REPORT n.4

Non-linear Separation Theorems, Duality and Optimality Conditions

Alberto Cambini

Pisa, 1987
This research is partially supported by the Ministery of Public Education.

Non-Linear Separation Theorems, Duality and Optimality Conditions

A. Cambini

1. INTRODUCTION

Recently [15, 16] a theorem of the alternative has been stated for generalized systems and it has been shown how to deduce, from such a theorem, known optimality conditions like saddle-point one, regularity conditions, known theorems of the alternative and new ones; furthermore connections among optimality conditions, duality, penalty functions have been shown. Some of these new ideas have been deepened by other people and new interesting results have been obtained [6, 22, 26, 27, 28, 29]. The aim of this paper is to deepen this unifying approach in such a way to give a survey of these recent studies and, at the same time, to obtain some new results.

2. THEOREMS OF THE ALTERNATIVE AND SEPARATION FUNCTIONS

Assume that we are given the positive integers n and ν, the non-empty sets $H \subseteq \mathbb{R}^{\nu}$, $X \subseteq \mathbb{R}^{n}$, and the real-valued function $F: X \to \mathbb{R}^{\nu}$. We want to study conditions for the generalized system

(1) $$F(x) \in H \ , \quad x \in X$$

to have (or not to have) solutions.

To this aim we introduce the following

Definition 1.1: $w: \mathbb{R}^\nu \to \mathbb{R}$ is called *weak separation function* iff(1)

(2a) $$\text{lev}_{>0}\, w \supseteq \mathcal{H}$$

$s: \mathbb{R}^\nu \to \mathbb{R}$ is called *strong separation function* iff

(2b) $$\text{lev}_{>0}\, s \subseteq \mathcal{H}$$

The following theorem holds:

Theorem 1.1

Let the sets $\mathcal{H}$, X and the function F be given.

i) The systems (1) and (3a)

(3a) $$w(F(x)) \leq 0 \quad , \quad \forall x \in X$$

are not simultaneously possible, whatever the weak separation function w might be.

ii) The systems (1) and (3b)

(3b) $$s(F(x)) \leq 0 \quad , \quad \forall x \in X$$

are not simultaneously impossible, whatever the strong separation function s might be.

(1) If ρ is a real-valued function $\rho: \mathbb{R}^n \to \mathbb{R}$, we set $\text{lev}_{>0}\, \rho = \{y \in \mathbb{R}^n: \rho(y) > 0\}$. Sets $\text{lev}_{\geq 0}\rho$, $\text{lev}_{<0}\rho$, $\text{lev}_{\leq 0}\rho$ are defined in a similar way.

Proof: i) If (1) is possible, there exists $\bar{x} \in X$ such that $F(\bar{x}) \in H$; from (2a) we have $w(F(\bar{x})) \geq 0$ so that (3a) is false. ii) If (1) is impossible, i.e., if $F(x) \notin H$, $\forall x \in X$ then (2b) implies $s(F(x)) \leq 0$ $\forall x \in X$ so that (3b) is true. This completes the proof. ⋕

Let us note that the impossibility of (1) is not equivalent to the possibility of (3a) (or(3b)). For this reason we say that *weak* (strong) *alternative* holds between (1) and (3a)(or(3b)).

When the possibility of (3b) implies the impossibility of (1) we say that *alternative* holds between (1) and (3b). Set $K = F(X)$; it is obvious that system (1) is impossible iff $K \cap H = \emptyset$. Thus, if we refer to $\mathbb{R}^{\nu}$ as the *image space*, the impossibility of (1) is equivalent to the disjunction between K and H in the image space.

When $K \subseteq \mathrm{lev}_{\leq 0} w$, that is (3a) is possible so that (1) is impossible, we say that *alternative* holds between (1) and (3a) or, equivalently, that w is a weak separation function which *guarantees alternative*.

In this order of ideas two important questions consist in choosing, for any given triplet (F,X,H), an appropriate class of weak separation functions and in finding, in this class, an element, if one exists, which guarantees alternative. These questions will be analyzed for a wide class of systems.

From now on we suppose that H is a convex cone, satisfying the following property[2]

(4) $$H + \mathrm{cl}\, H = H$$

For instance, (4) is verified in the following cases: a) H is open

(2) cl A denotes the closure of set A.

or H is closed; b) cl H is pointed, i.e. $(\mathrm{cl}\ H) \cap (-\mathrm{cl}\ H) = \{0\}$ and[3] $H = (\mathrm{cl}\ H) \setminus \{0\}$; c) H is the Cartesian product of two convex cones H_1, H_2 with H_1 and/or H_2 satisfying a) or b).

Consider now the *conic extension* of K with respect to cl H, i.e., the set $E \triangleq K - \mathrm{cl}\ H$.

The following lemma holds:

Lemma 1.1 <u>System</u> (1) <u>is impossible iff</u> $E \cap H \neq \emptyset$.

Proof: Sufficiency: $H \cap E = \emptyset \Rightarrow K \cap H = \emptyset \Rightarrow$ (1) impossible. *Necessity:* assume that (1) is impossible and that $E \cap H \neq \emptyset$. Then there exist $\bar{x} \in X$, $h \in \mathrm{cl}\ H$ such that $F(\bar{x}) - h \in H$ that is $F(\bar{x}) \in (\mathrm{cl} H) + H = H$ and this is absurd. ⧣

The reason of introducing E is based on the fact that E can have some properties which are not valid for K; for instance E turns out to be convex even if K is not. E is a key set in further analysis and we will see that its properties are very important in order to study optimality conditions, duality, regularity and so on.

As regards to the convexity of E, the following lemma holds:

Lemma 1.2 <u>Let</u> X <u>be convex</u>. <u>The following conditions are equivalent:</u>

i) F <u>is</u> (cl H)-<u>convex-like</u>[4]

ii) $E = F(X) - \mathrm{cl}\ H$ <u>is convex</u>.

(3) $A \setminus B \doteq \{a \in A : a \notin B\}$.

(4) Let $C \subset \mathbb{R}^s$ be a convex cone, and $X \subseteq \mathbb{R}^n$ a convex set. A function $F: X \to \mathbb{R}^s$ is C-convexlike iff $\forall x, y \in X$, there exist $z \in X$ such that $F(z) - (1-\alpha)F(x) - \alpha F(y) \in C$, $\forall \alpha \in [0,1]$.

Proof: We have

i) $\Leftrightarrow \forall x,y \in X,\ \forall \alpha \in [0,1],\ (1-\alpha)F(x)+\alpha F(y) \in F(X)-\text{cl}\, H$

$\Leftrightarrow \forall x,y \in X,\ \forall p_1,p_2 \in \text{cl}\, H,\ \forall \alpha \in [0,1]$

$(1-\alpha)(F(x)-p_1)+\alpha(F(y)-p_2) \in (F(X)-\text{cl}\, H \Leftrightarrow$ ii). #

Lemma 1.1 shows that the impossibility of (1) is equivalent to the disjunction between E and H in the image space; when E is convex it is natural to study this disjunction by means of the class of linear functionals $W = \{w: \mathbb{R}^\nu \to \mathbb{R},\ w \in H^*\}$ [5].

The following theorem of the alternative holds:

Theorem 1.2

Consider system (1)

(1) $$F(x) \in H, \quad x \in X$$

and suppose that F is (cl H)-convexlike. Then i) and ii) hold:

i) if (1) is impossible then

(5) $$\exists \bar{w} \in H^* : \bar{w}(F(x)) \le 0 \quad , \quad \forall x \in X$$

ii) if (5) holds and moreover

$$\{x \in X: \bar{w}(F(x)) = 0\} = \emptyset \quad \text{when} \quad \bar{w} \notin \text{int}\, H^*,$$

then system (1) is impossible.

[5] A^* denotes the *polar* of $A \subset \mathbb{R}^s$, i.e. the set $A^*=\{\lambda \in \mathbb{R}^s : \langle \lambda, y\rangle \ge 0$ $\forall y \in A\}$ where $\langle\ ,\ \rangle$, denotes the scalar product.

Proof: i) According to Lemma 1.2 and Lemma 1.1, E is convex and $E \cap H = \emptyset$. Since $\mathrm{ri}\, E \cap \mathrm{ri}(\mathrm{cl}\, H) = \emptyset$[6] there exists a hyperplane which separates E and $\mathrm{cl}\, H$ properly, that is there exists a linear functional $\bar{w}$ such that $\mathrm{lev}_{\geq 0} \bar{w} \supseteq \mathrm{cl}\, H$, $\mathrm{lev}_{\leq 0} \bar{w} \supseteq E$. The first inclusion implies $\bar{w} \in H^*$ and the second the inequalities $\bar{w}(F(x)) \leq 0 \quad \forall x \in X$. ii) If $\bar{w} \in \mathrm{int}\, H^*$[6] then $\bar{w}(h) > 0 \quad \forall h \in H$ so that $\bar{w}$ is a weak separation function and the thesis follows by Theorem 1.1. Suppose now $\bar{w} \notin \mathrm{int}\, H^*$ and that there exists $\bar{x} \in X$ with $F(\bar{x}) \in H$ so that $\bar{w}(F(\bar{x})) \geq 0$; then by (5), $\bar{w}(F(\bar{x}))=0$, and this is absurd. #

Consider now the following important particular case of (1):

$$v = \ell + m \quad , \quad H = (\mathrm{int}\, U) \times V$$

(6)

$$f: X \to \mathbb{R}^\ell \quad , \quad g: X \to \mathbb{R}^m \quad , \quad F(x)=(f(x),g(x)$$

where the positive integers ℓ and m, the closed convex cones $U \subset \mathbb{R}^\ell, V \subset \mathbb{R}^m$, with $\mathrm{int}\, U \neq \emptyset$ (otherwise $H = \emptyset$), and the functions f, g are given.

The generalized system (1) becomes

(7) $$f(x) \in \mathrm{int}\, U \quad , \quad g(x) \in V \quad , \quad x \in X$$

Let W_1 the class of functions

$$W_1=\{w:\mathbb{R}^\ell \times \mathbb{R}^m \to \mathbb{R} : w(u,v,\theta,\lambda)=<\theta,u>+<\lambda,v>, \theta \in U^*, \lambda \in V^*\}.$$

It is easy to show that $w \in W_1$ is a weak separation function when

[6] ri A, int A denote, respectively, the relative interior and the interior of A.

$\theta \in U^{*}\setminus\{0\}$. We may show that it is possible to choose $w \in W_1$ in such a way that it guarantees alternative for a wide class of systems (7).

As a consequence of Theorem 1.2 we have the following Corollaries.

Corollary 1.1 Let $F(x) = (f(x),g(x)$ be cl H-convexlike. Then i) and ii) hold:

i) if (7) is impossible then:

(8a) $\exists\bar{\theta} \in U^{*}$, $\exists\bar{\lambda} \in V^{*}$ with $(\bar{\theta},\bar{\lambda}) \neq 0$ such that

(8b) $<\bar{\theta},f(x)> + <\bar{\lambda},g(x)> \leq 0$, $\forall x \in X$

ii) if (8) holds and moreover

$$\{x \in X\colon f(x) \in \text{int } U,\ g(x) \in V\ ,\ <\bar{\lambda},g(x)>=0\} = \emptyset$$

when $\bar{\theta} = 0$, then system (7) is impossible.

Proof: Similar to the one given in Theorem 1.2. #

Corollary 1.2 Let X be convex, f be a U-function(7) and g be a V-function. Then i) and ii) hold:

(7) Let C be a convex cone. F is said to be a C-function on a convex set X, iff $\forall x,\ y \in X$

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

Note that a $(\mathbb{R}^n_+)$-function is a concave function and a $(\mathbb{R}^n_-)$-function is a convex function.

i) <u>if</u> (7) <u>is impossible then</u>:

(9a) $\exists\bar{\theta} \in U^*$, $\exists\bar{\lambda} \in V^*$, with $(\bar{\theta},\bar{\lambda}) \neq 0$ such that

(9b) $$<\bar{\theta},f(x)> + <\bar{\lambda},g(x)> \leq 0 \quad , \quad \forall x \in X$$

ii) <u>suppose that</u> (9) <u>holds and moreover</u>

$$\{x \in X: f(x) \in \text{int } U \ , \ g(x) \in V \ , \ <\bar{\lambda},g(x)>=0\}=\emptyset$$

<u>when</u> $\bar{\theta} = 0$. <u>Then system</u> (7) <u>is impossible</u>.

Proof: It is sufficient to note that $F(x) = (f(x),g(x))$ is (cl H)-convexlike. #

Some sufficient conditions for the convexlikeness of pairs of two functions can be found in Ref.12.

As outlined in Ref. 16, when $U = \mathbb{R}^{\ell}_{+}$, Corollary 1.2 becomes Theorem 1 of Ref. 3; if, in addition f and g are concave in the ordinary sense and $V = \mathbb{R}^m_+$, then Corollary 1.2 becomes Theorem 3 of Ref. 15.

Further instances of how theorems of the alternative can be derived from Corollary 1.2 are found in Ref. 15, 16.

3. WEAK ALTERNATIVE AND OPTIMALITY CONDITIONS

In this section we will see how weak alternative can be used to study optimality conditions.

Consider the following extremum problem

(10) $$P: \min \phi(x) \quad , \quad x \in R \triangleq \{x \in X: g(x) \geq 0\}$$

where $X \subseteq \mathbb{R}^n$, $\phi: X \to \mathbb{R}$, $g: X \to \mathbb{R}^m$.

A feasible solution $\bar{x} \in R$ is optimal for problem (10) iff the system

(11) $$f(x) \overset{\Delta}{=} \phi(\bar{x}) - \phi(x) > 0 \quad , \ g(x) \geq 0 \quad , \ x \in X$$

has not solutions.

Note that (11) is a particular case of (7) and (1).

Taking into account that for i) of Theorem 1.1, systems (1) and (3a) cannot be both possible, any assumption which ensures that an element of a given class of weak separation functions guarantees alternative, becomes a sufficient condition for $\bar{x}$ to be optimal.

In this way we can obtain some optimality conditions such as a generalized saddle-point condition.

With this aim consider the set of functions

(12) $$W_2 = \{w : \mathbb{R} \times \mathbb{R}^m \to \mathbb{R} \ , \ w(u,v,\theta,\omega) = \theta u + \gamma(v,\omega) \ , \ \theta \geq 0 \ , \quad \omega \in \Omega\}$$

where $\gamma : \mathbb{R}^m \to \mathbb{R}$ and Ω is the domain of parameter ω such that:

(13a) $$\forall \theta \geq 0 \ , \quad \forall \omega \in \Omega \ , \ \mathrm{lev}_{\geq 0} w \supseteq \mathrm{cl}\ H$$

(13b) $$\bigcap_{\theta \geq 0, \omega \in \Omega} \mathrm{lev}_{\geq 0}\, w = \mathrm{cl}\ H$$

(13c) $$w \in W_2 \ , \ k > 0 \quad \underline{\text{implies}} \ kw \in W_2$$

where $\mathrm{cl}\ H = \mathbb{R}_+ \times \mathbb{R}^m_+$.

It is easy to verify that the class of linear functionals

(14) $$w(u,v,\theta,\omega) = \theta u + \langle \omega , v \rangle \ , \quad \theta \in \mathbb{R}_+ \ , \ \omega \in \mathbb{R}^m_+$$

satisfies properties (13a, b , c).

The following Lemma shows some connections between the properties of w and γ.

Lemma 2.1 Consider the set of functions W_2. Then (13) is equivalent to (15):

(15a) $$\forall\, \omega \in \Omega\ ,\ \mathrm{lev}_{\geq 0}\, \gamma \supseteq \mathbb{R}^m_+$$

(15b) $$\bigcap_{\omega \in \Omega} \mathrm{lev}_{\geq 0}\, \gamma = \mathbb{R}^m_+$$

(15c) $$\forall \bar{\omega} \in \Omega,\ \forall k > 0\ ,\ \exists \hat{\omega} \in \Omega : \gamma(\cdot, \hat{\omega}) = k\gamma(\cdot, \bar{\omega})$$

Proof: (13a) $\Leftrightarrow$ (15a) .

(13a) $\Rightarrow w(u,v,\theta,\omega) \geq 0 \quad \forall u,\ \theta \geq 0\ , \quad \forall v, \omega \in \mathbb{R}^m_+$

$\Rightarrow w(0,v,\theta,\omega) = \gamma(v,\omega) \geq 0 \quad \forall v,\ \omega \in \mathbb{R}^m_+ \ \Rightarrow$ (15a) .

(15a) $\Rightarrow \gamma(v,\omega) \geq 0 \quad \forall v, \omega \in \mathbb{R}^m_+ \Rightarrow \theta u + \gamma(v,\omega) \geq 0 \quad \forall \theta, u \geq 0$

$\forall v,\ \omega \in \mathbb{R}^m_+ \ \Rightarrow$ (13a).

(13b) $\Leftrightarrow$ (15b). Suppose that (13b) holds and (15b) is false. Then $\exists \bar{v} \notin \mathbb{R}^m_+$ such that $\gamma(\bar{v},\omega) \geq 0 \quad \forall \omega \in \Omega$ and, consequently, $w(0,\bar{v},\theta,\omega) = \gamma(\bar{v},\bar{\omega}) \geq 0$ $\forall \theta \geq 0,\ \forall \omega \in \Omega$ so that $(0,\bar{v}) \in \mathrm{cl}\, H$. Thus $\bar{v} \in \mathbb{R}^m_+$ and this is absurd. Suppose now that (15b) holds and (13b) is false. Then $\exists (\bar{u},\bar{v}) \notin \mathrm{cl}\, H$ such that $\theta\bar{u} + \gamma(\bar{v},\omega) \geq 0 \quad \forall \theta \geq 0,\ \forall \omega \in \Omega$ and this relation implies $\bar{u} > 0$, otherwise $\theta\bar{u} \to -\infty$ when $\theta \to +\infty$. On the other hand $w(\bar{u},\bar{v},0,\omega) = \gamma(\bar{v},\omega) \geq 0 \ \forall \omega \in \Omega$ so that $\bar{v} \in \mathbb{R}^m_+$ for (15b), and this is absurd.

(13c) $\Leftrightarrow$ (15c). $\forall \bar{\omega} \in \Omega\ ,\ \forall k > 0$ consider the function $w(u,v,\theta,\bar{\omega}) = (\theta/k)u + \gamma(v,\bar{\omega})$. Then $kw = \theta u + k\gamma(v,\bar{\omega}) \in W_2$ iff (15c) holds. #

The following Lemma gives conditions under which $w \in W_2$ guarantees weak alternative between (1) and (3a) where, now, system (1) is of the form (11).

Lemma 2.2 i) When $\theta > 0$ or $\theta = 0$ and $\text{lev}_{\geq 0}\gamma \supseteq \mathbb{R}^m_+$, the function $w \in W_2$ guarantees weak alternative between (1) and (3a), with $Z =]-\infty, 0]$.
ii) When $\theta = 0$ and $\text{lev}_{>0}\gamma \not\supseteq \mathbb{R}^m_+$, w guarantees weak alternative between (1) and (3a), with $Z =]-\infty, 0[$.

Proof: (i) In the present case, namely (6) with $\ell = 1$, (2a) becomes $\text{lev}_{>0}w \supseteq]0,+\infty[\times \mathbb{R}^m_+$, or

$$(u,v) \in]0,+\infty[\times \mathbb{R}^m_+ \Rightarrow \theta u + \gamma(v;\omega) > 0 .$$

This relationship holds since now we have either $\theta > 0$ and $\text{lev}_{\geq 0}\gamma \supseteq \mathbb{R}^m_+$, or $\theta = 0$ and $\text{lev}_{>0}\gamma \supseteq \mathbb{R}^m_+$. Thus, the thesis follows from (i) of theorem 1.1.

(ii) In the present case, namely (4.2)-(5.3), (2a) becomes:

$$\text{lev}_{\geq 0}w \supseteq]0,+\infty[\times \mathbb{R}^m_+ ,$$

or:

$$(u,v) \in]0,+\infty[\times \mathbb{R}^m_+ \Rightarrow \theta u + \gamma(v;\omega) \geq 0 .$$

Since $\theta = 0$, this relationship is an obvious consequence of (15a). Again weak alternative follows from (i) of theorem 1.1. This completes the proof.

Taking into account Lemma 2.2, it is immediate to interprete (i) of Theorem 1.1 as a sufficient optimality condition for problem (10); this is contained in the following:

Corollary 2.1 If $\bar{x} \in \mathbb{R}^n$ fulfils conditions: (i) $\bar{x} \in R$; (ii) there exist $\bar{\theta} \in \mathbb{R}_+$ and $\bar{\omega} \in \Omega$, such that

$$(16) \qquad \bar{\theta}\,[\phi(\bar{x}) - \phi(x)] + \gamma(g(x);\bar{\omega}) \leq 0 , \quad \forall x \in X ,$$

<u>and moreover</u>

$$\{x \in X:\ \phi(x) < \phi(\bar{x}), g(x) \geq 0\ ,\ \gamma(g(x),\bar{\omega})=0\}=\emptyset$$

<u>when</u> $\theta=0$ <u>and</u> $\mathrm{lev}_{>0}\gamma \not\supseteq \mathbb{R}^m_+$; <u>then</u> $\bar{x}$ <u>is a global minimum point of</u> (10).

Now, introduce the function

$$L(x;\theta,\omega) \triangleq \theta\phi(x) - \gamma(g(x);\omega)\ ,$$

and let us prove the following:

Theorem 2.1

<u>Condition</u> (i)-(ii) <u>of Corollary</u> 2.1 <u>is equivalent to the other one</u> : <u>there exist</u> $\bar{x} \in X$, $\bar{\theta} \in \mathbb{R}_+$ <u>and</u> $\bar{\omega} \in \Omega$, <u>such that</u>

$$(17)\qquad L(\bar{x};\bar{\theta},\omega) \leq L(\bar{x};\bar{\theta},\bar{\omega}) \leq L(x;\bar{\theta},\bar{\omega})\ ,\quad \forall x \in X\ ,\ \forall \omega \in \Omega,$$

<u>and moreover</u>

$$\{x \in X:\ \phi(x) < \phi(\bar{x}), g(x) \geq 0\ , \gamma(g(x),\bar{\omega}) = 0\}=\emptyset$$

<u>if</u> $\theta = 0$ <u>and</u> $\mathrm{lev}_{>0}\gamma \not\supseteq \mathbb{R}^m_+$.

Proof: Let us prove that (i)-(ii) of Corollary 2.1 $\Rightarrow$(17).$\bar{\omega}\in\Omega$ and $\bar{x} \in R$ imply that $\gamma(g(\bar{x});\bar{\omega})\geq 0$ (since $\mathrm{lev}_{\geq 0}\gamma \supseteq \mathbb{R}^m_+$); at $x = \bar{x}$ (16) implies $\gamma(g(\bar{x});\bar{\omega}) \leq 0$; it follows $\gamma(g(\bar{x});\bar{\omega}) = 0$. Hence, (16) is equivalent to the 2-nd of (17).

Let us prove, now, that (17) implies i)-ii) of Corollary 2.1. The 1-st of (17) implies

$$(18)\qquad \gamma(g(\bar{x}),\bar{\omega}) \leq \gamma(g(\bar{x}),\omega)\ ,\ \forall \omega \in \Omega\ .$$

Suppose that $g(\bar{x}) \not\geq 0$. Then, by (15b) there exists $\hat{\omega}\in\Omega$ such that $\gamma(g(x),\bar{\omega})<0$;

hence by (15c)

$$\gamma(g(\bar{x}),\bar{\omega}) \le K(g(\bar{x}),\hat{\omega}) \quad \forall K > 0 \quad \text{and this is absurd.}$$

Thus $g(\bar{x}) \ge 0$ and i) of Corollary 2.1 is proven. Because of (15b) we have $\gamma(g(\bar{x}),\bar{\omega}) \ge 0$. Suppose that $\gamma(g(\bar{x}),\bar{\omega}) > 0$; by (15c) and (18) we have $\gamma(g(\bar{x}),\bar{\omega}) \le \frac{1}{2}\, \gamma(g(\bar{x}),\bar{\omega})$, so that $\gamma(g(\bar{x}),\bar{\omega}) = 0$. Account taken of this equality, it is easy to show that the second part of (17) implies (16). This completes the proof. #

Now, note that (17) can be regarded as a generalized saddle-point condition and L as a generalized Lagrangean function. When $\gamma(v,\omega) = \langle \omega, v \rangle$, (17) becomes the well-known John saddle-point condition and L the classic Lagrangean function.

Note that, when (16) or (17) holds, from the proof of Theorem 2.1 we have that $(\bar{x},\bar{\omega})$ fulfils the generalized complementarity condition $\gamma(g(x),\omega) = 0$ which collapses to the well-known ordinary one, when γ is linear and $\Omega = \mathbb{R}^m_+$.

Now, suppose that $\bar{x}$ is an optimal solution for problem (10), so that system (1) in the form of (11) has not solutions. The impossibility of (1) does not imply the possibility of (3a). In this sense, any condition which guarantees, in a given, suitable class of weak separation functions, the existence of an element w satisfying (3a) becomes a necessary optimality condition (besides the weak alternative).

When E is convex, choosing the class of linear functionals, for (i) of Corollary 1.1, (16) and (17) hold.

Consequently, when X is convex, $-\phi$, g are concave, (16) and (17) become necessary otimality conditions too.

When E is not convex the above mentioned class is not useful since a linear functional does not guarantee, in general, alternative.

A general approach is to consider a transformation of the constraints $T = (T_1,\dots,T_m)$ such that

$$T_i(v_i,\mu_i) \gtreqless 0 \quad \text{according to} \quad v_i \gtreqless 0 \ , \ \forall \mu_i \geq 0$$

and to choose, properly, w in the class of weak separation functions W_3, where

$$W_3 = \{w : \mathbb{R} \times \mathbb{R}^m \to \mathbb{R} \ , \ w(u,v,\lambda,\mu) = u + \sum_{i=1}^{m} \lambda_i T_i(v_i,\mu_i) \quad .$$

Note that $T(g(x),\mu)$ may have a certain property, for instance concavity, differentiability, which does not hold for g.

An appropriate choice of W_3 may be the exponential one, that is the case where $T_i(v_i,\mu_i) = v_i \exp(-\mu_1 v_i)$ $i = 1,\ldots,m$. By means of this transformation it has been possible (Ref. 27,28) to characterize a wide class of problems for which there exists $w \in W_3$ which guarantees alternative, i.e. such that (19) holds:

(19a) $$\exists \bar{\lambda},\bar{\mu} \geq 0, \ (\bar{\lambda},\bar{\mu}) \neq 0 \quad \text{such that}$$

(19b) $$\phi(\bar{x}) - \phi(x) + \sum_{i=1}^{m} \bar{\lambda}_i g_i(x) \exp(-\bar{\mu}_i g_i(x)) \leq 0 \ , \ \forall x \in X$$

Let us note that (16), (17), (19) are global optimality conditions since X is not specified. Some other results can be obtained by means of a further analysis of the weak alternative, based on local arguments (Ref. 16,27); in particular it is possible to deduce, in a very simple way, the classic necessary conditions of the Lagrange, Karush, Kuhn-Tucker type. With this end, consider a class of weak separation functions $W = \{w(u,v,\omega), \ \omega \in \Omega\}$, satisfying the following properties:

(20a) $$\bigcap_{\omega \in \Omega} \text{lev}_{\geq 0} w(u,v,\omega) = \text{cl } H$$

(20b) $$\bigcap_{\omega \in \Omega} \text{lev}_{> 0} w(u,v,\omega) = \text{int } H$$

It is easy to prove that the class of the linear functionals

$$w(u,v,\lambda,\mu) = \lambda u + \langle \mu, v \rangle \ , \quad \lambda \in \mathbb{R}^{+} \ , \quad \mu \in \mathbb{R}^{m}_{+}$$

verifies (20).

The following Lemma plays a fundamental role in the derivation of F. John conditions, and for necessary conditions of isoperimetric kind in the calculus of variations (see [17]). To this end denote by E_G the conic extension (in the sense of sect.1) of the cone generated by K at the origin 0 (i.e. the union of rays having 0 as common origin and non-empty intersection with K), and by E_G^* the (non-negative polar) of E_G.

Lemma 2.3 Let W be a class of weak separation functions satisfying property (20), and assume that the condition:

$$E_G^* \neq \{0\}$$

is fulfilled. If $\bar{x}$ is a local minimum point for problem (10), then there exists $\bar{\omega} \in \Omega$ such that

$$\limsup_{x \to \bar{x}} \frac{w(f(x),g(x),\bar{\omega}) - w(f(\bar{x}),g(\bar{x}),\bar{\omega})}{\|x - \bar{x}\|} \leq 0 \tag{21}$$

Remark If $w(u,v,\omega) = 0$ when $\omega = 0$, Lemma 2.3 holds with $\bar{\omega} \neq 0$.

Now we are able to prove the following classic theorem:

Theorem 2.2

Let $\bar{x}$ be a local optimal solution for problem (10) and set $I = \{i : g_i(\bar{x}) = 0\}$. Suppose that X is an open neighbourhood of $\bar{x}$, ϕ_i, g_j $i=1,\dots,\ell$ $j \in I$ are differentiable at $\bar{x}$ and g_j, if $j \notin I$ is continuous at $\bar{x}$. Then there exist

(23a) $\bar{\lambda} \geq 0$, $\bar{\mu} \geq 0$, $(\bar{\lambda}, \bar{\mu}) \neq 0$ such that

(23b) $\bar{\theta} \nabla \phi(\bar{x}) - \sum_{i=1}^{m} \bar{\mu}_i \nabla g_i(\bar{x}) = 0$

(23c) $$\bar{\mu}_i \cdot g_i(\bar{x}) = 0 \quad i=1,\ldots,m$$

Proof: The continuity of $g_j(x)$ at $\bar{x}$, $j \notin I$ implies that $\bar{x}$ is a local optimal solution for problem P'

$$P': \min \phi(x) \; , \; g(x) \geq 0 \; , \; x \in U(\bar{x})$$

where $U(\bar{x})$ is a suitable neighbourhood of $\bar{x}$ such that $g_j(x)>0$, $\forall x \in U(\bar{x})$. Consider the class of functions

$$w(u,v,\lambda,\mu) = \lambda u + <\mu,v> \; , \; \lambda \geq 0, \; \mu \in \mathbb{R}^m_+$$

and the Lagrangean function associated with P'

(24) $$L(x,\lambda,\mu) = \lambda\phi(x) - <\mu,g(x)>.$$

Taking into account that $\inf\limits_{x \to \bar{x}} \sup \frac{\psi(x)-\psi(\bar{x})}{\|x - \bar{x}\|} = |\nabla\psi(\bar{x})|$ when ψ is differentiable at $\bar{x}$, for Lemma 2.3 and (24), there exist $\bar{\lambda} \geq 0$, $\bar{\mu} \geq 0$ such that $|\nabla L(\bar{x},\bar{\lambda},\bar{\mu})| \leq 0$, i.e., $\nabla L(\bar{x},\bar{\lambda},\bar{\mu}) = 0$. Setting $\bar{\lambda}_i = 0$, $i \notin I$, the proof of the theorem is complete. #

4. STRONG ALTERNATIVE AND OPTIMALITY CONDITIONS

Let $s: \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}$, $s(u,v) = u - \delta(v)$ be a strong separation function with $\delta: \mathbb{R}^m \to \mathbb{R}$. It is easy to show that the condition $\text{lev}_{>0} s \subseteq \mathcal{H}$ implies $\delta(v) \geq 0 \quad \forall v \geq 0$ and $\delta(v) = +\infty \quad \forall v \ngeq 0$.

In order to avoid this kind of restrictions we will give a more general definition of a strong separation function than the one given in section 1.

Consider system (1) and let $\bar{K} \subseteq \mathbb{R}^\nu$ be such that $K \subseteq \bar{K}$, namely

(25) $$F(x) \in \bar{K} \quad , \quad \forall x \in X \quad .$$

$\bar{K}$ trivially exists since (25) is satisfied by at least $\bar{K} = \mathbb{R}^\nu$.

We say that $s: \mathbb{R}^\nu \to \mathbb{R}$ is a *strong separation function* iff we have:

(26) $$\operatorname{lev}_{>0} s \cap \bar{K} \subseteq H$$

Let us note that (26) reduces to (2b) when $\bar{K} = \mathbb{R}^\nu$.

The following Lemma generalizes ii) of Theorem 1.1.

Lemma 3.1 Let s be a strong separation function. The systems (1) and the following one:

(27) $$s(F(x)) \le 0 \quad , \quad \forall x \in X$$

cannot be both impossible.

Proof: Suppose that (27) is impossible i.e. there exists $\bar{x} \in X$ such that $s(F(\bar{x})) > 0$; from (27) it results $F(\bar{x}) \in \operatorname{lev}_{>0} s \cap \bar{K} \subseteq H$ and thus (1) is possible. #

Consider now problem (10) and let $\bar{K} \subseteq \mathbb{R} \times \mathbb{R}^m$ be such that

(28) $$(\phi(\bar{x})-\phi(x), g(x)) \in \bar{K} \quad , \quad \forall x \in X$$

For instance we can set $\bar{K} = [-\rho,\rho]^{1+m}$ if there exists a positive real number ρ such that $\|\phi(x)\| \le \rho/2$, $\|g(x)\| \le \rho$ $\forall x \in X$ or we can set $\bar{K} = \{(u,v) \in \mathbb{R} \times \mathbb{R}^m : u \le M, v \le M\}$ if $\phi(x) \le M$, $g_i(x) \le M$ $i=1,..,m$, $\forall x \in X$.

Let

(29) $$S=\{s: \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}, s(u,v,\omega)=u-\delta(v,\omega), \ \omega \in \Omega\}$$

be a class of strong separation functions, where $\delta(\cdot,\omega): \mathbb{R}^m \to \mathbb{R}$ and Ω is a set of parameters. For Lemma 3.1, the systems (11) and (30)

$$(30) \qquad s(f(x),g(x),\omega)=\phi(\bar{x})-\phi(x)-\delta(g(x),\omega) \le 0 \quad , \forall x \in X \ , \ \forall \omega \in \Omega$$

cannot be both impossible, so that the optimality of $\bar{x}$ implies the validity of (30). Consequently (30) becomes a necessary optimality condition too, when the class S satisfies (31)

$$(31) \qquad \bigcup_{\omega \in \Omega} (\bar{K} \cap \mathrm{lev}_{>0} s(u,v,\omega)) = \bar{K} \cap H \ .$$

The following theorem holds :

Theorem 3.1

Consider the class of strong separation functions (29) satisfying (31). Then (30) is a sufficient condition for $\bar{x}$ to be optimal.

Proof: Suppose that there exists $\hat{x} \in X$ such that $g(\hat{x}) \ge 0$, $\phi(\hat{x}) < \phi(\bar{x})$. Then $(\phi(\bar{x})-\phi(\hat{x}),g(\hat{x})) \in \bar{K} \cap H$ and, for (31), there exists $\bar{\omega} \in \Omega$ such that $(\phi(\bar{x})-\phi(\hat{x}),g(\hat{x})) \in \mathrm{lev}_{>0} s(u,v,\bar{\omega})$ and this contradicts (30) for $x = \hat{x}$. #

Consider now the case $\bar{K} = \mathbb{R} \times \mathbb{R}^m$ and suppose that the class (29) satisfies the following properties:

$$(32) \qquad \delta(v,\omega) = +\infty \ , \quad \forall v \not\geq 0 \ , \ \forall \omega \in \Omega$$

$$(33) \qquad \bigcup_{\omega \in \Omega} \mathrm{lev}_{>0} s(u,v,\omega) = \mathrm{int}\ H$$

The following Lemma holds :

Lemma 3.2 Conditions (33) and (34) are equivalent

$$(34) \qquad \forall v > 0 \ , \ \inf_{\omega \in \Omega} \delta(v,\omega) = 0$$

Proof: <u>(33) ⇒ (34)</u>. Ab absurdo, suppose that there exists $\bar{v}>0$ such that $\inf_{\omega} \delta(\bar{v},\omega) = \ell > 0$. Let $\bar{u}$ be such that $0<\bar{u}<\ell$. We have $\bar{u}-\delta(\bar{v},\omega)<0$ $\forall\, \omega\in\Omega$ and thus $(\bar{u},\bar{v}) \notin \mathrm{lev}_{>0} s(u,v,\omega)$ $\forall\, \omega\in\Omega$ with $(\bar{u},\bar{v}) \in \mathrm{int}\, H$ and this contradicts (33).

<u>(34) ⇒ (33)</u>. The condition $\mathrm{lev}_{>0} s \subseteq H$ implies $\delta(v,\omega) \geq 0$ $\forall v>0$, $\forall \omega\in\Omega$. By definition and from (32) we have $\bigcup_{\omega\in\Omega} \mathrm{lev}_{>0} s \subseteq \mathrm{int}\, H$, and thus it is sufficient to show that $\mathrm{int}\, H \subseteq \bigcup_{\omega} \mathrm{lev}_{>0} s$.

Let $(\bar{u},\bar{v}) \in \mathrm{int}\, H$. Since $\inf_{\omega} \delta(\bar{v},\omega) = 0$, there exists $\bar{\omega}$ such that $0 \leq \delta(\bar{v},\bar{\omega}) < \bar{u}$; consequently $(\bar{u},\bar{v}) \in \mathrm{lev}_{>0} s(u,v,\bar{\omega})$ and this implies $\mathrm{int}\, H \subseteq$ $\subseteq \bigcup_{\omega} \mathrm{lev}_{>0} s$. #

Consider again problem (10) and set $R^{\circ} = \{x \in X: g(x) > 0\}$, $R^{*} = \{\bar{x} \in R: \phi(\bar{x}) = \min_{x\in R} \phi(x)\}$.

The following theorem gives a necessary and sufficient condition for $\bar{x}$ to be optimal which is weaker than the one stated in Theorem 3.1.

Theorem 3.2

<u>Consider problem</u> (10) <u>and assume that</u> ϕ <u>is continuous</u>, $R^{*} \neq \emptyset$, $R^{\circ} \neq \emptyset$, $R = \mathrm{cl}\, R^{\circ}$. <u>Let</u> S <u>be the class of strong separation functions</u> (29) <u>satisfying</u> (32) <u>and</u> (33). <u>Then</u> i) <u>and</u> ii) <u>hold</u>:

i) $\bar{x}$ <u>is an optimal solution for</u> (10) <u>iff</u>

(35) $$\sup_{\omega\in\Omega} \sup_{x\in X} s(f(x),g(x),\omega) \leq 0$$

ii) <u>if</u> $\hat{x}$ <u>is a feasible solution of</u> (10) <u>such that</u>

(36) $$\phi(\hat{x}) = \inf_{\omega\in\Omega} \inf_{x\in X} [\phi(x)+\delta(g(x),\omega)]$$

<u>then</u> $\hat{x}$ <u>is an optimal solution for problem</u> (10).

Proof: i) *Necessity:* Suppose that $\bar{x}$ is an optimal solution for (10). Then (30) holds and this implies (35). *Sufficiency.* Ab absurdo, suppose that there exists $\hat{x}$ such that $g(\hat{x}) \geq 0$, $\phi(\hat{x}) < \phi(\bar{x})$. Since ϕ is continuous at $\hat{x}$ and $R = \mathrm{cl}\, R^\circ$, there exist a suitable neighbourhood $U(\hat{x})$ of $\hat{x}$ and $x^\circ \in R^\circ \cap U(\hat{x})$ such that $\phi(x^\circ) < \phi(\bar{x})$. Setting $\bar{u}=\phi(\bar{x})-\phi(x^\circ)>$ >0, $\bar{v}=g(x^\circ)>0$, we have $(\bar{u},\bar{v}) \in \mathrm{int}\, \mathcal{H}$ and, from (33), there exists $\bar{\omega}$ such that $(\bar{u},\bar{v}) \in \mathrm{lev}_{>0}\, s(u,v,\bar{\omega})$. Consequently $\sup_{x \in X} s(f(x),g(x),\bar{\omega}) > 0$ and this contradicts (35).

ii) Let $\bar{x}$ be an optimal solution for (10) and set $m = \phi(\bar{x})$, $\ell = \inf_{\omega \in \Omega} \inf_{x \in X} [\phi(x)+\delta(g(x),\omega)]$. Since ϕ is continuous at $\bar{x}$ and $R = \mathrm{cl}\, R^\circ$, for every $\varepsilon > 0$ there exist a suitable neighbourhood $U(\bar{x})$ of $\bar{x}$ and $x^\circ \in R^\circ \cap U(\bar{x})$, such that $\phi(x^\circ) \leq \phi(\bar{x}) + \varepsilon = m + \varepsilon$. From (34) there exists $\bar{\omega}$ such that $0 \leq \delta(g(x^\circ),\bar{\omega}) < \varepsilon$ and, consequently, $\ell \leq \inf_{x \in X} (\phi(x)+\delta(g(x),\bar{\omega}) \leq$ $\leq \phi(x^\circ)+\delta(g(x^\circ),\bar{\omega}) \leq m + 2\varepsilon$ and this implies $\ell \leq m$.
On the other hand, since $\delta(g(x),\omega) \geq 0$ $\forall x \in X$, $\forall \omega \in \Omega$ we have $m \leq \ell$. It follows $\ell = m$. #

5. LANGRANGEAN PENALTY APPROACHES

Penalty approaches are a natural extension of the original Lagrangean method and aim to get an optimal solution of a constrained extremum problem by solving a sequence of unconstrained ones.

More exactly, exterior penalty function methods usually solve problem (10) by a sequence of unconstrained minimization problems whose optimal solutions approach the solution of (10) *outside* the feasible set so that the sequence of unconstrained minima converges to a feasible point of the constrained problem that satisfies some sufficient optimality conditions. On the contrary, interior penalty function methods solve (10) through a sequence of unconstrained optimization problems whose minima are at points in the interior of feasible set; staying in the interior is ensured by formulating a barrier function by which an infinitely large penalty is imposed for crossing the boundary of the feasible set from the inside.

In this section it will be shown that these approaches can be viewed in terms of weak and strong separation functions. To this end consider problem (10), with $X = \mathbb{R}^n$ and the continuous functions $p_r : \mathbb{R}^n \to \mathbb{R}$; $r = 1,2,..$ such that

$$p_r(v) = 0 \quad \text{if} \quad v \geq 0; \; p_r(v) > 0 \quad \text{if} \quad v \ngeq 0$$

(37)

$$p_{r+1}(v) > p_r(v) \; ; \; \lim_{r\to+\infty} p_r(v) = +\infty \quad \text{if} \quad v \ngeq 0$$

The function $w(u,v,r) = u - p_r(v)$ is, for any r, a weak separation function and, moreover, it is easily seen that

(38a) $$w(\cdot,r) \text{ is continuous for any } r$$

(38b) $$\text{lev}_{>0} w(\cdot,r) \supset \text{lev}_{>0} w(\cdot,r+1)$$

(38c) $$\bigcap_{r=1}^{+\infty} \text{lev}_{>0} w(\cdot,r) = H$$

(38d) $$\forall h \in H, \; \exists K(h) > 0 \text{ such that } w(h,r) \geq K(h) \quad \forall r$$

The following theorem holds (Ref.16).

Theorem 4.1

Let W be a class of weak separation functions satisfying (38). Then system (1) is impossible iff

(39) $$\inf_r \sup_{x\in X} w(F(x),r) \leq 0$$

Since w(u,v,r) is, for any r, a weak separation function, i) of Theorem 1.1 can be applied and (3a) becomes

$$\phi(\bar{x})-\phi(x)-p_r(g(x)) \le 0 \quad , \quad \forall x \in \mathbb{R}^n , \tag{40}$$

and is a sufficient condition for the feasible $\bar{x}$ to be optimal . Such a condition can be weakened by applying theorem 4.1; (39) becomes

$$\lim_{r\to+\infty} \inf_{x\in\mathbb{R}^n} [\phi(x)+p_r(g(x))] \ge \phi(\bar{x}) \tag{41}$$

and is a sufficient condition weaker than (41). Denote by Φ_r the infimum in (41). From (37) we deduce

$$\Phi_1 \le \Phi_2 \le \dots \le \Phi \stackrel{\Delta}{=} \inf_{x\in R} \phi(x) \tag{42}$$

Assume that $\exists\bar{r}$ such that $\Phi_{\bar{r}} > -\infty$, and that there is a proper $x^r \in R^n$ such that $\phi(x^r) = \Phi_r$, $\forall r \ge \bar{r}$. If $\bar{x}$ is any limit point of sequence $\{x^r\}$, then condition (41) is fulfilled and theorem 1.1 gives the optimality of $\bar{x}$. The construction of sequence $\{x^r\}$ by solving the infimum problems in (41) is the well-known *exterior penalty method* and p_r is said a *penalty function*; if the above convergence can be ensured after a finite number of steps, i.e. if $\exists\bar{r}$ such that (40) is fulfilled at $r = \bar{r}$, then p_r is said *exact penalty function* (Ref.20). Hence, the conditions for a penalty function to be exact can be regarded as conditions which ensure (40) instead of (42).

A particular case of (38), corresponding to a well-known penalty function is

$$w(u,v;r,\alpha) = u-r \sum_{i=1}^{m} (-\min\{0,v_i\})^{\alpha} \quad , \quad \alpha \ge 1 .$$

A more general class of functions satisfying (38) is contained in Ref. 20, where the case of both equality and inequality constraints is considered. The latter requires only formal changes in the above reasoning. In fact, if the constraints of (10) are $g(x) = 0$, it is enough to replace in (7) $V = \mathbb{R}^m_+$ with $V = \{0\}$, so that now $\mathcal{H} = \{(u,v) \in \mathbb{R}^{1+m} : u>0; v=0\}$.

In such a case a weak separation function is for instance the following one

$$w(u,v,\lambda,r)=u+\langle\lambda,v\rangle-r\langle v,v\rangle, \text{ with } \lambda\in\mathbb{R}^m_+ , r\in\mathbb{R}_+ ,$$

which corresponds to the so-called augmented Lagrangean approach (Ref. 31).

It follows that *exterior penalty approach can be formulated in terms of weak separation.*

Now, consider again problem (10) and assume that ϕ, g are continuous, $R = \text{cl } R^o$, $R^* \neq \emptyset$.

Let $\{\mu_k\}$ be a sequence of real numbers tending to infinity such that, for each k, $k = 1,2,..$, $\mu_k > 0$, $\mu_{k+1} > \mu_k$. Assume that, for each k, problem (43) has a solution

(43) $$\min_{x\in R^o} [\phi(x) + \frac{1}{\mu_k} \delta(g(x)]$$

where δ is a continuous function such that $\delta(g(x)) \geq 0$ if $g(x) > 0$; $\delta(g(x)) = +\infty$ if $g(x) \not> 0$. Interior penalty function methods solve, for each k, problem (43) obtaining the point x_k; any limit point of the sequence $\{x_k\}$ is a solution to problrm (10). This procedure corresponds to find a feasible solution $\hat{x}$ satisfying (37).

It follows that *interior penalty approach can be formulated in terms of strong separation.*

In this section it is shown that a dual problem naturally arises when optimality is studied through alternative. In this way some generalizations are easily achieved.

Assume that $\bar{x}$ is an optimal solution for problem (10)and,moreover, that a constraint qualification holds.

Consider the class of weak separation functions (12) satisfying (13), with $\theta = 1$, and the class of strong separation functions (29).

Since $w(f(x),g(x),\omega) = \phi(\bar{x}) - \phi(x) + \gamma(g(x),\omega)$,it results $w(f(\bar{x}), g(\bar{x}),\omega) = \gamma(g(\bar{x}),\omega) \geq 0 \quad \forall\omega \in \Omega$, so that

$$\sup_{x \in X} w(f(x),g(x),\omega) \geq 0 \qquad \forall\omega \in \Omega$$

or, equivalently,

$$\phi(\bar{x}) \geq \inf_{x \in X}[\phi(x) - \gamma(g(x),\omega)] \qquad \forall\omega \in \Omega.$$

It follows

(44)
$$\phi(\bar{x}) \geq \sup_{\omega \in \Omega} \inf_{x \in X}[\phi(x) - \gamma(g(x),\omega)] \;;$$

In a similar way we can use strong alternative: since $s(f(x),g(x),\omega) = \phi(\bar{x}) - \phi(x) - \delta(g(x),\omega)$, it results $s(f(\bar{x}),g(\bar{x}),\omega) = -\delta(g(\bar{x}),\omega) \leq 0$ $\forall\omega \in \Omega$ so that $\inf_{x \in X} s(f(x),g(x),\omega) \leq 0 \quad \forall\omega \in \Omega$. It follows

(45)
$$\phi(\bar{x}) \leq \inf_{\omega \in \Omega} \inf_{x \in X}[\phi(x) + \delta(g(x),\omega)]$$

Set

$$L(x,\omega)=\phi(x)-\gamma(g(x),\omega) \quad ; \quad L_s(x,\omega)=\phi(x)+\delta(g(x),\omega)$$

and assume, for sake of simplicity, that any infimum (supremum) which appears in (44), (45), is achieved as a minimum (maximum).

Then two news problems can be associated to P:

$$D: \max_{\omega\in\Omega} \min_{x\in X} L(x,\omega) \quad ; \quad D_s: \min_{\omega\in\Omega} \min_{x\in X} L_s(x,\omega)$$

Problem D is called the generalized Lagrangian dual an it reduces to the usual Lagrangian dual when γ is linear. We refer to D and D_s, respectively, as the *weak dual* and the *strong dual* of the *primal* problem P.

Relation (44) is known as the *weak duality theorem*. The difference Δ between the left-hand side and the right-hand side of (44) is the *duality gap*.

The following theorem is a general formulation of the so-called strong duality theorem.

Theorem 5.1

Consider the pair of problems P and D. Then $\Delta = 0$ iff there exists $\bar{w}\in W_2$ which guarantees alternative.

Proof: $\Delta = 0 \Leftrightarrow \phi(\bar{x}) \le \max_{\omega\in\Omega} \min_{x\in X} L(x,\omega)$

$$\Leftrightarrow \exists\bar{\omega}\in\Omega:\ \phi(\bar{x}) \le \min_{x\in X} (\phi(x)-\gamma(g(x),\bar{\omega}))$$

$$\Leftrightarrow \exists\bar{\omega}\in\Omega:\ \phi(\bar{x})-\phi(x)+\gamma(g(x),\bar{\omega}) \le 0 \quad , \quad \forall x \in X$$

or, equivalently, iff the function $\bar{w}(u,v,\bar{\omega}) = u-\gamma(v,\bar{\omega})$ is a weak separation function which guarantees alternative. #

Let us note the strict connection between the strong duality theorem and the necessary optimality conditions. To this end note that ny condition which guarantees the validity of a necessary optimality condition is also a condition which ensuresto be zero the duality gap , when the same class of weak separation functions is adopted. As a consequence, the results obtained in section 2 can be used to characterize classes of problems having $\Delta = 0$.

For instance, when P is convex and a constraint qualification holds, if we choose the class of linear weak separation functions $w(u,v,\omega) = u + \langle\omega,v\rangle$, that is $L(x,\omega) = \phi(x) - \langle\omega,g(x)\rangle$, it results $\Delta = 0$; when P is the linear fractional problem or when the objective function ϕ is concave and g is affine, if we choose the class of exponential weak separation functions $w(u,v,\lambda,\mu) = u + \sum_{i=1}^{m} \lambda_i \exp(-\mu_i v_i)$, that is $L(x,\lambda,\mu) = \phi(x) - \sum_{i=1}^{m} \lambda_i \exp(-\mu_i g_i(x))$, it results $\Delta = 0$ (Ref.28).

Consider now problem D in the case where $L(x,\omega)$ is the usual Lagrangean function. It results, in general, $\Delta > 0$. In the image space the duality gap can be easily characterized.

For $\mu \in \mathbb{R}^m$ and $\alpha \in \mathbb{R}$ let us set

$$H(\mu,\alpha) = \{(u,v) \in \mathbb{R} \times \mathbb{R}^m : u + \langle\mu,v\rangle \le \alpha\}$$

$$A = \{\alpha \in \mathbb{R}_+ \ / \ \exists\bar{\mu} \in \mathbb{R}^m_+ : H(\bar{\mu},\alpha) \supseteq E\}$$

$$\alpha_0 = \inf A \quad \text{if} \quad A \neq \emptyset$$

The following theorem holds (Ref. 29):

Theorem 5.2

i) <u>If</u> $A = \emptyset$, <u>then</u> $\Delta = +\infty$. ii) <u>If</u> $A \neq \emptyset$, <u>then</u> $\Delta = \alpha_0$.

Working in the image space it is possible to establish an upper bound for Δ.

Setting $v^* = \max_{\omega\in\Omega} \min_{x\in X} L(x,\omega)$, it results (Ref.29)

(46) $$v^* \ge F(-\sigma(g)) - \rho(\phi)$$

where F is the perturbation function

$$F(\varepsilon) \triangleq \min \phi(x) \quad , \; x \in R_{\varepsilon} = \{x \in X: \; g(x) \geq \varepsilon\}$$

and $\rho(\phi)$, $\sigma(g)$ are, respectively, the lack of convexity of function ϕ and the lack of concavity of function g.

As a particular case of (46) we have, when g is concave,

(47) $$0 \leq \Delta \leq \rho(\phi) \quad .$$

Now, consider again the strong dual D_S and set

$$\Phi(\omega) \triangleq \min_{x \in X}[\phi(x) + \delta(g(x),\omega)].$$

We refer to (45) as the *strong duality theorem* .

The difference between the right-hand side and the left-hand one of (45) is called *strong duality gap.*

As an obvious consequence of Theorem 3.1, we have the following :

Theorem 5.3

Consider the pair of problems P and D_S and assume that the class of strong separation functions (29) satisfies (31). Then, the strong duality gap is zero.

7. REGULARITY

In section 2 we have pointed out that, if $w(f(x),g(x),\bar{\theta},\bar{\lambda}) = \bar{\theta}(\phi(\bar{x}) - \phi(x) + \langle\bar{\lambda},g(x)\rangle$ is a weak separation function which guarantees alternative, then

(48) $$\bar\theta(\phi(\bar x)-\phi(x)) + <\bar\lambda, g(x)> \le 0 \quad , \quad \forall x \in X$$

becomes a sufficient condition for $\bar x$ to be optimal for problem (10).For this optimality condition and any other which involves Lagrange multipliers the problem of *regularity* arises, that is the problem of finding conditions which guarantees that θ is different from zero. When such conditions involve only the constraints, they are referred to as *constraint qualifications*. Now, we will see how the image space suggests some simple ideas concerning conditions under which $\theta \neq 0$.

Let us note that the validity of (48) is equivalent to state the existence of an hyperplane $\Gamma \subset \mathbb{R} \times \mathbb{R}^m$ which separates E and H,that is, $E \subset \Gamma^-$, $H \subset \Gamma^+$ where

$$\Gamma^- = \{(u,v) \in \mathbb{R} \times \mathbb{R}^m : \bar\theta u + <\bar\lambda, v> \le 0\} \; ;$$

$$\Gamma^+ = \{(u,v) \in \mathbb{R} \times \mathbb{R}^m : \bar\theta u + <\bar\lambda, v> \ge 0\}, \; \Gamma = \Gamma^- \cap \Gamma^+ .$$

From a geometrical point of view, a regularity condition for (48)is equivalent to the one which ensure that Γ does not contain the line $r=\{(u,0: u \in \mathbb{R}\}$.

Consider now the simplest case where E is convex, that is the function $F(x) = (f(x),g(x))$ is cl H-convexlike and let T be the tangent cone(8) of E at the origin. It is easy to show that Γ separates E and

(8) The tangent cone $T(\bar h)$ to A at $\bar h \in A$ is defined as the set of h for which there exist a sequence $\{h^r\} \subset A$ and a positive sequence $\{\alpha_r\} \subset \mathbb{R}_+$, such that $\lim_{r\to+\infty} h^r = \bar h$, $\lim_{r\to+\infty} \alpha_r (h^r - \bar h) = h - \bar h$.

H iff Γ separates T and H. The reason of introducing T is given by the fact that we can characterize regularity in terms of disjunction between T and int U $=\{(u,0)\in \mathbb{R}\times\mathbb{R}^m : u>0\}$.

The following theorem holds (Ref.16).

Theorem 6.1

Consider problem (10) and assume that $F(x) = (f(x),g(x))$ is cl H-convexlike. Then (48) is fulfilled with $\bar{\theta}\neq 0$ iff

(49) $$T\cap \text{int } U = \emptyset .$$

Let us note that, in the convex case, constraint qualifications are sufficient conditions for (49) to be satisfied. Condition (49) is given in the image space and it is equivalent to the following one given in the original space (Ref.25).

Condition 1. For every sequence $\{x^r\}\subset X$ and for every positive sequence $\{\alpha_r\}\subset\mathbb{R}^+$, we have

(50) $$\lim_{r\to+\infty} \alpha_r(f(x^r),g(x^r)) \neq (u^\circ,0), \quad u^\circ\in \text{int } U$$

or such a limit does not exist.

In the case where E is not convex, (49) becomes a necessary condition to have regularity. Conditions under which (49) is sufficient too, and some other regularity conditions for the differentiable case, can be found in Ref.25.

The following theorem [35] states a necessary and sufficient regularity condition which generalizes (49).

Theorem 6.2

Consider problem (10) and assume that (48) holds. Then (48) is fulfilled with $\bar{\theta} \neq 0$ iff

(51) $$\bar{E} \cap \operatorname{int} U = \emptyset$$

where[9] $\bar{E} = \operatorname{cl}(\operatorname{conv}(\operatorname{con} E))$.

---

8. THE IMAGE OF A CONSTRAINED PROBLEM

In this section we will see how the image space can be used in order to find necessary and/or sufficient conditions under which an optimal solution for a constrained extremum problem exists.

With this end let m and n be positive integers; assume we are given $X \subseteq \mathbb{R}^n$, $\phi: X \to \mathbb{R}$, $g: X \to \mathbb{R}^m$; and let $V \subseteq \mathbb{R}^m$ be a closed convex cone, containing the origin 0 of $\mathbb{R}^m$.

We consider the following constrained extremum problem:

(P) $$\min \phi(x) \quad ; \quad x \in R \triangleq \{x \in X : g(x) \in V\} ,$$

and we assume that $R \neq \emptyset$.

Note that, when $V = \mathbb{R}^m_+$, (P) collapses to (10).

(9) con A denotes the cone generated by A.
conv B denotes the convex hull of B.

Given a point $\bar{x} \in R$, we set $f_{\bar{x}}(x) = \phi(\bar{x}) - \phi(x)$ and $F_{\bar{x}}(x)=(f_{\bar{x}}(x)$, $g(x))$, so that $F_{\bar{x}}: X \to \mathbb{R} \times \mathbb{R}^m$. Moreover, we define $K_{\bar{x}} \triangleq \{(u,v) \in \mathbb{R} \times \mathbb{R}^m : u = f_{\bar{x}}(x);\ v = g(x);\ x \in X\}$ and we call *image* of problem (P), with respect to the point $\bar{x}$, the problem:

$$(P_{\bar{x}}) \qquad \max\ (u) \quad , \quad \text{s.t.} \quad (u,v) \in R_{\bar{x}} \triangleq \{(u,v) \in K_{\bar{x}} : v \in V\}$$

Set $E_{\bar{x}} = K_{\bar{x}} - \text{cl}\, H$, we shall call *extended image* of problem (P), with respect to the point $\bar{x}$, the problem:

$$(P_{\bar{x}}^{(e)}) \qquad \max\ (u) \quad , \quad \text{s.t.} \quad (u,v) \in R_{\bar{x}}^{(e)} \triangleq \{(u,v) \in E_{\bar{x}} : v \geq 0\}.$$

The sets $K_{\bar{x}}$ and $E_{\bar{x}}$, and therefore the problems $(P_{\bar{x}})$ and $(P_{\bar{x}}^{(e)})$, obviously depend on the choice of $\bar{x}$ in R; even if such dependence is of a very particular kind. Indeed, if $\hat{x} \in R$, it is easily seen that:

$$K_{\hat{x}} = K_{\bar{x}} + \{(\phi(\hat{x}) - \phi(\bar{x}), 0)\} \quad ; \quad E_{\hat{x}} = E_{\bar{x}} + \{(\phi(\hat{x}) - \phi(\bar{x}), 0)\}.$$

Moreover, it can be easily verified that, if $\bar{x}, \hat{x} \in R$, problem $(P_{\bar{x}})$ [or $(P_{\bar{x}}^{(e)})$] has an optimal solution, iff $(P_{\hat{x}})$ [respect. $(P_{\bar{x}}^{(e)})$] does have.

For this reason, when we consider properties which hold independently on the choice of $\bar{x}$ over R, we shall drop $\bar{x}$ from the corresponding notation.

We shall now give some general results, which should be of some help in analysing the relations that hold between a constrained extremum problem and its (extended) image.

Lemma 7.1 The following equalities hold:

$$\phi(\bar{x}) - \inf_{x\in R} \phi(x) = \sup_{(u,v)\in \mathcal{R}} (u) = \sup_{(u,v)\in \mathcal{R}^{(e)}} (u) = \sup_{\substack{(u,v)\in E \\ v=0}} (u) ,$$

where the image and the extended image are considered with respect to the point $\bar{x} \in R$.

Proof: The first equality follows easily, observing that $x \in R \Longrightarrow (\phi(\bar{x})-\phi(x),g(x)) \in K$ and $(u,v)\in K \Rightarrow \exists\, x \in R$ such that $u = \phi(\bar{x}) - \phi(x)$, $v = g(x)$. To prove the other equalities, note first of all that, since $(u,v) \in \mathcal{R} \Rightarrow (u,0)=(u,v)-(0,v) \in E$, we have $\sup_{(u,v)\in\mathcal{R}} (u) \le \sup_{\substack{(u,v)\in E \\ v=0}} (u)$.

From the inclusion $\{(u,v)\in E : v = 0\} \subseteq \mathcal{R}^{(e)}$, we then get $\sup_{\substack{(u,v)\in E \\ v=0}} (u) \le$

$\le \sup_{(u,v)\in\mathcal{R}^{(e)}} (u)$. Observe finally that, for every $(u,v)\in \mathcal{R}^{(e)}$, there must

be, by definition of E, a point $(\tilde{u},\tilde{v})\in\mathcal{R}$ with $\tilde{u} \ge u$. From this we get $\sup_{(u,v)\in\mathcal{R}^{(e)}} (u) \ge \sup_{(u,v)\in\mathcal{R}} (u)$, and this is the last inquality needed to complete the proof. #

Theorem 7.1

The following conditions are equivalent:

(i) problem (P) has a global minimum point;

(ii) problem $(\mathcal{P})$ has a global minimum point;

(3i) problem $(\mathcal{P}^{(e)})$ has a global minimum point.

Moreover, if one of the above conditions is verified, and if the images are considered with respect to $\bar{x}$, we have:

(52) $$\min_{x\in R} \phi(x) = \phi(\bar{x}) - \max_{(u,v)\in \mathcal{R}} (u) = \phi(\bar{x}) - \max_{(u,v)\in \mathcal{R}^{(e)}} (u).$$

Proof: (i) $\Leftrightarrow$ (ii). Let $\hat{x} \in X$, $\hat{u} = f(\hat{x})$, $\hat{v} = g(\hat{x})$. We have $\hat{x} \in P$, iff $(\hat{u},\hat{v}) \in \mathcal{R}$. Let now $\tilde{x} \in R$, $\tilde{u} = f(\tilde{x})$, $\tilde{v} = g(\tilde{x})$, we then have $\phi(\tilde{x}) < \phi(\hat{x})$, iff $\tilde{u} > \hat{u}$. The thesis follows easily from the above remarks. (ii) $\Rightarrow$(3i). It is a consequence of lemma 7.1 noting that $\mathcal{R} \subseteq \mathcal{R}^{(e)}$. (3i) $\Rightarrow$(2i). Let $(\hat{u},\hat{v}) \in \mathcal{R}^{(e)}$ be such that $\hat{u} = \max_{(u,v)\in\mathcal{R}^{(e)}} (u)$, then there exists $(\tilde{u},\tilde{v}) \in \mathcal{R}$, such that $\tilde{u} \geq \hat{u}$, hence the thesis follows from lemma 7.1. The second part of the theorem is a consequence of the first part and of lemma 7.1. #

Now we shall study the existence of the minimum of problem (P), by means of its relationships with the extended image problem $(\mathcal{P}^{(e)})$. More precisely, we shall give some sufficient condition, one of which generalizes the well-known Weierstrass condition (of semicontinuity of functions and compactness of domain) to the kind of problem that we consider.

First of all observe that, as an easy consequence of Theorem 7.1, we get the following:

Theorem 7.2

Problem (P) has a global minimum point, iff the set $D = \mathcal{E} \cap \{(u,v) \in \mathbb{R} \times \mathbb{R}^m : v=0\}$ is closed and one has $D \neq \emptyset$ and $\sup_{(u,v)\in D} (u) < +\infty$.

Note that the assumption $D \neq \emptyset$ is equivalent to requiring $R \neq \emptyset$, whereas, account taken of Lemma 7.1, the assumption $\sup_{(u,v)\in D} (u) < +\infty$ is equivalent to $\inf_{x\in R} \phi(x) > -\infty$.

From this proposition we get immediately the following:

Corollary 7.1 Let $D \neq \emptyset$ and $\sup_{(u,v)\in D} (u) < +\infty$. Then, if E is closed, problem (P) has a global minimum point.

This corollary allows us to claim that, once sure that $D \neq \emptyset$ and $\sup_{(u,v)\in D} (u) < +\infty$, every condition ensuring the closure of E, or of a subset containing D, ensures also the existence of the minimum of (P). Sufficient conditions for the closure of E can be found in Ref.7,8,33.

The following theorem holds (Ref.33).

Theorem 7.3
Let $R \neq \emptyset$. Then, if X is compact and $F = (f,g)$ is (cl H)-upper semicontinuous, problem (P) has a global minimum point.

Note that, when $V = \mathbb{R}^m_+$, and, therefore cl $H = \mathbb{R}^{1+m}_+$, this result collapses to the well-known Weierstrass condition.

Some other necessary and/or sufficient conditions can be found in Ref.33.

9. CONCLUDING REMARKS

Theorems of the alternative can be considered as a general framework within which optimality conditions and related topics can be studied. By means of a theorem of the alternative we have introduced weak and strong alternative and we have shown how sufficient optimality conditions of the saddle-point type, exterior penalty method and weak duality can be deduced and/or interpreted by weak alternative (see Table 1). Furthermore, we have seen that strong alternative (see Table 1)produces strong duality, penalty interior scheme and necessary optimality conditions (different from the stationary type; these, on the contrary, turn out to be a further deepending of weak analysis and are based on local arguments). Connections among the topics which appear in the first (second) column of table 1 are shown.

The general approach consists in introducing the image space, in studying a certain question on it, and then, when a result has been obtained on the image space, to obtain its counterimage, namely the corresponding result in the original space.

Other kinds of problems, which can be reduced to the above scheme are combinatorial problem, discrete optimization problems, variational inequalities (Ref.16). This general scheme can be also extended to vector extremum problems (Ref.9) and to multifunctions.

LOGICAL CORRESPONDENCE

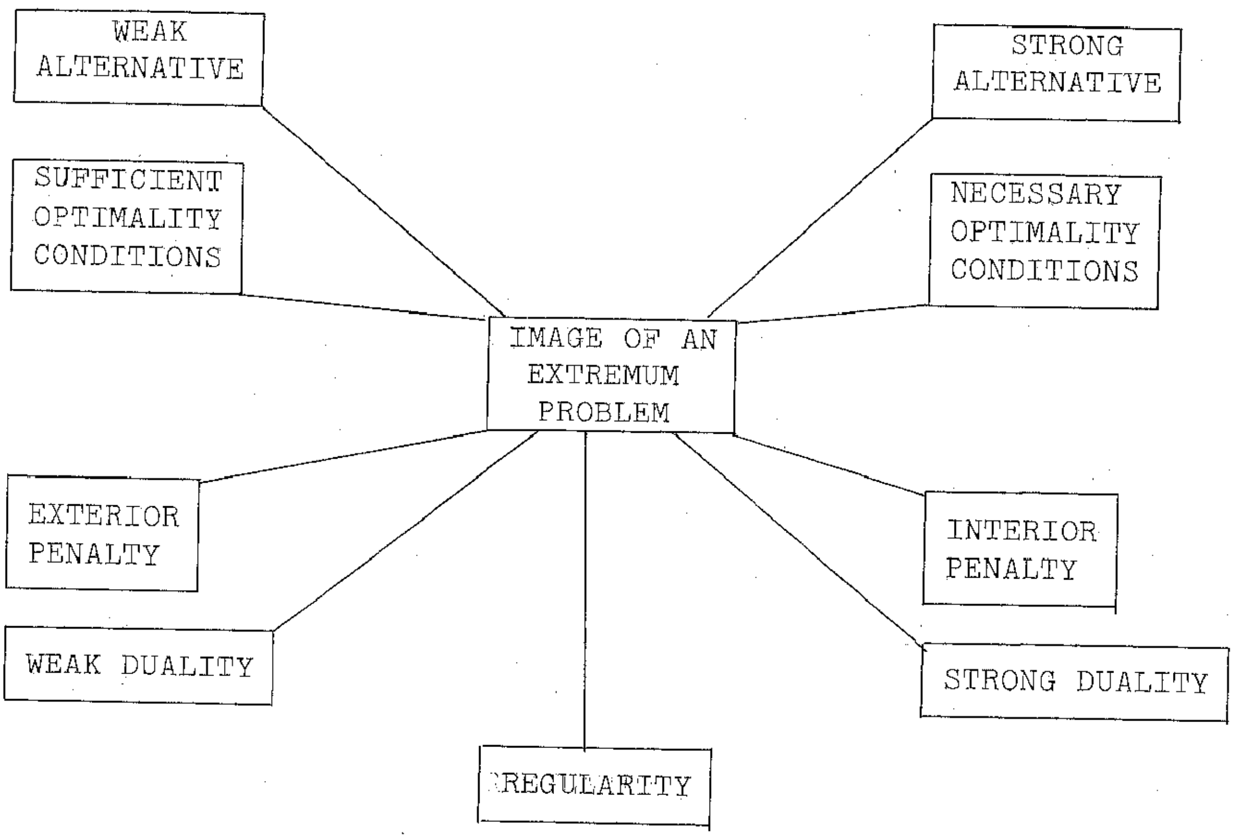

TABLE 1

REFERENCES

[1] M.Avriel. 'Non linear programming. Analysis and methods'. Prentice Hall, (1976).

[2] M.S.Bazaraa, C.M.Shetty. 'Foundations of optimizations'. *Lecture Notes in Economics and Mathematical Systems*. Springer-Verlag(1976).

[3] M.S.Bazaraa. 'A theorem of the alternative with applications to convex programming: optimality, duality and stability'. *J. of Math. Analysis and Appl.*, vol.41, (1973), 701-715.

[4] A.Ben-Israel, A.Ben-Tal, and S.Zlobec. 'Optimality in nonlinear programming. A feasible directions approach'. *J.Wiley*, New York, (1981).

[5] A.Ben-Israel, A.Charnes, and K.O.Kortanek. 'Asymptotic duality over closed convex sets'. *J. Math. Analysis Appl.*, vol.35, (1971), 677-691.

[6] C.Beoni. 'A generalization of Fenchel duality theory'. *Jou Optimization Th. Appl.*, vol.49, N.3, June 1986.

[7] G.R.Bitran, T.Magnanti. 'The structure of admissible points with respect to Cone Dominance'. *J.O.T.A.*, vol.29, n.4, (1979), 573-614.

[8] O.Calligaris, P.Oliva. 'Necessary and sufficient conditions for Pareto problems'. *Boll. UMI* (5), 18B (1981), 177-216.

[9] A.Cambini, L.Martein. 'Separation functions and optimality conditions in vector extremum problems'. Technical report A-120, Dept. of Math., Univ. of Pisa.

[10] A.Y.Dubovitskii and A.A.Milyutin. 'The extremum problem in the presence of constraints'. Doklady Akademiie Nauk SSSR, vol.149, (1963), 759-762.

[11] E.Eisenberg. 'On cone functions'. In "Recent advances in mathematical programming". P.Wolfe and R.L.Graves (eds.), *McGraw-Hill*, (1963), 27-33.

[12] K.H.Elster, R.Nehse. 'Optimality conditions for some non- convex problems'. *Lecture Notes in Control and Information Sciences*, n.23, Springer-Verlag, (1980).

[13] H.Everett. 'Generalized Lagrange multiplier method for solving problems of optimum allocation of resources'. *Oper. Res.*, vol.11, (1963), 399-417.

[14] A.V.Fiacco, and G.P.McCormick. 'Nonlinear programming: sequential unconstrained minimization techniques'. *J. Wiley*, New York, (1968).

[15] F.Giannessi. 'Theorems of alternative, quadratic programs and complementarity problems'. In "Variational inequalities and complementarity problems". Cottle-Giannessi-Lions (eds), *J.Wiley*, New York, (1980), 151-186.

[16] F.Giannessi. 'Theorems of the alternative and optimality conditions'. *J.O.T.A.*, vol.42, n.3, March (1984).

[17] F.Giannessi. 'On Lagrangian non-linear multipliers theory for constrained optimization and related topics'. Dept. Math., Appl.Math. Sect. (Optimization Group), Res. Report n.123, (1984).

[18] F.J.Gould. 'Extensions of Lagrange multipliers in nonlinear programming'. *SIAM J. Applied Math.*, vol.17, (1969), 1280-1297.
[19] F.J.Gould. 'Nonlinear pricing: applications to concave programming'. *Operat. Research*, vol.19, n.4, (1971), 1026-1035.
[20] S.P.Han, and O.L.Mangasarian. 'Exact penalty functions in nonlinear programming'. *Math. Programming*, vol.17, (1979), 251-269.
[21] M.Hayashi, H.Komiya. 'Perfect duality for Convexlike programs'. *J. O.T.A.*, vol.38, n.2 (1982).
[22] V.Jeyakumar. 'Convexlike alternative theorems and mathematical programming'. Dept. Math., Univ. of Melbourne, Res.Rep. n.13,(1984). To appear on "*Optimization, Mathematische Operationsforschung und Statistik*".
[23] L.McLinden. 'Duality theorems and theorem of the alternative'.*Proc. Ann. Math. Soc.*, vol.53, n.1, (1975), 172-175.
[24] O.L.Mangasarian. 'Nonlinear programming'. *McGraw-Hill*, New York , (1969).
[25] L.Martein. 'Regularity conditions for constrained extremum problems'.*Jou.Optimization Th.Appl.*,vol.47,N.2,Oct.1985.
[26] L.Martein. 'A necessary and sufficient regularity condition for convex extremum problems'. Technical Report A-90, Dept. of Math., Univ. of Pisa.
[27] L.Martein. 'Sulla separabilità locale in problemi di estremo vincolato'. Proceed. of the VI Meeting "Associazione per la Matematica Applicata alle Scienze Economiche e Sociali", (1982).
[28] L.Martein. 'Sulla dualità lagrangiana esponenziale'. Technical Report A-114, Dept. of Math., Univ. of Pisa.
[29] M.Pappalardo. 'On the duality gap in non convex optimization'. *Mathematics of Operations Research* (to appear).
[30] R.T.Rockafellar. 'Convex Analysis'. Princeton, (1970).
[31] R.T.Rockafellar. 'Penalty methods and augmented Lagrangians in nonlinear programming'. Proceedings of 5-th Conference on Optim.Tech., Springer-Verlag, Berlin, (1973), 418-425.
[32] R.T.Rockafellar. 'Augmented Lagrange multiplier functions and duality in nonconvex programmin'. *SIAM J. on Control*, vol.12, (1974), 268-283.
[33] F.Tardella. 'On the image of a constrained extremum problems'.Tech. Report A-106, Dept. of Math., Univ. of Pisa.
[34] J.Tind,and L.A.Wolsey.'An elementary survey of general duality theory'. *Mathematical Programming*, vol.21, n.3, (1981),241-261.
[35] F.Tardella. (private communication).

PUBBLICAZIONI
del
Dipartimento di Statistica e Matematica Applicata all Economia

Report n. 1 - Some Optimality Conditions in Vector Optimization.(A.Cambini- L.Martein), 1987

Report n. 2 - On Maximizing a Sum of Ratios. (A.Cambini-L.Martein-S.Schaible), 1987

Report n.3 - On the Charnes-Cooper Transformation in Linear Fractional Programming. (G.Gasparotto), 1987

Report n. 4 - Non-linear Separation Theorems, Duality and Optimality. (A.Cambini), 1987