.), Plenum Press, New York, 1996, pp. 13–26.

[7] CONTI R. et al. (Eds.), "Optimization and related fields". Lecture Notes in Mathematics No. 1190, Springer–Verlag, Berlin, 1986, pp. 57–93.

[8] CRAVEN B.D., "Nonsmooth multiobjective programming". Numer. Funct. Anal. and Optimiz., Vol. 10, 1989, pp. 49–64.

[9] DAUER J.P., "Analysis of the objective space in multiple objective linear programming". Jou. Mathem. Analysis and Appls., Vol. 126, 1987, pp. 579–593.

[10] DAUER J.P., "Solving multiple objective linear programs in the objective space". Europ. Jou. of Operational Research, Vol. 46, 1990, pp. 350–357.

[11] DIEN P.H., MASTROENI G., PAPPALARDO M. and QUANG P.H., "Regularity conditions for constrained extremum problems via image space: the linear case". In Lecture Notes in Sc. and Mathem. Systems, No. 405, Komlosi, Rapcsáck, Schaible Eds., Springer–Verlag, 1994, pp. 145–152.

[12] DINH THE LUC, "Theory of Vector Optimization". Lecture Notes in Ec. and Mathem. Systems No. 319, Springer–Verlag, Berlin, 1989.

[13] DI PILLO G. et al. (Eds.), “Nonlinear optimization and Applications”. Plenum, New York, 1996, pp. 13–26 and 171–179.

[14] FAVATI P. and PAPPALARDO M.,“On the reciprocal vector optimization problems”. Jou. Optimiz. Theory Appls., Plenum, New York, Vol. 47, No. 2, 1985, pp. 181–193.

[15] FERREIRA P.A.V. and MACHADO M.E.S., “Solving multiple–objective problems in the objectice space”. Jou. Optimiz. Theory Appls., Plenum, New York, Vol. 89, No. 3, 1996, pp. 659–680.

[16] GALPERIN E.A., “Nonscalarized multiobjective global optimization”. Jou. Optimiz. Theory Appls., Plenum, New York, Vol. 75, No. 1, 1992, pp. 69–85.

[17] GALPERIN E.A., “Pareto analysis vis–à–vis balance space approach in multiobjective global optimization”. Jou. Optimiz. Theory Appls., Vol. 93, No. 3, 1997, pp. 533–545.

[18] GIANNESSI F., “Theorems of the alternative, quadratic programs and complementarity problems”. In “Variational Inequalities and complementarity problems”, R.W. Cottle et al. Eds., J. Wiley, 1980, pp. 151–186.

[19] GIANNESSI F., “Theorems of the alternative and optimality conditions”.Jou. Optimiz. Theory Appls., Plenum, New York, Vol. 42, No. 11, 1984. pp. 331–365.

[20] GIANNESSI F., “Semidifferentiable functions and necessary optimality conditions”. Jou. Optimiz. Theory Appls., Vol. 60, 1989, pp. 191–241.

[21] GIANNESSI F., “On Minty variational principle”. In “New trends in mathematical programming”, F. Giannessi, S. Komlosi and T. Rapcsáck Eds., Kluwer Acad. Publ., Dordrecht, 1998, pp. 93–99.

[22] GIANNESSI F. and MAUGERI A. (Eds.), “Variational Inequalities and Networks equilibrium problems”. Plenum, New York, 1995, pp. 1–7, 21–31, 101–121 and 195–211.

[23] GIANNESSI F. and PELLEGRINI L., “Image space Analysis for Vector Optimization and Variational Inequalities. Scalarization”. In “Advances in Combinatorial and Global Optimization”, A. Migdalas, P. Pardalos and R. Burkard, Eds., Worlds Scientific Publ., To appear.

[24] ISERMANN H., “On some relations between a dual pair of multiple objective linear programs”. Zeitschrift für Operations Re-

search, Vol. 22, 1978, pp. 33–41.

[25] KINDERLEHERER D. and STAMPACCHIA G., "An introduction to Variational inequalities". Academic Press, New York, 1980.

[26] KOMLOSI S., "On the Stampacchia and Minty Vector Variational Inequalities". In "Generalized Convexity and Optimization for economic and financial decisions", G. Giorgi and F. Rossi Eds., Editrice Pitagora, Bologna, Italy, 1999, pp. 231–260.

[27] LEITMANN G., "The Calculus of Variations and Optimal Control". Plenum Press, New York, 1981.

[28] MAEDA T., "Constraints Qualifications in Multiobjective Optimization Problems: Differentiable Case". Jou. of Optimiz. Theory and Appls., Vol. 80, No. 3, 1994, pp. 483–500.

[29] MANGASARIAN O.L., "Nonlinear Programming". Series "Classics in Applied Mathematics", No.10, SIAM, Philadelphia, 1994.

[30] MARTEIN L., "Stationary points and necessary conditions in Vector extremum problems". Tech. Report No. 133, Dept. of Mathem., Optimiz. Group, Univ. of Pisa, 1986. Published with the same title in Jou. of Informations and Optimization Sciences, Vol. 10, No. 1, 1989, pp. 105–128.

[31] MARTEIN L., "Lagrange multipliers and generalized differentiable functions in vector extremum problems". Tech. Report No. 135, Dept. of Mathem., Optimiz. Group, Univ. of Pisa, 1986. Published with the same title in Jou. of Optimiz. Theory Appls., Vol. 63, No. 2, 1989, pp. 281–297.

[32] MASTROENI G., "Separation methods for Vector Variational Inequalities. Saddle–point and gap function". To appear in "Nonlinear Optimization and Applications *2*", G. Di Pillo at al. (Eds.), Kluwer Acad. Publ., Dordrecht, 1999.

[33] MASTROENI G. and RAPCSÁK T., "On Convex Generalized Systems". Jou. of Optimiz. Theory and Appls., 1999. To appear.

[34] PAPPALARDO M., "Some Calculus Rules for semidifferentiable functions and related topics". In "Nonsmooth Optimization. Methods and Applications", F. Giannessi Ed., Gordon & Breach, 1992, pp. 281–294.

[35] PAPPALARDO M., "Stationarity in Vector Optimization". Rendiconti del Circolo Matematico di Palermo, Serie II, No. 48,

1997, pp. 195–200.

[36] PASCOLETTI A. and SERAFINI P., “Scalarizing Vector Optimization problems”. Jou. Optimiz. Theory Appls., Plenum, New York, Vol. 42, No. 4, 1984, pp. 499–523.

[37] RAPCSÁCK T., “Smooth Nonlinear Optimization in $\mathbb{R}^n$”. Series “Nonconvex Optimization and its Applications”, No. 19, Kluwer Acad. Publ., Dordrecht, 1997.

[38] SONG W., “Duality for Vector Optimization of Set Valued Functions”. Jou. of Mathem. Analysis and Appls., Vol. 201, 1996, pp. 212–225.

[39] SAWARAGI Y., NAKAYAMA H. and TANINO T., “Theory of multiobjective Optimization”. Academic Press, New York, 1985

[40] TARDELLA F., “On the image of a constrained extremum problem and some applications to the existence of a minimum”. Jou. of Optimiz. Theory Appls., Plenum, New York, Vol. 60, No.1, 1989, pp. 93–104.

[41] WANG S. and LI Z., “Scalarization and Lagrange duality in multiobjective optimization”. Optimization, Vol. 26, Gordon and Breach Publ., 1992, pp. 315–324.