

ON GENERALIZED CONSTRAINED OPTIMIZATION AND SEPARATION THEOREMS

Letizia Pellegrini

Abstract. In this paper a generalized format for a constrained extremum problem is considered. Subsequently, the paper investigates and deepens some aspects concerning the linear separation between two sets in the Euclidean space, that are a convex cone and a generic set. A condition equivalent to their linear separation is given. Moreover, a condition equivalent to regular linear separation is proposed; this condition includes also the nonconvex case and it is finalized to the application to the generalized constrained extremum problems.

1. INTRODUCTION

The Separation Theory, which is a fundamental topic in Mathematics, takes a particular importance in Optimization since it can be assumed as basis for the theory of constrained extremum problems and related fields. For example, the necessary optimality conditions for an extremum problem can be expressed in terms of separation between two suitable sets defined by the objective function and the constraints. In particular, in the image space associated with a constrained extremum problem [4], these two sets are a suitable conic approximation of the image of the functions involved in the problem and a convex cone that only depends on the type of conditions, which define the feasible region of the extremum problem.

A classic kind of separation between two sets is the linear separation obtained by means of a hyperplane that separates the space in two halfspaces, each of them containing one of the two sets. In this paper, we aim at deepening the analysis of the linear separation between two particular sets in the Euclidean space. In Sect.2, we give a condition equivalent to the linear separation between a convex cone C and a generic set S ; this condition can be called of "Helly-type" because, if each

Received August 14, 2009, accepted September 18, 2009.

Communicated by F. Giannessi.

2000 *Mathematics Subject Classification*: 90C, 49K.

Key words and phrases: Separation, Optimality conditions.

subset of S of finite cardinality enjoys a separability property, then S itself enjoys a separability property. The relationships between linear separation and regularity are clarified at the beginning of Sect.3, where we propose a condition equivalent to the regular linear separation between C and S ; it is expressed in terms of the tangent cone to a suitable approximation of the set, which allows us to include also the nonconvex case. All these results are illustrated through several examples.

We stress the fact that the separation results obtained in this work have been conceived for their application to constrained optimization. Hence, now, we introduce a well known format of the constrained extremum problem and we briefly recall the separation approach based on the analysis in the image space [4]. For this, assume we are given the integer m , with $m \geq 0$, the nonempty subset X of a Banach space B and the functions $f : X \rightarrow \mathbb{R}$, $g : X \rightarrow \mathbb{R}^m$. We consider the following constrained extremum problem

$$(1) \quad \min f(x), \text{ subject to } x \in R := \{x \in X : g(x) \in D\},$$

where D is a closed convex cone in \mathbb{R}^m . The format (1) has been largely considered in the literature (see Chapter 3, Sect.3, page 178 of [1]; Chapter 3 of [2]; Sect.4 of [5]). It embeds several particular formulations, including the classic case where, given the integer p with $0 \leq p \leq m$, the condition $g(x) \in D$ is:

$$(2) \quad g_i(x) = 0, \quad i \in \mathcal{I}^0 := \{1, \dots, p\}, \quad g_i(x) \geq 0, \quad i \in \mathcal{I}^+ := \{p+1, \dots, m\}.$$

It is immediate to observe that $\bar{x} \in R$ is a (global) minimum point of (1) iff $\mathcal{H} \cap \mathcal{K}_{\bar{x}} = \emptyset$, where $\mathcal{H} := \{(u, v) \in \mathbb{R} \times \mathbb{R}^m : u > 0, v \in D\}$ and $\mathcal{K}_{\bar{x}} := \{(u, v) \in \mathbb{R} \times \mathbb{R}^m : u = f(\bar{x}) - f(x), v = g(x), x \in X\}$. Obviously, a sufficient condition for the disjunction of \mathcal{H} and $\mathcal{K}_{\bar{x}}$, and hence for the optimality of \bar{x} , is the existence of a hyperplane such that \mathcal{H} and $\mathcal{K}_{\bar{x}}$ are contained in two of its disjoint level sets. Such a condition is also necessary for the optimality, under regularity assumptions; i.e., under a condition guaranteeing that the separation hyperplane does not contain $\mathcal{H}_u := \{(u, v) \in \mathbb{R} \times \mathbb{R}^m : u = 0, v \in D\}$. An important case is that where $\mathcal{K}_{\bar{x}}$ is replaced by its linear approximation or, more generally, by its homogenization [3]. In such a case, when D has nonempty interior – as, for example, when $p = 0$ in (2) – the disjunction between \mathcal{H} and $\mathcal{K}_{\bar{x}}$ implies the disjunction between \mathcal{H} and the approximation of $\mathcal{K}_{\bar{x}}$. Unfortunately, this implication does not hold anymore when the interior of the cone D is empty (see Example 5.10 of [7]); in other words, there are cases where the optimality of \bar{x} for (1) does not imply the optimality for the linearized problem.

Now, we recall the main notations and definitions that will be used in the sequel. O_n denotes the n -tuple, whose entries are zero; when there is no fear of confusion the suffix is omitted; for $n = 1$, the 1-tuple is identified with its element, namely, we set $O_1 = 0$; $\langle \cdot, \cdot \rangle$ is the usual scalar product in \mathbb{R}^n . Let $M \subseteq \mathbb{R}^n$; $\dim M$,

$\text{cl } M$, $\text{conv } M$, $\text{aff } M$, $\text{int } M$ and $\text{ri } M$ denote the dimension, the closure, the convex hull, the affine hull, the interior and the relative interior of M , respectively. The vectors $k_1, \dots, k_{m+1} \in \mathbb{R}^n$, with $m \leq n$ are *affinely independent* iff $\dim \text{aff } \{k_1, \dots, k_{m+1}\} = m$. If $\bar{x} \in \mathbb{R}^n$ and $M \neq \{\bar{x}\}$, the *cone generated by M from \bar{x}* is the set

$$\text{cone } (\bar{x}; M) := \{x \in \mathbb{R}^n : x = \bar{x} + \alpha(y - \bar{x}), y \in M, \alpha > 0\}.$$

If $M \neq \emptyset$ and $\bar{x} \in \text{cl } M$, then the set of $\bar{x} + x \in \mathbb{R}^n$ for which $\exists \{x^i\} \subseteq \text{cl } M$, with $\lim_{i \rightarrow +\infty} x^i = \bar{x}$, and $\exists \{\alpha_i\} \subset \mathbb{R}_+ \setminus \{0\}$ such that $\lim_{i \rightarrow +\infty} \alpha_i(x^i - \bar{x}) = x$, is called *tangent cone* to M at \bar{x} and denoted by $TC(\bar{x}; M)$. We stipulate that $TC(\bar{x}; \emptyset) = \emptyset$. If $\bar{x} = O$, then \bar{x} is omitted from the notation of the cones. For a cone C with apex at \bar{x} , the (positive) *polar cone* associated to C is

$$C^* := \{x \in \mathbb{R}^n : \langle x, y - \bar{x} \rangle \geq 0, \forall y \in C\}.$$

Let $a \in \mathbb{R}^n \setminus \{O\}$ and $b \in \mathbb{R}$; in the sequel we will consider the hyperplane

$$H^0 := \{x \in \mathbb{R}^n : \langle a, x \rangle = b\}$$

and the related halfspaces

$$H^- := \{x \in \mathbb{R}^n : \langle a, x \rangle \leq b\}, \quad H^+ := \{x \in \mathbb{R}^n : \langle a, x \rangle \geq b\}.$$

Definition 1.1. The nonempty sets K_1 and $K_2 \subset \mathbb{R}^n$ are *linearly separable* iff there exists a hyperplane $H^0 \subset \mathbb{R}^n$, such that:

$$(3) \quad K_1 \subseteq H^-, \quad K_2 \subseteq H^+.$$

H^0 is called *separating hyperplane*. The separation is *strict* iff

$$K_1 \subseteq \text{int } H^-, \quad K_2 \subseteq \text{int } H^+;$$

proper iff, besides (3), we have

$$K_1 \cup K_2 \not\subseteq H^0.$$

Definition 1.2. A hyperplane $H^0 \subset \mathbb{R}^n$ is called *supporting hyperplane* of $K \subset \mathbb{R}^n$, iff

$$K \subseteq H^+, \text{ (or } K \subseteq H^-) \quad \text{and} \quad H^0 \cap \text{cl } K \neq \emptyset.$$

Definition 1.3. Let $K \subset \mathbb{R}^n$. $F \subset \text{cl } K$ is a *face* of K iff it is the intersection of $\text{cl } K$ with a supporting hyperplane H^0 of K , or

$$F := H^0 \cap \text{cl } K.$$

2. A HELLY-TYPE CONDITION FOR LINEAR SEPARATION

In this section, we will give a necessary and sufficient condition for the linear separation between two sets of \mathbb{R}^n and, in a particular case, sufficient for their proper separation. We suppose that one of the two sets is a nonempty convex cone C with apex at $O \notin C$, and the other is any nonempty subset S of \mathbb{R}^n ; set $s := \dim S$. Let $z \in \mathbb{R}^n$; denote by $\text{proj } z$ its projection on $C^\perp := \{x \in \mathbb{R}^n : \langle x, k \rangle = 0, \forall k \in C\}$, the orthogonal complement of C . Let $p := \dim C^\perp$ so that $\dim C = n - p$. Let us consider the following condition.

Condition 2.1. For every set $\{z^1, \dots, z^{s+1}\}$ of affinely independent vectors of S , such that

$$(4) \quad \left. \begin{array}{l} \dim \text{conv} \{\text{proj } z^1, \dots, \text{proj } z^{s+1}\} = p \text{ and} \\ O \in \text{ri conv} \{\text{proj } z^1, \dots, \text{proj } z^{s+1}\} \end{array} \right\},$$

we have:

$$(5) \quad (\text{ri } C) \cap \text{ri conv} \{z^1, \dots, z^{s+1}\} = \emptyset.$$

In the above Condition 2.1, if $p = 0$, we stipulate that (4) – (5) shrinks to (5). We stipulate also that a singleton coincides with its relative interior. Let us consider the case $p > 0$: every set $\{z^1, \dots, z^{s+1}\}$ of affinely independent vector of S satisfying (4) is a set for which we have to check condition (5), that is equivalent to the linear separation between C and $\{z^1, \dots, z^{s+1}\}$. In other words, it is enough to check condition (5) for every set $\{z^1, \dots, z^{s+1}\}$ "representative" of S with respect to the linear separation from C . If it is not possible to find such a set, then, of course, condition (4) – (5) is meant to be satisfied; otherwise, Condition 2.1 requires that (4) implies (5). This distinction originates two possible cases where Condition 2.1 is fulfilled, that are treated separately in the following lemma. For a geometric interpretation of Condition 2.1, see Example 4.6.1 on page 278 of [4].

Lemma 2.1. *Suppose that C^\perp be a coordinate subspace of \mathbb{R}^n of dimension p such that $1 \leq p \leq s$. If Condition 2.1 holds, then C and S are linearly separable and, moreover, the separation is proper.*

Proof. There are two possible cases that we consider separately.

(A) (4) does not hold, in the sense that no set of affinely independent vectors of S verifies (4). Denote by $\text{proj } S \subset \mathbb{R}^n$ the projection of S on C^\perp . Since for every set of $s + 1$ affinely independent vectors of S , relation (4) does not hold, then

$$(6) \quad O \notin \text{ri conv proj } S.$$

In fact, if ab absurdo $O \in \text{ri conv proj } S$, then $\exists \alpha_1, \dots, \alpha_{p+1} > 0$ with $\sum_{i=1}^{p+1} \alpha_i = 1$ and $\exists x^1, \dots, x^{p+1} \in \text{proj } S$ affinely independent, such that $O = \sum_{i=1}^{p+1} \alpha_i x^i$. Thus, we would have $p + 1$ affinely independent vectors of S such that $x^i = \text{proj } z^i, z^i \in S, i = 1, \dots, p + 1$ and $O = \sum_{i=1}^{p+1} \alpha_i \text{proj } z^i$. Since $\dim S = s$, then the set $\{z^1, \dots, z^{p+1}\}$ could be augmented (if $p < s$) to form a set $\{z^1, \dots, z^{s+1}\}$ of affinely independent vectors of S which would satisfy (4), this contradicts the initial assumption. Since C^\perp is a coordinated subspace, then (6) becomes:

$$O_p \notin \text{int conv proj } S.$$

Applying the Hahn-Banach Theorem, we get the existence of a hyperplane of \mathbb{R}^p through O_p whose equation is $\sum_{i=1}^p a_i x_i = 0$ and such that $\sum_{i=1}^p a_i w_i \leq 0, \forall (w_1, \dots, w_p) \in \text{conv proj } S$. Setting $a_i = 0, i = p + 1, \dots, n$, it follows that $\sum_{i=1}^n a_i w_i \leq 0, \forall (w_1, \dots, w_n) \in \text{conv } S$ because conv and proj are permutable.

The hyperplane $H^0 = \{x \in \mathbb{R}^n : \sum_{i=1}^n a_i x_i = 0\}$ contains the cone C and therefore separates C and S . Moreover, the separation is proper since S cannot be included in the hyperplane H^0 , otherwise (6) would be contradicted.

(B) (4) holds, in the sense that there exists a set $\{z^1, \dots, z^{s+1}\}$ of affinely independent vectors of S which verifies (4). We prove that (5) implies

$$(7) \quad \text{ri } C \cap \text{ri conv } S = \emptyset.$$

Suppose that (7) does not hold, i.e., there exists $\bar{z} \in \text{ri } C \cap \text{ri conv } S$. Because of a well known Carathéodory Theorem, \bar{z} can be expressed as a convex combination of $s + 1$ affinely independent vectors of S , say w^1, \dots, w^{s+1} , that is $\bar{z} = \sum_{j=1}^{s+1} \alpha_j w^j,$

with $\alpha_j > 0, \forall j = 1, \dots, s + 1$ and $\sum_{j=1}^{s+1} \alpha_j = 1$. If these vectors verify (4), then (5) is contradicted. Therefore we have:

$$O \notin \text{ri conv } \{\text{proj } w^1, \dots, \text{proj } w^{s+1}\}.$$

Since C^\perp is a coordinated subspace, the previous relation becomes:

$$O_p \notin \text{int conv}\{\text{proj } w^1, \dots, \text{proj } w^{s+1}\}$$

and thus there exists $(a_1, \dots, a_p) \neq O_p$ with $\sum_{i=1}^p a_i(\text{proj } w^j)_i \leq 0, \forall j = 1, \dots, s+1$.

If we set $a_i = 0, \forall i = p+1, \dots, n$ we get $\sum_{i=1}^n a_i(w^j)_i \leq 0$, and therefore also

$\sum_{i=1}^n a_i \alpha_j(w^j)_i \leq 0, \forall j = 1, \dots, s+1$. On the other hand, $\bar{z} \in \text{ri } C$ and thus

$\langle a, \bar{z} \rangle = 0$. Since the coefficients α_i are all positive, it follows that $\sum_{i=1}^n a_i(w^j)_i = 0,$

$\forall j = 1, \dots, s+1$. This implies that $\sum_{i=1}^p a_i(w^j)_i = 0, \forall j = 1, \dots, s+1$, which contradicts $O_p \notin \text{int conv}\{\text{proj } w^1, \dots, \text{proj } w^{s+1}\}$. Therefore (7) is satisfied and this implies proper separation between C and S . ■

Theorem 2.2. *C and S are linearly separable, if and only if Condition 2.1 holds. The separation is proper if $0 \leq p \leq s$.*

Proof. If. In the proof of the sufficiency we will consider three different cases.

(A) $p = 0$. C is a convex body and thus, obviously, (5) implies linear separation (even proper) between C and S .

(B) $0 \leq s \leq p-1$. Let B_C and B_S be bases for $\text{aff } C$ and $\text{aff } S$, respectively; $\dim B_C = n-p, \dim B_S = s$ and $\dim(B_C \cup B_S) \leq n-p+s \leq n-1$. This shows that there exists a hyperplane of \mathbb{R}^n which contains C and is parallel to $\text{aff } S$, so that separation holds.

(C) $1 \leq p \leq s$. If C^\perp is a coordinated subspace, then by Lemma 2.1 it results that C and S are linearly separable and, moreover, that the separation is proper. If C^\perp is not a coordinated subspace, then by its definition we have that there exists a suitable rotation ρ which transforms C into a cone C^ρ such that $(C^\rho)^\perp$ is a coordinated subspace and, after having applied the rotation ρ , by Lemma 2.1 we obtain (proper) separation between C and S .

Only if. By assumption, $\exists a \in \mathbb{R}^n \setminus \{O\}$ and $b \in \mathbb{R}$, such that

$$\langle a, x \rangle \geq b, \forall x \in C \quad \text{and} \quad \langle a, y \rangle \leq b, \forall y \in S.$$

Since $O \in \text{cl } C$, we can put $b = 0$. Set $H^0 := \{x \in \mathbb{R}^n : \langle a, x \rangle = 0\}$. If no set of $s+1$ affinely independent vectors of S exists, such that (4) is satisfied, then the thesis is trivial. Let us assume that there exists a set $\{z^1, \dots, z^{s+1}\}$ of affinely independent vectors of S such that (4) holds while (5) is not valid, i.e.

$$(8) \quad O \in \text{ri conv } \{\text{proj } z^1, \dots, \text{proj } z^{s+1}\},$$

and

$$(9) \quad (\text{ri } C) \cap \text{ri conv } \{z^1, \dots, z^{s+1}\} \neq \emptyset.$$

Let \bar{z} belong to the left-hand side of (9); thus there exist $\alpha_i > 0, i = 1, \dots, s + 1$ with $\sum_{i=1}^{s+1} \alpha_i = 1$ such that $\bar{z} = \sum_{i=1}^{s+1} \alpha_i z^i \in \text{ri } C$. From $\bar{z} \in \text{ri } C$ we have $\text{proj } \bar{z} = O$ and from (8) we have that there exists $J \subseteq \{1, \dots, s + 1\}$ with $\text{card } J = p + 1$ such that $\text{proj } z^i \neq O$ for $i \in J$. Therefore, it results $\text{proj } \bar{z} = \text{proj} \left(\sum_{i=1}^{s+1} \alpha_i z^i \right) = \sum_{i=1}^{s+1} \alpha_i \text{proj } z^i = \sum_{i \in J} \alpha_i \text{proj } z^i = O$. Since $z^1, \dots, z^{s+1} \in S$, then $\langle a, z^i \rangle \leq 0, i = 1, \dots, s + 1$ and hence $\langle a, \sum_{i=1}^{s+1} \alpha_i z^i \rangle \leq 0$. On the other hand, $\bar{z} \in \text{ri } C$ and thus $\langle a, \sum_{i=1}^{s+1} \alpha_i z^i \rangle \geq 0$. It follows $\bar{z} \in H^0$. From $\bar{z} \in \text{ri } C$ and C convex, we have that $\exists \beta_i > 0, i = 1, \dots, n - p + 1$ with $\sum_{i=1}^{n-p+1} \beta_i = 1$ and $\exists k^i \in C, i = 1, \dots, n - p + 1$ affinely independent, such that $\bar{z} = \sum_{i=1}^{n-p+1} \beta_i k^i$. Since $\bar{z} \in H^0$, then $\sum_{i=1}^{n-p+1} \beta_i \langle a, k^i \rangle = 0$, which implies $\langle a, k^i \rangle = 0, i = 1, \dots, n - p + 1$. Thus, $\text{conv } \{k^1, \dots, k^{n-p+1}\} \subseteq H^0$ and, consequently, $C \subseteq H^0$; it follows that $a \in C^\perp$. Moreover, from $S \subseteq H^-$ we have $\text{proj } S \subseteq H^-$. Using $O = \text{proj } \bar{z}$, we obtain

$$\langle a, O \rangle = \langle a, \text{proj } \bar{z} \rangle = \langle a, \sum_{i=1}^{s+1} \alpha_i \text{proj } z^i \rangle = \sum_{i=1}^{s+1} \alpha_i \langle a, \text{proj } z^i \rangle.$$

Since $\alpha_i > 0, i = 1, \dots, s + 1$, we get $\langle a, \text{proj } z^i \rangle = 0, i = 1, \dots, s + 1$; hence we have also $\{\text{proj } z^1, \dots, \text{proj } z^{s+1}\} \subseteq H^0$ and, obviously, $\text{conv } \{\text{proj } z^1, \dots, \text{proj } z^{s+1}\} \subseteq H^0$. Let us denote by $B(O_n, \varepsilon)$ an open ball of center O_n and radius $\varepsilon > 0$ in \mathbb{R}^n such that $\dim B(O_n, \varepsilon) = p$. From (8) we have that $\exists \bar{\varepsilon} > 0$ such that

$$B(O_n, \bar{\varepsilon}) \subseteq \text{conv } \{\text{proj } z^1, \dots, \text{proj } z^{s+1}\} \subseteq H^0,$$

i.e. $\langle a, y \rangle = 0, \forall y \in B(O_n, \bar{\varepsilon})$. By assumption $a \neq O$; hence, for $\gamma := \frac{\bar{\varepsilon}}{\|a\|} > 0$, it turns out $\bar{y} := \frac{1}{2} \gamma a \in B(O_n, \bar{\varepsilon})$. Consequently, we have

$$0 = \langle a, \bar{y} \rangle = \frac{\gamma}{2} \langle a, a \rangle = \frac{\gamma}{2} \|a\|^2,$$

which contradicts the assumption $a \neq O$. ■

A classic result about separation and proper separation between convex sets is given by the following theorem (see Theorem 2.39 of [9]).

Theorem 2.3. *Two nonempty, convex sets C_1 and C_2 in \mathbb{R}^n are linearly separable, if and only if $O \notin \text{int}(C_1 - C_2)$. The separation must be proper if also $\text{int}(C_1 - C_2) \neq \emptyset$.*

Remark 2.4. Obviously

$$(10) \quad O \notin \text{int}(\text{conv } S - C)$$

is equivalent to Condition 2.1, because both are equivalent to the linear separation between $\text{conv } S$ and the (convex) cone C . Hence, it is necessary to compare such two conditions. The introduction of Condition 2.1 is motivated by the study of necessary optimality conditions in constrained optimization by means of the separation approach. In view of applying separation results to the constrained extremum problem (1), the cone C that we will consider is $\mathcal{H} := \{u \in \mathbb{R} : u > 0\} \times D$. Condition 2.1 is articulated in such a way of distinguish the case where $\text{int } C$ is empty from the case where it is nonempty. Hence, when $\text{int } C = \emptyset$, Condition 2.1 can be used, while (10) cannot. Obviously, if $\text{int } D = \emptyset$ then also $\text{int } C = \emptyset$. A particular case where $\text{int } D = \emptyset$ is given by (2) with $p > 0$; in fact, in this case we have $D = \{O_p\} \times \mathbb{R}_+^{m-p}$. Nevertheless, in general, Condition 2.1 does not require that the cone D be the cartesian product between a suborthant and the origin of its orthogonal complement.

Remark 2.5. Both in Theorem 2.2 and Theorem 2.3 there is a sufficient condition for proper separation. First of all, observe that the sufficient condition $\text{int}(\text{conv } S - C) \neq \emptyset$ implies $0 \leq p \leq s$; in fact, if this double inequality does not hold, then $p \geq s+1 > 0$ so that $\text{int}(\text{conv } S - C) = \emptyset$. In general, the converse implication does not hold, as can be easily shown by choosing $C = \{x \in \mathbb{R}^3 : x_1 > 0, x_2 \geq 0, x_3 = 0\}$ and $S = \{x \in \mathbb{R}^3 : x_1 = -x_2, x_3 = 1\}$. Nevertheless, in this example, the distance $d(S, C)$ between S and C is positive and this implies the strict separation, and hence, trivially, the proper separation. A different example could be the one where $C = \{x \in \mathbb{R}^3 : x_1 > 0, x_2 \geq 0, x_3 = 0\}$ and $S = \{x \in \mathbb{R}^3 : x_1 = -x_2, x_3 = 0\}$; also this case is not very interesting, since it is enough to consider the smallest linear subspace containing both S and C and there Theorem 2.3 does apply. No other cases are included in the condition $0 \leq p \leq s$ and not in the condition $\text{int}(\text{conv } S - C) \neq \emptyset$, as it is proved in the next proposition.

Proposition 2.6. *Let us suppose that C and S are linearly separable and moreover that $0 \leq p \leq s$ and $\text{int}(\text{conv } S - C) = \emptyset$. Then one of the following conditions holds:*

- (i) $d(\text{conv } S, C) > 0$ and hence the separation between S and C is strict;
- (ii) the smallest linear subspace containing $\text{conv } S$ and C has dimension $< n$ and in this subspace $\text{int}(\text{conv } S - C) \neq \emptyset$.

Proof. If $d(\text{conv } S, C) > 0$, then (i) is proved. Hence, let us suppose $d(\text{conv } S, C) = 0$. Since $\text{int}(\text{conv } S - C) = \emptyset$, we have that $s < n$ and $n - p < n$. These two inequalities and $d(\text{conv } S, C) = 0$ imply that the smallest linear subspace containing $\text{conv } S$ and C has dimension $k < n$ and $\dim S = k$ or $\dim C = k$. Therefore, we obtain $\text{int}(\text{conv } S - C) \neq \emptyset$ in the subspace. ■

3. REGULAR SEPARATION BETWEEN A SET AND A FACE OF A CONE

As already observed, the separation approach based on the analysis in the image space can be applied to the study of necessary optimality conditions for the constrained extremum problem (1). In this approach, the image $\mathcal{K}_{\bar{x}}$ is replaced by a suitable approximation of $\mathcal{K}_{\bar{x}}$, as, for example, its linearization or, more generally, its homogenization [3]. If the approximation of $\mathcal{K}_{\bar{x}}$ contains \mathcal{H}_u , then it can happen that the existence of a minimum point of (1) does not imply the existence of Lagrange multipliers; in other words, a necessary optimality condition is not fulfilled in correspondence of a minimum point. A condition which guarantees the existence of Lagrange multipliers is called regularity condition (or constraint qualification if the condition does not involve the objective function).

In [4] Giannessi states a special separation theorem, namely a disjunctive separation between a face F of a convex cone C and a set S by means of a hyperplane which does not contain the face; referring to the above problem, such a separation will be called regular (with respect to the face F). In this section, we will generalize to the nonconvex case the results established by Giannessi in [4].

Let us consider Theorem 2.2.7 of [4].

Theorem 3.1. *Let $C \subseteq \mathbb{R}^n$ be a nonempty and convex cone, with apex at $O \notin C$ such that $C + \text{cl } C = C$, and F be any face of C . Let $S \subseteq \mathbb{R}^n$ be nonempty with $O \in \text{cl } S$ and such that $S - \text{cl } C$ is convex. F is contained in every hyperplane which separates C and S , if any, if and only if*

$$F \subseteq TC(S - \text{cl } C),$$

where $TC(S - \text{cl } C)$ is the tangent cone to $S - \text{cl } C$ at O .

Theorem 3.1 assumes the convexity of $S - \text{cl } C$. The following example shows that if we remove such an assumption, then the necessity in the theorem does not hold.

Example 3.2. Let C be the following convex cone in \mathbb{R}^3 :

$$C = \{x \in \mathbb{R}^3 : x_1 > 0, x_2 = 0, x_3 = 0\}$$

and

$$S = \{x \in \mathbb{R}^3 : x_1 = x_2 \geq 0, x_3 = -x_1^2 - x_2^2\} \cup$$

$$\{x \in \mathbb{R}^3 : x_1 = -x_2 \geq 0, x_3 = -x_1^2 - x_2^2\}.$$

Choose $F = C$. Obviously S and $S - \text{cl } C$ are not convex. The plane $H^0 = \{x \in \mathbb{R}^3 : x_3 = 0\}$ is the unique plane which separates C and S and it contains the face F , nevertheless F is not contained in $TC(S - \text{cl } C)$.

In order to extend Theorem 3.1 to nonconvex case, we have to consider $TC(\text{conv}(S - \text{cl } C))$ instead of $TC(S - \text{cl } C)$. Before giving the result that extend Theorem 3.1, let us state a preliminary property by means of the following lemma.

Lemma 3.3. *Let $C \subseteq \mathbb{R}^n$ be a nonempty and convex cone with apex at O and S be a nonempty subset of \mathbb{R}^n with $O \in \text{cl}(S - \text{cl } C)$. The following statements are equivalent:*

- (i) *a hyperplane separates C and S ;*
- (ii) *the same hyperplane separates C and $TC(\text{conv}(S - \text{cl } C))$.*

Proof. (i) \Rightarrow (ii) Let H^0 be any hyperplane which separates C and S . By Lemma 2.2.1 of [4], we have that H^0 separates C and $S - \text{cl } C$ and hence, obviously also C and $\text{conv}(S - \text{cl } C)$; i.e., $C \subseteq H^+$ and $\text{conv}(S - \text{cl } C) \subseteq H^-$. Now we will prove that $\text{conv}(S - \text{cl } C) \subseteq H^-$ implies $TC(\text{conv}(S - \text{cl } C)) \subseteq H^-$. Let $t \in TC(\text{conv}(S - \text{cl } C))$; then there exist a sequence $\{x^n\}_{n \geq 1} \subseteq \text{conv}(S - \text{cl } C)$ with $\lim_{n \rightarrow +\infty} x^n = 0$ and a sequence $\{\alpha_n\}_{n \geq 1} \subset \mathbb{R}_+ \setminus \{0\}$ such that $\lim_{n \rightarrow +\infty} \alpha_n x^n = t$. Since $x^n \in \text{conv}(S - \text{cl } C)$, $\forall n \geq 0$, then $\langle a, x^n \rangle \leq 0$, and hence $\langle a, \alpha_n x^n \rangle \leq 0$, $\forall n \geq 0$. Letting $n \rightarrow +\infty$ we obtain $\langle a, t \rangle \leq 0$ and thus $TC(\text{conv}(S - \text{cl } C)) \subseteq H^-$.

(ii) \Rightarrow (i) This is an obvious consequence of the inclusions $S \subseteq S - \text{cl } C \subseteq \text{conv}(S - \text{cl } C) \subseteq TC(\text{conv}(S - \text{cl } C))$. ■

Now, we give the generalization of Theorem 3.1 to nonconvex case.

Theorem 3.4. *Let $C \subseteq \mathbb{R}^n$ be a nonempty and convex cone with apex at O and S be a nonempty subset of \mathbb{R}^n with $O \in \text{cl}(S - \text{cl } C)$. Let F be any face of C . The following statements are equivalent:*

- (i) *There exists at least one hyperplane which separates S and C and does not contain F ;*
- (ii) *$F \not\subseteq TC(\text{conv}(S - \text{cl } C))$.*

Proof. (i) \Rightarrow (ii) The hypotheses imply the existence of a hyperplane of equation $H^0 := \{x \in \mathbb{R}^n : \langle a, x \rangle = 0\}$, such that $\langle a, x \rangle \leq 0$, $\forall x \in S$ and $\langle a, x \rangle \geq 0$, $\forall x \in C$ and that there exists $\bar{f} \in F$ with $\langle a, \bar{f} \rangle > 0$.

Ab absurdo, suppose $F \subseteq TC(\text{conv}(S - \text{cl } C))$. From Lemma 3.3 we have that H^0 separates also $TC(\text{conv}(S - \text{cl } C))$ and C , i.e. $\langle a, x \rangle \leq 0$, $\forall x \in TC(\text{conv}(S - \text{cl } C))$. Thus also $\langle a, \bar{f} \rangle \leq 0$, $\forall \bar{f} \in F$, which contradicts the hypothesis.

(ii) \Rightarrow (i) From $F \not\subseteq TC(\text{conv}(S - \text{cl } C))$ it follows that $\exists f^0 \in F \setminus TC(\text{conv}(S -$

$\text{cl } C$). Since $TC(\text{conv}(S - \text{cl } C))$ is closed and convex, then there exists a hyperplane H^0 of equation $\langle a, x \rangle = b$ with $a \in \mathbb{R}^n \setminus \{O\}$ such that $\langle a, x \rangle \leq b < \langle a, f^0 \rangle, \forall x \in TC(\text{conv}(S - \text{cl } C))$. Because of $O \in TC(\text{conv}(S - \text{cl } C))$, we can set $b = 0$ and thus we have

$$(11) \quad \langle a, x \rangle \leq 0 < \langle a, f^0 \rangle, \forall x \in TC(\text{conv}(S - \text{cl } C)).$$

The inclusion $S - \text{cl } C \subseteq TC(\text{conv}(S - \text{cl } C))$ implies that $\langle a, x \rangle \leq 0, \forall x \in S - \text{cl } C$. Now we prove that $\langle a, x \rangle \geq 0, \forall x \in C$. Ab absurdo, suppose that $\exists k \in C$ such that $\langle a, k \rangle < 0$ and let $s \in S$. Then we have $s - \alpha k \in S - \text{cl } C, \forall \alpha \in \mathbb{R}_+$ so that $\lim_{\alpha \rightarrow +\infty} \langle a, s - \alpha k \rangle = +\infty$, which contradicts $\langle a, x \rangle \leq 0, \forall x \in S - \text{cl } C$. Therefore H^0 separates C and $S - \text{cl } C$; obviously, H^0 separates also C and S and it does not contain F because of (11). ■

We call the separation between S and C *regular with respect to the face F* , iff F is not contained in at least one separating hyperplane.

Notice that in Theorem 3.4 the tangent cone $TC(\text{conv}(S - \text{cl } C))$ can be replaced by $\text{cl cone conv}(S - \text{cl } C)$; in fact, if A is a convex set, then $TC(A) = \text{cl cone } A$. Moreover, observe that in Theorem 3.4 it is not possible to replace $TC(\text{conv}(S - \text{cl } C))$ by $\text{conv } TC(S - \text{cl } C)$; in such a case, without the convexity assumption, it may exist a hyperplane which separates C and $TC(S - \text{cl } C)$ but does not separate C and $S - \text{cl } C$. This situation is illustrated by the following example.

Example 3.5. Let C be the following convex cone in \mathbb{R}^3 :

$$C = \{x \in \mathbb{R}^3 : x_1 > 0, x_2 = 0, x_3 = 0\} \text{ and}$$

$$S = \{x \in \mathbb{R}^3 : x_1 = x_2 \geq 0, x_3 \leq 0, x_3 = (x_1 - 1)^2 + (x_2 - 1)^2 - 2\} \cup \\ \{x \in \mathbb{R}^3 : x_1 = -x_2 \geq 0, x_3 \leq 0, x_3 = (x_1 - 1)^2 + (x_2 + 1)^2 - 2\}.$$

Choose $F = C$. Obviously S and $S - \text{cl } C$ are not convex. The plane $H^0 = \{x \in \mathbb{R}^3 : x_3 = 0\}$ is the unique plane which separates C and S and it contains the face F . It turns out:

$$TC(S - \text{cl } C) = \{x \in \mathbb{R}^3 : x_1 = x_2, x_3 \leq 0, x_3 \leq -4x_1\} \cup \\ \{x \in \mathbb{R}^3 : x_1 = -x_2, x_3 \leq 0, x_3 \leq -4x_1\}.$$

$TC(S - \text{cl } C)$ is not convex and we have that $F \not\subseteq \text{conv } TC(S - \text{cl } C)$. Moreover, every plane $H_a = \{x \in \mathbb{R}^3 : ax_1 + x_3 = 0\}$, with $0 < a \leq 4$, separates C and $TC(S - \text{cl } C)$ (and hence also C and $\text{conv } TC(S - \text{cl } C)$), but does not separate C and S and does not contain the face F .

Both in Example 3.2 and 3.5 we have $\text{int } C = \emptyset$. Similar examples with $\text{int } C \neq \emptyset$ can be given by putting $C = \{x \in \mathbb{R}^3 : x_1 \geq 0, -10x_1 \leq x_2 \leq 0, 0 \leq x_3 \leq 10x_1\}$ and choosing $F = \{x \in \mathbb{R}^3 : -10x_1 \leq x_2 \leq 0, x_3 = 0\} \subset C$.

4. CONCLUDING REMARKS

We have considered a generalized format of a constrained extremum problem and we have stressed the fact that, for the investigation of optimality conditions, it is of fundamental importance the study of the separation between two suitable sets. In particular, we have given a condition equivalent to the linear separation between a convex cone C and a generic set S . Moreover, we have proposed a regularity condition for the linear separation between C and S ; this condition includes also the nonconvex case and it is finalized to the application to constrained optimization. In fact, given a constrained optimization problem, the application of the separation results obtained in this work to the convex cone \mathcal{H} and the image set $\mathcal{K}_{\bar{x}}$ allows to achieve the existence of regular saddle points; i.e., a regularity condition that plays the role of a sufficient optimality condition [6]. On the other hand, the regular separation can be applied to a constrained extremum problem also by replacing the image of the problem with its homogenization. In this alternative approach, it is possible to prove that it is equivalent to the existence of Lagrangian multipliers with a positive multiplier associated to the objective function; i.e., to a necessary optimality condition [7].

REFERENCES

1. V. Barbu and Th. Precupanu, *Convexity and Optimization in Banach Spaces*, D. Reidel Publ. Co., Dordrecht, 1986.
2. J. F. Bonnans and A. Shapiro, *Perturbation Analysis of Optimization Problems*, Springer, New York, 2000.
3. F. Giannessi, Semidifferentiable Functions and Necessary Optimality Conditions, *J. Optim. Theory and Appl.*, **60** (1989), 191-214.
4. F. Giannessi, *Constrained Optimization and Image Space Analysis*, Vol. 1, Separation of Sets and Optimality Conditions, Springer, 2005.
5. V. Jeyakumar and N. D. Yen, Solution Stability of Nonsmooth Continuous Systems with Applications to Cone-Constrained Optimization, *SIAM J. Optim.*, **14** (2004), 1106-1127.
6. A. Moldovan and L. Pellegrini, On Regularity for Constrained Extremum Problems, Part 1: Sufficient Optimality Conditions, *J. Optim. Theory and Appl.*, **142** (2009), 147-163.

7. A. Moldovan and L. Pellegrini, On Regularity for Constrained Extremum Problems, Part 2: Necessary Optimality Conditions, *J. Optim. Theory and Appl.*, **142** (2009), 165-183.
8. P. H. Quang and N. D. Yen, New Proof of a Theorem of F. Giannessi, *J. Optim. Theory and Appl.*, **68** (1991), 385-387.
9. T. R. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Springer-Verlag, Berlin, 1998.

Letizia Pellegrini
Department of Economics
University of Verona
Viale dell'Università 3
Palazzina 32
37129 Verona
Italy
E-mail: letizia.pellegrini@univr.it