INSTITUTUL
DE
MATEMATICA

INSTITUTUL NATIONAL
PENTRU CREATIE
STIINTIFICA SI TEHNICA

ISSN 0250 3638

TWO PAPERS ON LEXICOGRAPHICAL SEPARATION AND VECTOR OPTIMIZATION

by

Juan Enrique Martinez-Legaz & Ivan Singer

PREPRINT SERIES IN MATHEMATICS

No.73/1985

BUCURESTI

TWO PAPERS ON LEXICOGRAPHICAL SEPARATION AND VECTOR OPTIMIZATION

Juan Enrique Martinez-Legaz*) & Ivan Singer**)

*) Department of Functional Equations, University of Barcelona, Gran Via, 585, Barcelona 08007, Spain.

**) The National Institute for Scientific and Technical Creation, Department of Mathematics, Bd.Pacii 220, 79622 Bucharest, Romania.

LEXICOGRAPHICAL SEPARATION IN R^n

Juan-Enrique Martínez-Legaz and Ivan Singer

Department of Functional Equations, University of Barcelona, Gran Via, 585, Barcelona 08007, Spain, and Department of Mathematics, INCREST, Bd.Păcii 220, Bucharest 79622, Romania

Abstract

We prove a new separation theorem for any two subsets of R^n, with disjoint convex hulls, by linear operators, or isomorphisms, or isometries, in the sense of the lexicographical order of R^n (or, equivalently, by "generalized half-spaces"). Also, we prove that if one of the sets is the non-positive orthant, then the isometry can be taken lexicographically non-negative, or the isomorphism non-negative (in the usual order). We give some applications.

§0. Introduction

The aim of the present paper is to give some theorems on separation, of a new ("lexicographical") type, of two subsets of R^n, by linear operators, in view of applications to vector optimization (see [12]).

The usual separation theorems of two convex sets in R^n, by linear functionals (or, equivalently, hyperplanes), and the various known extensions of these theorems to separation by linear operators, in the sense of the usual order of R^n, require rather strong assumptions (see e.g. [3], [4], [14] and the references therein). A theorem of a new type, on the separation of an arbitrary convex set in R^n and any outside point, by orthogonal matrices, in the sense of the lexicographical order of R^n (instead of the usual order of R^n), has been given in [9]; we shall recall it in theorem 1 below. Some geometric versions and some applications of this separation theorem have been given in [9], [15] (see also remark 1.1 below) and [10], [11] (applications to generalized Lagrangian duality for vector optimization).

In the present paper, we shall extend the case of separation of a convex set and an outside point, in the above sense, to the case of separation of two subsets of R^n, with disjoint convex hulls, by linear operators, or isomorphisms (non-singular matrices) or orthogonal matrices, in the sense of the lexicographical order of R^n (theorem 2). Furthermore, we shall give some results on the particular case when the second set is the non-positive orthant in R^n (theorem 3), or the non-positive orthant without the origin (theorem 4). Finally, we shall give two applications: a characterization of the elements of the "infimum" (in the sense of [1],[2]) of a set in R^n (theorem 5) and an improvement of the mean value theorem for the integral of a vector function with values in a convex subset of R^n; for further applications to surrogate duality in vector optimization and to characterizations of convex sets with convex complements, see [12] and [16] respectively.

§1. Preliminaries. Separation of a convex set from a point

Let us recall now some notions, notations and results, which we shall use in the sequel.

The elements of $\overline{R}^n$ (where $\overline{R}=[-\infty,+\infty]$) will be considered column vectors and the superscript T will mean transpose. We recall that $x=(\xi_1 \dots \xi_n)^T \in \overline{R}^n$ is said to be "lexicographically less than" $y=(\eta_1 \dots \eta_n)^T \in \overline{R}^n$, in symbols, $x <_L y$, if $x \neq y$ and if for $k=\min \{i \in \{1,\dots,n\} \mid \xi_i \neq \eta_i\}$ we have $\xi_k < \eta_k$. We write $x \leqq_L y$ if $x <_L y$ or $x=y$. The notations $y >_L x$, respectively, $y \geqq_L x$, will be also used.

We shall denote by $\mathcal{L}(R^n)$, $\mathcal{U}(R^n)$ and $\mathcal{O}(R^n)$, the families of all linear operators, all isomorphisms, and all linear isometries

$v:R^n \to R^n$, respectively. We shall identify each $v \in \mathcal{L}(R^n)$ with its matrix with respect to the unit vector basis $\{e_j\}_{j=1}^n$ of R^n, that is, we shall write

$$v=(m_{ij})_{i,j=1}^n=(m_1 \ldots m_n)^T=(c_1 \ldots c_n), \qquad (1.1)$$

where $m_i^T=(m_{i1} \ldots m_{in})$ $(i=1,\ldots,n)$ are the rows of $(m_{ij})_{i,j=1}^n$ and $c_j=(m_{1j} \ldots m_{nj})^T=v(e_j)$ $(j=1,\ldots,n)$ are its columns, with e_j being the j-th unit vector $(0 \ldots 0\ 1\ 0 \ldots 0)^T \in R^n$.

We shall consider two orderings of $\mathcal{L}(R^n)$ (hence also of $\mathcal{U}(R^n)$, $\mathcal{O}(R^n)$), namely, the usual order relation $\geqslant$ (in termwise sense) and the lexicographical order $v \geqslant_L 0$ in the sense of [10], defined columnwise (i.e., $v \geqslant_L 0$ if and only if all columns of v are $\geqslant_L 0$). Let us recall now two properties of the lexicographical order, proved in [10], which we shall need in the sequel (for the sake of completeness, we include simple proofs for them).

Lemma 1.1 ([10], corollary 2.3). *For $v \in \mathcal{L}(R^n)$, we have $v \geqslant_L 0$ if and only if*

$$v(x) \geqslant_L 0 \qquad (x \in R^n,\ x \geqslant 0). \qquad (1.2)$$

Proof. Let $v=(c_1 \ldots c_n)$, so $c_1,\ldots,c_n$ are the columns of v. If (1.2) holds, then, since $e_j \geqslant 0$, we have

$$c_j=v(e_j) \geqslant_L 0 \qquad (j=1,\ldots,n), \qquad (1.3)$$

i.e., $v \geqslant_L 0$. Conversely, if (1.3) holds and $x=\sum_{j=1}^n \alpha_j e_j \in R^n$, $x \geqslant 0$, then $\alpha_j \geqslant 0$ $(j=1,\ldots,n)$, whence $v(x)=\sum_{j=1}^n \alpha_j v(e_j) \geqslant_L 0$. ∎

Lemma 1.2 (a particular case of [10], corollary 2.2 b)). *Let $\ell=(\ell_{ij})_{i,j=1}^n \in \mathcal{L}(R^n)$ be a unitary lower triangular matrix. Then*

$$\ell(x) \geqslant_L 0 \qquad (x \in R^n,\ x \geqslant_L 0). \qquad (1.4)$$

Proof. We have

$$\ell(x)=(\xi_1\quad \ell_{21}\xi_1+\xi_2 \ldots \sum_{j=1}^{n-1}\ell_{nj}\xi_j+\xi_n)^T \qquad (x=(\xi_1 \ldots \xi_n)^T \in R^n). \qquad (1.5)$$

If $x >_L 0$ and $i_o=\min\{i \mid \xi_i \neq 0\}$, then $\xi_{i_o} > 0$ and

$$\sum_{j=1}^{k-1} \ell_{kj}\xi_j+\xi_k=0 \quad (k=1,\dots,i_o-1), \quad \sum_{j=1}^{i_o-1} \ell_{i_oj}\xi_j+\xi_{i_o}=\xi_{i_o}>0,$$

whence, by (1.5), $\ell(x) \geqslant_L 0$. ▨

For any subset G of R^n, we shall denote by $\sup_L G$, $\inf_L G$, the supremum, respectively, the infimum of G, for the lexicographical order of $\overline{R}^n$, by co G the convex hull of G and, by INF G, the "infimum" of G in the sense of [1], [2], i.e., the subset of R^n defined as follows: $x \in \mathrm{INF}\, G$ if $x \in \overline{G}$ (the closure of G in R^n) and if there exists no $g \in G$ such that $g < x$ (i.e., such that $g \leqslant x$, $g \neq x$).

For two subsets G_1, G_2 of R^n, we shall say that an operator $v \in \mathcal{L}(R^n)$ _separates_ G_1 _from_ G_2 (in the sense of the lexicographical order of R^n), if

$$v(y_1) <_L v(y_2) \qquad (y_1 \in G_1,\ y_2 \in G_2); \qquad (1.6)$$

clearly, this happens if and only if $-v$ separates G_2 from G_1, so we can speak about separation of the sets G_1 _and_ G_2.

Finally, for simplicity, we shall not assume that the sets occurring in our separation theorems are $\neq \emptyset$ (where $\emptyset$ denotes the empty set), but, instead, we shall make the convention that if one of them is empty, then the separation properties will be considered to hold (vacuously).

The following separation theorem has been proved in [9]:

Theorem 1.1 ([9], p.258). _Let G be a convex subset of_ R^n _and_ $x_o \notin G$. _Then there exists_ $v \in \mathcal{O}(R^n)$ _such that_

$$v(y) <_L v(x_o) \qquad (y \in G). \qquad (1.7)$$

Remark 1.1. a) Theorem 1.1 admits the following geometric interpretation (and proof). We recall that, following Hammer [6], a set S in a linear space E is called a _semi-space at_ x_o (a _hypercone_, in the terminology of [8]), if S is a maximal convex cone with vertex x_o, and $x_o \notin S$, or, equivalently, if S is a maximal convex set such that $x_o \notin S$; in R^n this concept has been also defined, independently, by Motzkin ([13], lecture III). Now, by [6], theorem 1 (or [13], lecture III), _if_ $G \subset E$ _is convex and_ $x_o \notin G$, _then there exists a semi-_

space S at x_o such that $G \subseteq S$. On the other hand, by [15], lemma 1.1, a set $S \subset R^n$ is a semi-space at $x_o \in R^n$ if and only if there exists $v \in \mathcal{O}(R^n)$ such that

$$S=\{y \in R^n | v(y) <_L v(x_o)\} \quad (1.8)$$

(note that in [15] this result has been stated only with $v \in \mathcal{U}(R^n)$, but the proof is the same for $v \in \mathcal{O}(R^n)$; note also that in (1.8) above we have corrected a misprint of [15], namely, we have replaced x_o of [15] by $v(x_o)$). Combining these two results, one obtains again the separation theorem 1.1 above, as has been observed, essentially, in [15]; however, in [9], theorem 1.1 above is obtained by a simple induction proof.

b) If $v \in \mathcal{O}(R^n)$, then for each $z \in R^n$ we have $z=v(x_o)$, where $x_o=v^{-1}(z) \in R^n$. Hence, a set $S \subset R^n$ is a semi-space if and only if there exist $v \in \mathcal{O}(R^n)$ and $z \in R^n$ such that

$$S=\{y \in R^n | v(y) <_L z\}. \quad (1.9)$$

Hence, since $v \in \mathcal{O}(R^n)$ if and only if $-v \in \mathcal{O}(R^n)$, it follows that a set $S \subset R^n$ is a semi-space if and only if there exist $v \in \mathcal{O}(R^n)$ and $z \in R^n$ such that

$$S=\{y \in R^n | v(y) >_L z\}. \quad (1.10)$$

c) Any set of the form (1.9), with $v \in \mathcal{L}(R^n)$ and $z \in \overline{R}^n$ (instead of $v \in \mathcal{O}(R^n)$ and $z \in R^n$) will be called a generalized half-space (such sets have been called "half-spaces" in [9]). As has been observed in [9], if $z=(\xi_1 \ldots \xi_n)^T$ in (1.9), and $\xi_1=-\infty$ or $+\infty$, then (1.9) is the empty set or the whole space respectively; furthermore [9], if $n>2$, $\xi_1 \in R$ and $\xi_2=-\infty$ or $+\infty$, then, for any $v=(m_1 \ldots m_n)^T \in \mathcal{L}(R^n)$ with $m_1^T \neq 0$, (1.9) is the open half-space $\{y \in R^n | m_1^T y < \xi_1\}$ or the closed half-space $\{y \in R^n | m_1^T y \leq \xi_1\}$, respectively.

§2. Separation of two sets

Theorem 2.1. For any sets $G_1, G_2 \subset R^n$, the following statements are equivalent:

1°. $\operatorname{co} G_1 \cap \operatorname{co} G_2 = \emptyset$.

2°-4°. There exists $v \in \mathcal{L}(R^n)$, $v \in \mathcal{U}(R^n)$, $v \in \mathcal{O}(R^n)$, respectively, satisfying (1.6).

5°. There exists either a generalized half-space $S' \subset R^n$ such that

$$G_1 \subseteq S', \quad G_2 \subseteq R^n \setminus S', \tag{2.1}$$

or a generalized half-space $S'' \subseteq R^n$ such that

$$G_1 \subseteq R^n \setminus S'', \quad G_2 \subseteq S''. \tag{2.2}$$

Proof. $1^\circ \Rightarrow 4^\circ$. If 1° holds, then $0 \notin \operatorname{co} G_1 - \operatorname{co} G_2$, which is a convex set. Hence, by theorem 1.1, there exists $v \in \mathcal{O}(R^n)$ such that

$$v(y) <_L 0 \qquad (y \in \operatorname{co} G_1 - \operatorname{co} G_2), \tag{2.3}$$

which implies 4°.

The implications $4^\circ \Rightarrow 3^\circ \Rightarrow 2^\circ$ are obvious.

$2^\circ \Rightarrow 1^\circ$. Assume that $v \in \mathcal{L}(R^n)$ is as in 2°, but 1° does not hold, say $\sum_{i=1}^{k} \lambda_i y_i^1 = \sum_{j=1}^{m} \mu_j y_j^2$, where $y_i^1 \in G_1$, $y_j^2 \in G_2$, $\lambda_i, \mu_j \geq 0$ $(i=1,\dots,k; j=1,\dots,m)$, $\sum_{i=1}^{k} \lambda_i = \sum_{j=1}^{m} \mu_j = 1$. Then, by (1.6),

$$v(y_i^1) <_L v(y_j^2) \qquad (i=1,\dots,k; j=1,\dots,m),$$

whence

$$v\Big(\sum_{i=1}^{k} \lambda_i y_i^1\Big) = \sum_{i=1}^{k} \lambda_i v(y_i^1) \leq_L \max_{1 \leq i \leq k} v(y_i^1) <_L$$

$$<_L \min_{1 \leq j \leq m} v(y_j^2) \leq_L \sum_{j=1}^{m} \mu_j v(y_j^2) = v\Big(\sum_{j=1}^{m} \mu_j y_j^2\Big),$$

in contradiction with $\sum_{i=1}^{k} \lambda_i y_i^1 = \sum_{j=1}^{m} \mu_j y_j^2$.

$4^\circ \Rightarrow 5^\circ$. If $v \in \mathcal{O}(R^n)$ is as in 4°, then, by (1.6),

$$\sup_L v(G_1) \leq_L \inf_L v(G_2). \tag{2.4}$$

If $\sup_L v(G_1) \notin v(G_1)$, then, by (2.4), for each of the generalized half-spaces

$$S_1' = \{y \in R^n \mid v(y) <_L \sup_L v(G_1)\} \subseteq S_2' = \{y \in R^n \mid v(y) <_L \inf_L v(G_2)\}, \tag{2.5}$$

we have (2.1). Similarly, if $\inf_L v(G_2) \notin v(G_2)$, then, by (2.4), for each of the generalized half-spaces

$$S_1'' = \{y \in R^n \mid v(y) >_L \sup_L v(G_1)\} \supseteq S_2'' = \{y \in R^n \mid v(y) >_L \inf_L v(G_2)\}, \tag{2.6}$$

we have (2.2). Finally, if $\sup_L v(G_1) \in v(G_1)$, $\inf_L v(G_2) \in v(G_2)$, then, by (1.6), there holds

$$\sup_L v(G_1) <_L \inf_L v(G_2), \tag{2.7}$$

and hence, for the semi-spaces

$$S_2'=\{y\in R^n \mid v(y) <_L \inf_L v(G_2)\},\ S_1''=\{y\in R^n \mid v(y) >_L \sup_L v(G_1)\}, \tag{2.8}$$

we have (2.1) and (2.2), respectively.

$5^\circ \Rightarrow 2^\circ$. If S' is a generalized half-space of the form (1.9), satisfying (2.1), where $v\in \mathcal{L}(R^n)$ and $z\in \bar{R}^n$, then

$$v(y_1) <_L z \leqslant_L v(y_2) \qquad (y_1\in G_1,\ y_2\in G_2).$$

Finally, the case of a generalized half-space S'' satisfying (2.2) is similar. ∎

Remark 2.1. a) Let us also give the following alternative proof of the implication $2^\circ \Rightarrow 1^\circ$, which does not use the characterization of co G as the set of convex combinations of points of G: Assume 2° and let

$$H=\{y\in R^n \mid v(y) <_L v(y_2)\ \ (y_2\in G_2)\}, \tag{2.9}$$

$$K=\{y\in R^n \mid v(y_1) <_L v(y)\ \ (y_1\in \text{co } G_1)\}. \tag{2.10}$$

Then H,K are convex sets and, by (1.6), we have $G_1 \subseteq H$. Hence, co $G_1 \subseteq H$, that is, $G_2 \subseteq K$. Hence, co $G_2 \subseteq K$, which implies 1°. ∎

b) One cannot replace in 5° "generalized half-space" by "semi-space", as shown e.g. by the (convex) sets

$$G_1=\{y=(\eta_1\ \ \eta_2)^T\in R^2 \mid \eta_1 \leqslant 0\},\ G_2=R^2\setminus G_1, \tag{2.11}$$

or by the (convex) sets

$$G_1=\{y=(\eta_1\ \ \eta_2)^T\in R^2 \mid \eta_1 < 0\},\ G_2=R^2\setminus G_1. \tag{2.12}$$

Note also that, for (2.11), the identity operator $v=I$ satisfies (1.6) and $\sup_L v(G_1)=(0\ +\infty)^T=\inf_L v(G_2)$, while for (2.12), $v=I$ satisfies (1.6) and $\sup_L v(G_1)=(0\ -\infty)^T=\inf_L v(G_2)$.

c) In the particular case when $G_2=\{x_o\}$, the implication $1^\circ \Rightarrow 4^\circ$ of theorem 2.1 yields again theorem 1.1. Moreover, by remark 1.1 a), in this case there exists a semi-space S' (namely, $S'=S$ of (1.8)), satisfying (2.1) (with $G_2=\{x_o\}$); note that this also follows from the above proof of theorem 2.1, since for $G_2=\{x_o\}$ we have

$\inf_L v(G_2) \in v(G_2)$, so only the cases (2.5) or (2.8) can hold.

d) Since S', S'', $R^n \setminus S'$ and $R^n \setminus S''$ of (2.1) and (2.2) are cones, the implication $1^o \Rightarrow 4^o$ is a "cone separation theorem", in a strong sense (since, usually, in such a theorem, the complement of a cone which "separates" G_1 from G_2, need not be a convex set; see e.g. [7]).

e) <u>If G_1, $G_2 \subset R^n$ are convex sets, condition 1^o becomes</u>

1'. $G_1 \cap G_2 = \emptyset$.

In this case, the implication $1' \Rightarrow 3^o$ is equivalent to the well-known result (see e.g. [6], corollary 2 or [8], §17, theorem 1), according to which, for any two convex sets G_1, G_2 in a linear space F, with $G_1 \cap G_2 = \emptyset$, there exists a convex set H such that $F \setminus H$ is convex, $G_1 \subseteq H$ and $G_2 \subseteq F \setminus H$, namely,

$$H = \bigcap_{y_2 \in G_2} \{y_2 + S\}, \qquad (2.13)$$

where S is a certain semi-space at 0. Indeed, since for any semi-space S at 0, H of (2.13) and $F \setminus H$ are convex, and $G_2 \subseteq F \setminus H$ ([16], [8]), it remains to show that, for $F = R^n$, there exists $v \in \mathcal{U}(R^n)$ as in 3^o if and only if there exists a semi-space S at 0 such that H of (2.13) satisfies $G_1 \subseteq H$. Now, if $v \in \mathcal{U}(R^n)$ is as in 3^o, then for $S = \{y \in R^n \mid v(y) <_L 0\}$ we have, by (1.6), $G_1 - y_2 \subseteq S$ $(y_2 \in G_2)$, whence $G_1 = y_2 + (G_1 - y_2) \subseteq y_2 + S \subseteq H$. Conversely, if for $S = \{y \in R^n \mid v(y) <_L 0\}$, where $v \in \mathcal{U}(R^n)$, the set H of (2.13) satisfies $G_1 \subseteq H$, then

$$G_1 \subseteq \bigcap_{y_2 \in G_2} \{y_2 + y \mid y \in R^n,\ v(y) <_L 0\} = \bigcap_{y_2 \in G_2} \{y \in R^n \mid v(y) <_L v(y_2)\},$$

so (1.6) holds, which proves our assertion.

f) For any two sets $G_1, G_2 \subset R^n$, we have 1' if and only if

$$0 \notin G_1 - G_2 . \tag{2.14}$$

If (2.14) holds and $G_1 - G_2$ is convex, then, by theorem 1.1, there exists $v \in \mathcal{O}(R^n)$ such that

$$v(y) <_L 0 \qquad (y \in G_1 - G_2), \tag{2.15}$$

whence we obtain (1.6). Thus, in this case, *we also have* 1° *and* 5° *of theorem* 2.1.

§3. Separation of a set and the non-positive orthant

Now we shall consider the particular case in which $G_2 = -R^n_+$, where $R^n_+ = \{y \in R^n \mid y \geqslant 0\}$.

Theorem 3.1. For any set $G \subset R^n$ *such that* $G + R^n_+$ *is convex, the following statements are equivalent:*

1°. $0 \notin G + R^n_+$.

2°. $G \cap (-R^n_+) = \emptyset$.

3°. $(\mathrm{co}\ G) \cap (-R^n_+) = \emptyset$.

4°-6°. *There exists* $v \in \mathcal{L}(R^n)$, $v \in \mathcal{U}(R^n)$, $v \in \mathcal{O}(R^n)$, *respectively*, such that

$$v(g) >_L v(y) \qquad (g\in G,\ y\in -R^n_+). \qquad (3.1)$$

7^o-9^o. Same as 4^o, 5^o and 6^o respectively, with $v \geqslant_L 0$.

10^o-11^o. Same as 4^o and 5^o respectively, with $v \geqslant 0$.

Proof. The equivalences $1^o \Leftrightarrow \ldots \Leftrightarrow 6^o$ hold by remark 2.1 f) with $G_1=G$, $G_2=-R^n_+$, replacing v of (1.6) by -v.

$6^o \Rightarrow 9^o$. For $v\in\mathcal{O}(R^n)$ as in 6^o, let us consider the matrix representation (1.1) of v and let

$$i_o=\min \{i\in\{1,\ldots,n\} \mid (m_1\ldots m_i)^T g >_L (m_1\ldots m_i)^T y \quad (g\in G, z\in -R^n_+)\}. \quad (3.2)$$

Since (by (3.1)) the set on the right hand side contains i=n, we have $1\leqslant i_o\leqslant n$. We claim that $(m_1\ldots m_{i_o})^T \geqslant_L 0$, i.e., that

$$(m_{1j}\ldots m_{i_o j})^T \geqslant_L 0 \qquad (j=1,\ldots,n). \qquad (3.3)$$

Indeed, assume first that $i_o > 1$. Then, by (3.1) and (3.2), there exist $g_o\in G$ and $y_o\in -R^n_+$ such that

$$(m_1\ldots m_{i_o-1})^T g_o = (m_1\ldots m_{i_o-1})^T y_o. \qquad (3.4)$$

But, since $y_o+y\in -R^n_+$ $(y\in -R^n_+)$, we have, by (3.1),

$$(m_1\ldots m_{i_o-1})^T g_o \geqslant_L (m_1\ldots m_{i_o-1})^T (y_o+y) \qquad (y\in -R^n_+). \quad (3.5)$$

From (3.4) and (3.5) it follows that

$$(m_1\ldots m_{i_o-1})^T y \leqslant_L 0 \qquad (y\in -R^n_+), \qquad (3.6)$$

whence, by lemma 1.1,

$$(m_1\ldots m_{i_o-1})^T \geqslant_L 0, \qquad (3.7)$$

and thus, by (1.1),

$$(m_{1j}\ldots m_{i_o-1,j})^T \geqslant_L 0 \qquad (j=1,\ldots,n). \qquad (3.8)$$

If $(m_{1j}\ldots m_{i_o-1,j})^T >_L 0$, then $(m_{1j}\ldots m_{i_o j})^T >_L 0$, so (3.3) holds for this j. Assume now that

$$(m_1\ldots m_{i_o-1})^T e_j = (m_{1j}\ldots m_{i_o-1,j})^T = 0. \qquad (3.9)$$

Then, by (3.4) and (3.9),

$$(m_1\ldots m_{i_o-1})^T g_o = (m_1\ldots m_{i_o-1})^T (y_o-\lambda e_j) \qquad (\lambda\in R_+). \qquad (3.10)$$

On the other hand, since $y_o-\lambda e_j\in -R^n_+$ $(\lambda\in R_+)$, we have, by $g_o\in G$ and (3.2),

$$(m_1\dots m_{i_o})^T g_o >_L (m_1\dots m_{i_o})^T (y_o-\lambda e_j) \qquad (\lambda\in R_+). \quad (3.11)$$

From (3.11) and (3.10) it follows that

$$m_{i_o}^T g_o > m_{i_o}^T y_o - \lambda m_{i_o}^T e_j \qquad (\lambda\in R_+), \quad (3.12)$$

whence $m_{i_o j}=m_{i_o}^T e_j \geqslant 0$ (since otherwise, taking $\lambda\to+\infty$, we arrive at a contradiction with (3.12)). Hence, by (3.9), we obtain that (3.3) holds for this j, too.

Finally, assume that $i_o=1$. Then, by (3.2), for any $g_o\in G$ and all $j=1,\dots,n$ we have (3.12) (with $i_o=1$), whence, as above, it follows that $m_{1j}\geqslant 0$ $(j=1,\dots,n)$. Thus, $m_1^T \geqslant_L 0$, which proves the claim (3.3).

Next, we claim that there exists $\tilde{v}=(\tilde{m}_{ij})_{i,j=1}^n\in\mathcal{O}(R^n)$, $\tilde{v}\geqslant_L 0$, such that

$$\tilde{m}_{ij} = m_{ij} \qquad (i=1,\dots,i_o; j=1,\dots,n). \quad (3.13)$$

Indeed, we shall construct the $\tilde{m}_{ij}$'s by induction on i. By $v\in\mathcal{O}(R^n)$, $(m_{i1}\dots m_{in})^T$ $(i=1,\dots,i_o)$ are mutually orthogonal non-zero vectors; also, they satisfy (3.3). Assume now that for some k with $0\leqslant k\leqslant n-i_o-1$ we have constructed mutually orthogonal non-zero vectors $(\tilde{m}_{i1}\dots\tilde{m}_{in})^T$ $(i=1,\dots,i_o+k)$ satisfying (3.13) and

$$(\tilde{m}_{1j}\dots\tilde{m}_{i_o+k,j})^T \geqslant_L 0 \qquad (j=1,\dots,n). \quad (3.14)$$

Then, by the orthogonality assumption, we have

$$\operatorname{rank}(\tilde{m}_{ij})_{i=1,\dots,i_o+k;\ j=1,\dots,n} = i_o+k, \quad (3.15)$$

i.e., there exist indices $1\leqslant\pi(1)<\dots<\pi(i_o+k)\leqslant n$ such that

$$\det(\tilde{m}_{i,\pi(j)})_{i,j=1,\dots,i_o+k} \neq 0. \quad (3.16)$$

Choose any $\pi(i_o+k+1)\in\{1,\dots,n\}\setminus\{\pi(i)\}_{i=1}^{i_o+k}$. Then, by (3.16), there exist (unique) real numbers $\xi_1,\dots,\xi_{i_o+k}$, such that

$$\sum_{j=1}^{i_o+k}\tilde{m}_{i,\pi(j)}\xi_j = -\tilde{m}_{i,\pi(i_o+k+1)} \qquad (i=1,\dots,i_o+k). \quad (3.17)$$

Let us define

$$\tilde{m}_{i_o+k+1,\, \pi(j)} = \xi_j \qquad (j=1,\ldots,i_o+k), \quad (3.18)$$

$$\tilde{m}_{i_o+k+1,\, \pi(i_o+k+1)} = 1, \quad (3.19)$$

$$\tilde{m}_{i_o+k+1,\, j} = 0 \qquad (j\in\{1,\ldots,n\} \setminus \{\pi(i)\}_{i=1}^{i_o+k+1}). \quad (3.20)$$

Then, by the induction assumption and by (3.17)-(3.19), the vectors $(\tilde{m}_{i1}\ldots\tilde{m}_{in})^T$ $(i=1,\ldots,i_o+k+1)$ are mutually orthogonal; also, by (3.14), (3.16), (3.19) and (3.20), we have

$$(\tilde{m}_{1j}\ldots\tilde{m}_{i_o+k+1,j})^T \geqslant_L 0 \qquad (j=1,\ldots,n), \quad (3.21)$$

which proves the claim on the existence of $\tilde{v}\in\mathcal{O}(R^n)$, $\tilde{v} \geqslant_L 0$, satisfying (3.13).

Finally, from (3.2) it follows that

$$(m_1\ldots m_{i_o})^T g >_L (m_1\ldots m_{i_o})^T y \qquad (g\in G, y\in -R^n_+), \quad (3.22)$$

and hence, by (3.13),

$$\tilde{v}(g) >_L \tilde{v}(y) \qquad (g\in G,\ y\in -R^n_+). \quad (3.23)$$

The converse implication $9^o \Rightarrow 6^o$ is obvious.

$8^o \Rightarrow 11^o$. If $v\in\mathcal{U}(R^n)$ is as in 8^o, then, by $v_o \geqslant_L 0$ and [10], corollary 2.1, there exist a unitary lower triangular matrix $\ell\in\mathcal{U}(R^n)$ and a matrix $p\in\mathcal{L}(R^n)$, $p\geqslant 0$, such that $v=\ell p$. Then $p=\ell^{-1}v\in \mathcal{U}(R^n)$ and, since ℓ^{-1} is also unitary and lower triangular, we have, by (3.1) and lemma 1.2,

$$p(x)=\ell^{-1}(v(x)) \geqslant_L 0 \qquad (x\in G + R^n_+). \quad (3.24)$$

But, by (3.1) and since $\ell^{-1}\in\mathcal{U}(R^n)$, we have $p(x)=\ell^{-1}(v(x))\neq 0$ $(x\in G + R^n_+)$ and hence, by (3.24),

$$p(x) >_L 0 \qquad (x\in G + R^n_+). \quad (3.25)$$

Thus, we may take p as the operator v required in 11^o.

Finally, the implications $11^o \Rightarrow 8^o \Rightarrow 7^o \Rightarrow 4^o$ and $11^o \Rightarrow 10^o \Rightarrow 7^o$ are obvious. ∎

Remark 3.1. a) One can give a direct proof of the implication $5^o \Rightarrow 8^o$, slightly simpler than the above proof of $6^o \Rightarrow 9^o$. Namely, after proving (3.3) as above, note that, by $v\in\mathcal{U}(R^n)$, we have rank $(m_1\ldots m_{i_o})^T = i_o$, i.e., there exist indices $1\leqslant\pi(1) < \ldots < \pi(i_o)\leqslant n$

such that

$$\det (m_{i,\pi(j)})_{i,j=1,\dots,i_o} \neq 0. \tag{3.26}$$

Let

$$\tilde{v} = (m_1 \dots m_{i_o} \quad e_{\pi(i_o+1)} \dots e_{\pi(n)})^T \in \mathcal{L}(R^n), \tag{3.27}$$

where $\{\pi(i_o+1),\dots,\pi(n)\}=\{1,\dots,n\}\setminus\{\pi(1),\dots,\pi(i_o)\}$. Then $\det \tilde{v}\neq 0$, so $\tilde{v}\in\mathcal{U}(R^n)$. Also, by (3.27) and (3.3), we have $\tilde{v} \geqslant_L 0$. Finally, from (3.2) there follows (3.22), whence, by (3.27), we obtain (3.23). ∎

One can also give a direct proof of the implication $4^o \Rightarrow 7^o$, which is even simpler. Namely, after proving (3.3) as above, it is enough to take

$$\tilde{v} = (m_1 \dots m_{i_o} \quad 0 \dots 0)^T \in \mathcal{L}(R^n); \tag{3.28}$$

indeed, $\tilde{v} \geqslant_L 0$ and, by (3.22) and (3.28), we have (3.23). ∎

b) One cannot replace, in 11^o, $v\in\mathcal{U}(R^n)$, $v\geqslant 0$, by the stronger properties $v\in\mathcal{O}(R^n)$, $v\geqslant 0$, since it is easy to see that the only nonnegative orthogonal matrices are the permutation matrices.

Let us consider now the particular case $G_2=-R^n_+\setminus\{0\}$ of the situation of theorem 2.1.

<u>Theorem 3.2. For any set</u> $G\subset R^n$ <u>such that</u> $G + R^n_+$ <u>is convex, the following statements are equivalent:</u>

1^o. $0\notin G + (R^n_+\setminus\{0\})$.

2^o. $G\cap(-R^n_+\setminus\{0\})=\emptyset$.

3^o. $(\mathrm{co}\ G)\cap(-R^n_+\setminus\{0\})=\emptyset$.

4^o-6^o. <u>There exists</u> $v\in\mathcal{L}(R^n)$, $v\in\mathcal{U}(R^n)$, $v\in\mathcal{O}(R^n)$, <u>respectively, such that</u>

$$v(g) >_L v(y) \qquad (g\in G,\ y\in -R^n_+\setminus\{0\}). \tag{3.29}$$

7^o-9^o. <u>Same as</u> 4^o, 5^o <u>and</u> 6^o <u>respectively, with</u> $v \geqslant_L 0$.

10^o-11^o. <u>Same as</u> 4^o <u>and</u> 5^o <u>respectively, with</u> $v\geqslant 0$.

<u>Proof</u>. The equivalences $1^o \Leftrightarrow \dots \Leftrightarrow 6^o$ hold by remark 2.1 f) with $G_1=G$, $G_2=-R^n_+\setminus\{0\}$, observing that $-G_2=R^n_+\setminus\{0\}$ and replacing v of (1.6) by $-v$.

$6^\circ \Rightarrow 9^\circ$. If $0 \notin \mathrm{co}\, G$, then, by 3°, we have $(\mathrm{co}\, G) \cap (-R^n_+) = \emptyset$, so we can apply theorem 3.1, implication $6^\circ \Rightarrow 9^\circ$. On the other hand, if $0 \in \mathrm{co}\, G$, then, by theorem 2.1 (see also (2.3)), there exists $v \in \mathcal{O}(R^n)$ satisfying

$$v(g) >_L v(y) \qquad (g \in \mathrm{co}\, G, y \in -R^n_+ \setminus \{0\}), \quad (3.30)$$

whence (3.29). In particular, since $0 \in \mathrm{co}\, G$, we have

$$0 >_L v(y) \qquad (y \in -R^n_+ \setminus \{0\}), \quad (3.31)$$

whence $0 \geqslant_L v(y)$ $(y \in -R^n_+)$, and hence, by lemma 1.1, $v \geqslant_L 0$.

The converse implication $9^\circ \Rightarrow 6^\circ$ is obvious.

The proof of the implication $8^\circ \Rightarrow 11^\circ$ is similar to that of theorem 3.1, implication $8^\circ \Rightarrow 11^\circ$, replacing, in (3.24) and (3.25), $G + R^n_+$ by $G + (R^n_+ \setminus \{0\})$. Finally, the implications $11^\circ \Rightarrow 8^\circ \Rightarrow 7^\circ \Rightarrow 4^\circ$ and $11^\circ \Rightarrow 10^\circ \Rightarrow 7^\circ$ are obvious. ▨

§4. Applications

As an application of theorem 3.2, let us give here the following characterization of the elements of INF G:

Theorem 4.1. *Let $G \subset R^n$ be such that $G + R^n_+$ is convex, and let $x \in R^n$. The following statements are equivalent:*

1°. $x \in \mathrm{INF}\, G$.

2°. $x \in \overline{G}$ *and there exists* $v \in \mathcal{U}(R^n)$, $v \geqslant 0$, *such that*

$$v(g) >_L v(y) \qquad (g \in G,\ y \in R^n,\ y < x). \quad (4.1)$$

Proof. $1^\circ \Rightarrow 2^\circ$. We may assume, without loss of generality (replacing G by G-x) that x=0. Now, if 1° holds for x=0, then $G \cap (-R^n_+ \setminus \{0\}) = \emptyset$, whence, by theorem 3.2, there exists $v \in \mathcal{U}(R^n)$, $v \geqslant 0$, satisfying (4.1) with x=0.

$2^\circ \Rightarrow 1^\circ$. If $v \in \mathcal{L}(R^n)$ satisfies (4.1), then $G \cap \{y \in R^n \mid y < x\} = \emptyset$, and hence, by $x \in \overline{G}$, we have 1°. ▨

Remark 4.1. a) Using theorem 3.2, one can also give other equivalent conditions, which we omit.

b) One cannot replace (4.1) of 2° by

$$v(y) >_L v(x) \qquad (y \in G \setminus \{x\}), \tag{4.2}$$

as shown e.g. by the (convex) set

$$G = \{(\eta_1\ \eta_2)^T \in R^2 \mid \eta_1 \geq 0,\ \eta_2 \geq 0,\ \eta_1 + \eta_2 \geq 1\} \subset R^2 \tag{4.3}$$

and the element $x = \left(\frac{1}{2}\ \ \frac{1}{2}\right)^T \in G \cap \mathrm{INF}\ G$. Indeed, $e_1 = (1\ 0)^T \in G \setminus \{x\}$, $e_2 = (0\ \ 1)^T \in G \setminus \{x\}$, so if (4.2) holds for some $v \in \mathcal{L}(R^n)$, then

$$v(x) = v\left(\frac{e_1 + e_2}{2}\right) = \frac{1}{2}v(e_1) + \frac{1}{2}v(e_2) >_L v(x),$$

which is impossible.

Finally, let us give an application of theorem 1.1.

<u>Theorem 4.2.</u> <u>Let</u> G <u>be a convex subset of</u> R^n <u>and let</u> $f = (f_1, \ldots, f_n)$ <u>be a function from a finite interval</u> $[a,b] \subset R$ <u>into</u> G, <u>with the components</u> $f_1, \ldots, f_n$ <u>integrable (Lebesgue). Then</u>

$$x_o = \frac{1}{b-a} \int_a^b f(t)\,dt \in G. \tag{4.4}$$

<u>Proof</u>. Assume that $x_o \notin G$. Then, by theorem 1.1, there exists $v \in \mathcal{O}(R^n)$ satisfying (1.7), whence, in particular,

$$v(f(t)) <_L v(x_o) \qquad (t \in [a,b]). \tag{4.5}$$

Let us consider the matrix representation (1.1) of v and let

$$A_i = \{t \in [a,b] \mid m_j^T f(t) = m_j^T x_o\ (j=1,\ldots,i-1),\ m_i^T f(t) < m_i^T x_o\} \quad (i=1,\ldots,n). \tag{4.6}$$

The sets A_i are measurable (Lebesgue) and $A_i \cap A_j = \emptyset$ $(i \neq j)$, $\bigcup_{i=1}^{n} A_i = [a,b]$ (by (4.5)), whence $\sum_{i=1}^{n} \mu(A_i) = b-a$, where μ denotes the Lebesgue measure. Let

$$i_o = \min\ \{i \in \{1,\ldots,n\} \mid \mu(A_i) > 0\}. \tag{4.7}$$

Then, by (4.5)-(4.7), we have

$$m_i^T f(t) = m_i^T x_o \qquad \mu\text{-a.e. in } [a,b] \qquad (i=1,\ldots,i_o-1), \tag{4.8}$$

$$m_{i_o}^T f(t) \leq m_{i_o}^T x_o \qquad \mu\text{-a.e. in } [a,b], \tag{4.9}$$

the inequality (4.9) being strict on some set of positive measure. Hence, by integration of (4.9), we obtain

$$m_{i_o}^T \int_a^b f(t)\,dt < m_{i_o}^T x_o (b-a),$$

whence, by the definition (4.4) of x_o,

$$m_{i_o}^T x_o < m_{i_o}^T x_o ,$$

which is impossible. ∎

Remark 4.2. Theorem 4.2 improves a result stated e.g. in [5], pp.200-201, where the conclusion is $x_o \in \bar{G}$, the closure of G.

REFERENCES

[1] A.Bacopoulos, G.Godini and I.Singer, On best approximation in vector-valued norms, in Fourier analysis and approximation theory, Coll.Math.Soc.J.Bolyai 19 (G.Alexits and P.Turán, Eds.), North Holland, Amsterdam-Oxford-New York, 1978, pp.89-100.

[2] A.Bacopoulos, G.Godini and I.Singer, Infima of sets in the plane and applications to vectorial optimization, Rev. roum. math. pures et appl. 23:343-360 (1978).

[3] R.Deumlich, K.-H.Elster and R.Nehse, Recent results on separation of convex sets, Math. Operationsforsch. Stat. Ser.Optim. 9:273-296(1978).

[4] K.-H.Elster and R.Nehse, Separation of two convex sets by operators, Comment. Math. Univ. Carol. 19:191-206(1978).

[5] W.H.Fleming and R.W.Rishel, Deterministic and stochastic optimal control, Springer-Verlag, Berlin-Heidelberg-New-York, 1975.

[6] P.C.Hammer, Maximal convex sets, Duke Math.J.22:103-106(1955).

[7] M.I.Henig, A cone separation theorem, J.Optim.Theory Appl.36:451-455(1982).

[8] G.Köthe, Topologische lineare Räume. I, Springer-Verlag, Berlin-Göttingen-Heidelberg, 1960.

[9] J.-E.Martínez-Legaz, Exact quasiconvex conjugation, Z.Oper.Res. 27:257-266(1983).

[10] J.-E.Martínez-Legaz, Lexicographical order, inequality systems and optimization, in Proc.11th IFIP Confer. on "System modelling

and optimization", Lecture Notes in Control and Information Sciences 59 (P.Thoft-Christensen, Ed.), Springer-Verlag, Berlin-Heidelberg-New-York-Tokyo, 1984, pp.203-212.

[11] J.-E.Martínez-Legaz, Lexicographical order and duality in multiobjective programming, European J.Oper.Res., to appear.

[12] J.-E.Martínez-Legaz and I.Singer, Surrogate duality in vector optimization, to appear. Preprint INCREST 73/1985.

[13] T.S.Motzkin, Linear inequalities (mimeographed lecture notes), Univ.of California, Los Angeles, 1951.

[14] R.Nehse, Separation of two sets in a product space. Math.Nachr. 97:179-187(1980).

[15] I.Singer, Generalized convexity, functional hulls and applications to conjugate duality in optimization, in Selected topics in operations resarch and mathematical economics, Lecture Notes in Econ. and Math. Systems 226 (G.Hammer and D.Pallaschke,Eds.), Springer-Verlag,Berlin-Heidelberg-New York-Tokyo, 1984, pp.49-79.

[16] J.-E.Martínez-Legaz and I.Singer, The structure of hemi-spaces (in preparation).

SURROGATE DUALITY FOR VECTOR OPTIMIZATION

Juan-Enrique Martínez-Legaz and Ivan Singer

Department of Functional Equations, University of Barcelona, Gran Vía, 585, Barcelona 08007, Spain, and Department of Mathematics, INCREST, Bd. Păcii 220, Bucharest 79622, Romania

ABSTRACT

Using our theorems (of [12]) on separation of convex sets by linear operators, in the sense of the lexicographical order on R^n, we prove some theorems of surrogate duality for vector optimization problems with convex constraints (but no regularity assumption), where the surrogate constraint sets are generalized half-spaces and the surrogate multipliers are linear operators, or isomorphisms, or isometries. In the case of inequality constraints, we prove that the surrogate multipliers can be taken lexicographically non-negative isometries or non-negative (in the usual order) linear isomorphisms.

0. INTRODUCTION

We recall that, given a subset G of R^n (assumed non-empty, throughout the sequel), called "constraint set", a functional $h:R^n \to \overline{R}=[-\infty,+\infty]$, called "objective functional", and the primal scalar optimization problem

$$(P) \qquad \alpha=\inf h(G), \tag{0.1}$$

the set of "solutions" of (P) is defined by

$$\mathcal{S}_G(h)=\{g_o\in G \mid h(g_o)\leq h(g) \quad (g\in G)\}=$$

$$=\{g_o\in G \mid G\cap A_{h(g_o)}(h)=\emptyset\}, \tag{0.2}$$

where $\emptyset$ is the empty set and $A_{h(g_o)}(h)$ is the level set

$$A_{h(g_o)}(h)=\{y\in R^n \mid h(y)<h(y_o)\}. \tag{0.3}$$

Furthermore, a "surrogate dual problem" to (P), in the sense of [14] (see also [19]), is any optimization problem of the form

$$\text{(Q)} \quad \beta=\sup \lambda((R^n)^*); \; \lambda(\Psi)=\inf h(\Delta_{G,\Psi}) \quad (\Psi\in(R^n)^*), \tag{0.4}$$

where $(R^n)^*$ denotes the family of all linear functionals $\Psi:R^n\to R$ and where $\Delta_{G,\Psi}\subseteq R^n$ $(\Psi\in(R^n)^*)$; the sets $\Delta_{G,\Psi}$ are called "surrogate constraint sets"; for the classical particular cases, see e.g. [3], [7], [4] (see also [2], [8] and the references therein). Due to the second equality in (0.2), a suitable tool for obtaining characterizations of the solutions of (P) consists (see [7], [15]) in separation of the constraint set G from the level set $A_{h(g_o)}(h)$, by linear functionals $\Psi\in(R^n)^*$.

Let us consider now the case when the above h is replaced by an "objective mapping" $h:R^n\to\overline{R}^m$ and $\mathcal{S}_G(h)$ is replaced by the set of all "Pareto-solutions" of the corresponding primal vector optimization problem (P), i.e., by

$$\mathcal{V}_G(h)=\{g_o\in G \mid h(g_o)\not> h(g) \quad (g\in G)\}=$$

$$=\{g_o\in G \mid G\cap A_{h(g_o)}(h)=\emptyset\}, \tag{0.5}$$

where $\geq$ is the natural partial order of $\overline{R}^m$, $\not>$ denotes the negation of $>$ (where $x>z$ means that $x\geq z$, $x\neq z$), and $A_{h(g_o)}(h)$ is still the "level set" (0.3), with $<$ in the sense of the natural partial order on $\bar{R}^m$. Again, the second equality in (0.5) suggests that a suitable tool for obtaining characterizations of the Pareto-solutions of (P), of "surrogate duality" type, should be the separation of the constraint set G from the level set $A_{h(g_o)}(h)$; generalizing the terminology of [14]-[16], this may be called "vector optimization by level set methods".

Since the usual separation theorems of two convex sets in R^n, and the various known extensions of these theorems to separation by linear operators, in the usual partial order of R^n, require rather strong assumptions, the duality theorems for (scalar or vector) optimization, obtained from them, require some "regularity assumptions". Therefore, a theorem of a new type (given in [9]), on the separation of an arbitrary convex set in R^n and any outside point, by linear operators, in the sense of the lexicographical order of R^n (instead of the usual order of R^n), has been used in [11], to obtain theorems of generalized "Lagrangian" duality for vector optimization.

In the present paper, we shall develop a theory of surrogate duality for vector optimization, based on the theorems of [12] on the separation of two subsets of R^n by linear operators $v:R^n \to R^n$, in the sense of the lexicographical order of R^n. In §§1-5, we shall consider abstractly given subsets G of R^n and we shall replace the surrogate constraint sets $\Delta_{G,\Psi}$ mentioned above by suitable sets $\Delta_{G,v}$, which correspond to those given, for scalar optimization, in [16], formula (2.5). Besides characterizations of the Pareto-solutions of (P) in terms of separation and surrogate constraint sets (§1), we shall also give characterizations of them in terms of surrogate subdifferentials (§2), surrogate Lagrangians (§3), and surrogate Lagrange multipliers (i.e., solutions of a surrogate dual problem of van Slyke-Wets [21] -Jahn [6] type; see §4), which we shall introduce here, as well as characterizations of pairs of primal-dual solutions in terms of the saddle-points (in the sense of [20] and [13]) of our surrogate Lagrangians. In §6, we shall consider subsets G of R^n given by (a finite number of) inequality constraints, i.e., of the form

$$G=\{g\in R^n \mid u(g)\leqslant 0\}=u^{-1}(-R^k_+), \tag{0.6}$$

where $u:R^n \to R^k$ and $R^k_+=\{x\in R^k \mid x\geqslant 0\}$; thus, in this case, by (0.6) and (0.5), for any $h:R^n\to\overline{R}^m$ we have

$$\mathcal{V}_G(h)=\{g_o\in G \mid u(A_{h(g_o)}(h))\cap(-R^h_+)=\emptyset\}. \tag{0.7}$$

In this case, we shall introduce other surrogate constraint sets $\Lambda_{u,v}\subseteq R^n$ (where $v:R^k\to R^k$ are linear operators), corresponding to those given, for scalar optimization, by Luenberger [7] and we shall give for them some results corresponding to those of §§1-5; moreover, we shall show that it is sufficient to consider non-negative linear operators v.

Note that the results of the present paper are new even in the particular case of scalar optimization, since even in that case we use linear operators instead of linear functionals.

Let us recall now some notions, notations and results, which we shall use in the sequel.

The elements of $\overline{R}^n$ will be considered column vectors and the superscript T will mean transpose. We recall that $x=(\xi_1\ldots\xi_n)^T\in\overline{R}^n$ is said to be "lexicographically less than" $y=(\eta_1\ldots\eta_n)^T\in\overline{R}^n$, in symbols, $x <_L y$, if $x\neq y$ and if for $k=\min\{i\in\{1,\ldots,n\} \mid \xi_i\neq\eta_i\}$ we have $\zeta_k < \eta_k$. We write $x \leq_L y$ if $x <_L y$ or $x=y$; the notations $y >_L x$, respectively $y \geq_L x$, will be also used. Furthermore, we shall denote by C_L the "lexicographical non-negative orthant" of R^n, i.e.,

$$C_L = \{y \in R^n \mid y \geq_L 0\}. \tag{0.8}$$

We shall denote by $\mathcal{L}(R^n)$, $\mathcal{U}(R^n)$ and $\mathcal{O}(R^n)$, the families of all linear operators, all isomorphisms, and all linear isometries $v: R^n \to R^n$, respectively. We shall consider two orderings of $\mathcal{L}(R^n)$ (hence also of $\mathcal{U}(R^n)$, $\mathcal{O}(R^n)$), namely, the usual order relation $\geq$, in which

$$v \geq 0 \Leftrightarrow v(x) \geq 0 \qquad (x \in R^n,\ x \geq 0), \tag{0.9}$$

and the lexicographical order in the sense of [10], in which

$$v \geq_L 0 \Leftrightarrow v(x) \geq_L 0 \qquad (x \in R^n,\ x \geq 0). \tag{0.10}$$

We recall that any subset S of R^n of the form

$$S = \{y \in R^n \mid v(y) <_L z\}, \tag{0.11}$$

where $v \in \mathcal{L}(R^n)$ and $z \in \overline{R}^n$, is called [12] a "generalized half-space" (such sets have been called "half-spaces" in [9]); since $v \in \mathcal{L}(R^n)$ if and only if $-v \in \mathcal{L}(R^n)$, (0.11) is equivalent to

$$S = \{y \in R^n \mid v'(y) >_L z'\}, \tag{0.12}$$

where $v' \in \mathcal{L}(R^n)$ and $z' \in \overline{R}^n$. As has been observed in [9] and [12], the generalized half-spaces (0.11) with $z = (\zeta_1 \ldots \zeta_n) \in \overline{R}^n$ include, among others, the empty set (for $\zeta_1 = -\infty$), the whole space (for $\zeta_1 = +\infty$), the open half-spaces (for $\zeta_1 \in R$, $\zeta_2 = -\infty$), the (usual) -spaces (for $\zeta_1 \in R$, $\zeta_2 = +\infty$) and the semi-spaces. We recall that a subset S of R^n is called [5] a "semi-space at x_0", where $x_0 \in R^n$, if S is a maximal convex set such that $x_o \notin S$; by [12], remark 1.2 a), a set $S \subset R^n$ is a semi-space (at some $x_o \in R^n$) if and only if there exist $v \in \mathcal{O}(R^n)$ (or $v \in \mathcal{U}(R^n)$) and $z \in R^n$ such that we have (0.11) (or, equivalently, (0.12)). In §§1-5, the surrogate constraint sets $\Lambda_{G,v}$ will be generalized half-spaces or complements of generalized half-spaces.

For two subsets G_1, G_2 of R^n, we shall say that an operator $v \in \mathcal{L}(R^n)$ "separates G_1 from G_2" (in the sense of the lexicographical order of R^n), if

$$v(g_1) <_L v(g_2) \qquad (g_1 \in G_1,\ g_2 \in G_2); \quad (0.13)$$

clearly, this happens if and only if $-v$ separates G_2 from G_1.

Let us recall some parts of the two main separation theorems of [12], which we shall use in the sequel.

Theorem 0.1 (part of [12], theorem 2.1). For any convex sets $G_1, G_2 \subset R^n$, the following statements are equivalent:

1^o. $G_1 \cap G_2 = \emptyset$.

2^o-4^o. There exists $v \in \mathcal{L}(R^n)$, $v \in \mathcal{U}(R^n)$, $v \in \mathcal{O}(R^n)$, respectively, satisfying (0.13).

Theorem 0.2 (part of [12], theorem 3.1). For any set $\Omega \subset R^k$ such that $\Omega + R^k_+$ is convex, the following statements are equivalent:

1^o. $\Omega \cap (-R^k_+) = \emptyset$.

2^o-4^o. There exists $v \in \mathcal{L}(R^k)$, $v \in \mathcal{U}(R^k)$, $v \in \mathcal{O}(R^k)$, respectively, such that

$$v(\omega) >_L v(z) \qquad (\omega \in \Omega,\ z \in -R^k_+). \qquad (0.14)$$

5^o-7^o. Some as 2^o, 3^o and 4^o respectively, with $v \geqslant_L 0$.

8^o-9^o. Same as 2^o and 3^o respectively, with $v \geqslant 0$.

We recall [12] that the sets G_1, G_2 and Ω in theorems 0.1 and 0.2 are not assumed to be $\neq\emptyset$, but, instead, the convention is made that if one of them is empty, then (0.13), respectively, (0.14), are considered to hold (vacuously).

We shall use the "upper addition" $\dot{+}$ on $\overline{R}^m$ defined componentwise (we recall that the usual upper addition $\dot{+}$ on $\overline{R}$ is defined by $a\dot{+}b=+\infty$ if $a=-b=\pm\infty$ and $a\dot{+}b=a+b$ for the other $a,b\in\overline{R}$).

1. SURROGATE CONSTRAINT SETS. SEPARATION

Proposition 1.1. Let $G\subset R^n$, $v\in\mathcal{L}(R^n)$ and let

$$\Delta_{G,v}=v^{-1}(v(G)+C_L)=\{y\in R^n|v(y)\in v(G)+C_L\}=$$

$$=\{y\in R^n|\ g\in G,\ v(y)\geqslant_L v(g)\}. \tag{1.1}$$

a) We have

$$G\subseteq\Delta_{G,v}. \tag{1.2}$$

b) We have

$$\Delta_{G,v}=\{y\in R^n|v(y)\rho\inf_L v(G)\}, \tag{1.3}$$

where $\inf_L v(G)$ denotes the infimum of $v(G)$ for the lexicographical order of R^n and where

$$\rho=\begin{cases}\geqslant_L & \text{if } \inf_L v(G)\in v(G)\\ >_L & \text{if } \inf_L v(G)\notin v(G),\end{cases} \tag{1.4}$$

so $\Delta_{G,v}$ is either a generalized half-space, or the complement of a generalized half-space.

Proof. a) If $g\in G$, then for $y=g\in R^n$ we have $v(y)\geqslant_L v(g)$, and thus, by (1.1), $g\in\Delta_{G,v}$.

b) Let $y\in\Delta_{G,v}$. Then, by (1.1), there exists $g\in G$ such that $v(y)\geqslant_L v(g)\geqslant_L \inf v(G)$. Moreover, if $\inf_L v(G)\notin v(G)$, then $v(y)\geqslant_L v(g)>_L \inf v(G)$. Thus,

$v(y)\rho \inf_L v(G)$.

Conversely, let $y\in R^n$, $v(y)\rho \inf_L v(G)$. If $\inf_L v(G)\in v(G)$, say, $\inf_L v(G)=v(g)$, where $g\in G$, then, by (1.4), $v(y) \geqslant_L v(g)$ and hence, by (1.1), $y\in\Delta_{G,v}$. If $\inf_L v(G)\notin v(G)$, then, by (1.4), $v(y) >_L \inf_L v(G)$, and hence there exists $g\in G$ such that $v(y) \geqslant_L v(g)$. Then, by (1.1), $y\in\Delta_{G,v}$.

Remark 1.1. a) By (1.3), (1.4), the sets $\Delta_{G,v}$ correspond to those given, for scalar optimization, in [16], formula (2.5) (see also [17], formula (3.78)).

b) One can also write (1.1) in the form

$$\Delta_{G,v}=\bigcup_{g\in G}\mathcal{H}_g(v), \qquad (1.5)$$

where

$$\mathcal{H}_g(v)=\{y\in R^n \mid v(y) >_L v(g)\} \qquad (g\in G). \qquad (1.6)$$

Then $\{\mathcal{H}_g(v)\}_{g\in G}$ is a family of "parallel" complements of generalized half-spaces, totally ordered by inclusion. Indeed, if $g,g'\in G$, then either $v(g) \geqslant_L v(g')$, or $v(g') \geqslant_L v(g)$. But, if $v(g) \geqslant_L v(g')$, then $\mathcal{H}_g(v)\subseteq \mathcal{H}_{g'}(v)$ (since if $y\in\mathcal{H}_g(v)$, then $v(y) \geqslant_L v(g) \geqslant_L v(g')$, so $y\in\mathcal{H}_{g'}(v)$). On the other hand, if $v(g') \geqslant_L v(g)$, then $\mathcal{H}_{g'}(v)\subseteq\mathcal{H}_g(v)$.

c) If $v\in\mathcal{U}(R^n)$, then, by (1.1), we have

$$\Delta_{G,v}=v^{-1}(v(G))+v^{-1}(C_L)=G+v^{-1}(C_L), \qquad (1.7)$$

where $v^{-1}(C_L)=\{y\in R^n \mid v(y) \geqslant_L 0\}$ is the complement of a semi-space at 0.

Theorem 1.1. Let G be a convex subset of R^n, let $g_o\in G$ and let $h:R^n\to\overline{R}^m$ be such that $A_{h(g_o)}$ is convex. The following statements are equivalent:

1^o. $g_o\in\mathcal{V}_G(h)$.

2^o-4^o. <u>There exists</u> $v\in\mathcal{L}(R^n)$, $v\in\mathcal{U}(R^n)$, $v\in\mathcal{O}(R^n)$, <u>respectively, such that</u>

$$v(y) <_L v(g) \qquad (y\in A_{h(g_o)}(h),\ g\in G). \quad (1.8)$$

5^o-7^o. <u>There exists</u> $v\in\mathcal{L}(R^n)$, $v\in\mathcal{U}(R^n)$, $v\in\mathcal{O}(R^n)$, <u>respectively, such that</u>

$$g_o\in\mathcal{V}_{\Delta_{G,v}}(h), \quad (1.9)$$

<u>where</u> $\Delta_{G,v}\subseteq R^n$ <u>is the set defined by</u> (1.1).

<u>Proof</u>. Since G and $A_{h(g_o)}(h)$ are convex, the equivalences $1^o \Leftrightarrow \ldots \Leftrightarrow 4^o$ follow from (0.5) and theorem 0.1.

$4^o \Rightarrow 7^o$. If $v\in\mathcal{O}(R^n)$ of 4^o does not satisfy (1.9), then there exists $y'\in\Delta_{G,v}\cap A_{h(g_o)}(h)$. Then, by $y'\in\Delta_{G,v}$ and (1.1), there exists $g'\in G$ such that $v(y') \geq_L v(g')$, which violates (1.8).

The implications $7^o\Rightarrow 6^o\Rightarrow 5^o$ are obvious.

$5^o\Rightarrow 2^o$. If, for $v\in\mathcal{L}(R^n)$ as in 5^o, there exist $y'\in A_{h(g_o)}(h)$ and $g'\in G$ such that $v(y') \geq_L v(g')$, then, by (1.1), we have $y'\in\Delta_{G,v}$. Thus, $\Delta_{G,v}\cap A_{h(g_o)}(h)\neq\emptyset$, whence, by (0.5), $g_o\notin\mathcal{V}_{\Delta_{G,v}}(h)$.

<u>Remark 1.2</u>. a) By the equivalence $1^o\Leftrightarrow 5^o$, we shall call $\Delta_{G,v}\subseteq R^n$ a <u>surrogate constraint set</u>.

b) The convexity of G and $A_{h(g_o)}(h)$ is needed only in the proofs of $1^o\Rightarrow n^o$ $(n=2,\ldots,7)$.

c) The fact that $g_o\in G$ and (1.9) imply 1^o, means that

$$G\cap\mathcal{V}_{\Delta_{G,v}}(h)\subseteq\mathcal{V}_G(h), \quad (1.10)$$

and can be also proved directly, as follows: $g_o\in G\subseteq\Delta_{G,v}$ and (1.9) imply $\Delta_{G,v}\cap A_{h(g_o)}(h)=\emptyset$, whence, by (1.2), $G\cap A_{h(g_o)}(h)=\emptyset$, so $g_o\in\mathcal{V}_G(h)$. Thus, as shown by the above proofs, the implications $n^o \Rightarrow 1^o$ $(n=2,\ldots,7)$ remain valid for $\Delta_{G,v}$ replaced by any set $\Delta \subseteq R^n$ such that $G \subseteq \Delta$.

Corollary 1.1. *Let G be a convex subset of* R^n *and let* $h: R^n \to \overline{R}^m$ *be such that* $A_{h(g_o)}(h)$ *is convex for each* $g_o \in G$. *Then*

$$\mathcal{V}_G(h) = G \cap \bigcup_{v \in \mathcal{L}(R^n)} \mathcal{V}_{\Delta_{G,v}}(h) =$$

$$= G \cap \bigcup_{v \in \mathcal{U}(R^n)} \mathcal{V}_{\Delta_{G,v}}(h). \qquad (1.11)$$

Corollary 1.2. *Let G be a convex subset of* R^n, *let* $h_1,\ldots,h_m: R^n \to \overline{R}$ *be explicitly quasi-convex functionals, and let*

$$h(y) = (h_1(y),\ldots,h_m(y)) \in \overline{R}^m \qquad (y \in R^n). \quad (1.12)$$

Then we have (1.11).

Proof. The explicit quasi-convexity of the h_i's implies the convexity of $A_{h(g_o)}(h)$, so corollary 1.1 applies.

In the next corollary, we shall use the standard identification of each $v \in \mathcal{L}(R^n)$ with its matrix with respect to the unit vector basis $\{e_j\}_{j=1}^n$ of R^n, i.e., we shall write

$$v = (m_{ij})_{i,j=1}^n = (m_1 \ldots m_n)^T, \qquad (1.13)$$

where $m_i^T = (m_{i1} \ldots m_{in})$ $(i=1\ldots n)$ are the rows of $(m_{ij})_{i,j=1}^n$ and $(m_{1j} \ldots m_{nj})^T = v(e_j)$ $(j=1,\ldots,n)$ are its columns. Also, we recall that, for a linear subspace G of R^n, $G^\perp$ denotes the set

$$G^\perp = \{\mu^T = (\mu_1 \ldots \mu_n) \in R^n \mid \mu^T g = 0 \ (g \in G)\}. \qquad (1.14)$$

Corollary 1.3. *Let G be a linear subspace of* R^n, *let* $h: R^n \to \overline{R}^m$ *be such that* $A_{h(g_o)}(h)$ *is convex, and let*

$g_o \in G$. Then we have $1^\circ \Leftrightarrow 2^\circ \Leftrightarrow 5^\circ$ of theorem 1.1, where we can take $v=(m_1 \dots m_n)^T$ such that $m_1^T, \dots, m_n^T \in G^\perp$.

Proof. Let $v_o=(m_1 \dots m_n)^T$ satisfy (1.8). Then, clearly,

$$(m_1 \dots m_i)^T y \leq_L (m_1 \dots m_i)^T g$$

$$(y \in A_{h(g_o)}(h),\ g \in G,\ i=1,\dots,n). \qquad (1.15)$$

In particular, for i=1, we must have

$$m_1^T y \leq m_1^T g \qquad (y \in A_{h(g_o)}(h),\ g \in G), \qquad (1.16)$$

whence, since G is a linear subspace of R^n, we obtain $m_1^T \in G^\perp$. Assume now that $i-1<n$ and $m_1^T, \dots, m_{i-1}^T \in G^\perp$. By (1.15), there are two cases:

Case (a). There exist $y' \in A_{h(g_o)}(h)$ and $g' \in G$ such that

$$(m_1 \dots m_i)^T y' = (m_1 \dots m_i)^T g' = (0 \dots 0\ m_i^T g')^T. \qquad (1.17)$$

Then, since G is a linear subspace of R^n, we have $g'+g \in G$ for all $g \in G$, whence, by (1.17), (1.15) and $m_1^T, \dots, m_{i-1}^T \in G^\perp$,

$$(0 \dots 0\ m_i^T g')^T = (m_1 \dots m_i)^T y' \leq_L$$
$$\leq_L (m_1 \dots m_i)^T (g'+g) = (0 \dots 0\ m_i^T (g'+g))^T \qquad (g \in G).$$

Consequently, $m_i^T g' \leq m_i^T (g'+g)$ $(g \in G)$, whence, since G is a linear subspace of R^n, we obtain $m_i^T \in G^\perp$.

Case (b). All inequalities in (1.15) are strict. Then, clearly, any matrix $v'=(m_1 \dots m_{i-1}\ r_i \dots r_n)^T \in \mathcal{L}(R^n)$ satisfies (1.8), so we can take $r_i^T, \dots, r_n^T \in G^\perp$.

2. SURROGATE SUBDIFFERENTIALS

Definition 2.1. For $h: R^n \to \bar{R}^m$ and $y_o \in R^n$, we shall call surrogate subdifferential of h at y_o the set

$\partial^\gamma h(y_o) \subseteq \mathcal{L}(R^n)$ defined by

$$\partial^\gamma h(y_o)=\{v\in\mathcal{L}(R^n)\,|\,y_o\in\mathcal{V}_{\Delta_{\{y_o\},v}}(h)\}=$$

$$=\{v\in\mathcal{L}(R^n)\,|\,y_o\in\mathcal{V}_{\{y\in R^n|v(y)\geqslant_L v(y_o)\}}(h)\}; \qquad (2.1)$$

here the last equality holds by proposition 1.1 b).

Proposition 2.1. We have

$$\partial^\gamma h(y_o)=\{v\in\mathcal{L}(R^n)\,|\,v(y)<_L v(y_o)\quad (y\in A_{h(y_o)}(h))\}. \quad (2.2)$$

Proof. By (2.1), $v\in\partial^\gamma h(y_o)$ if and only if

$$h(y)\nless h(y_o) \qquad (y\in R^n,\ v(y)\geqslant_L v(y_o)), \quad (2.3)$$

which is equivalent to

$$v(y)<_L v(y_o) \qquad (y\in R^n,\ h(y)<h(y_o)). \quad (2.4)$$

Remark 2.1. Geometrically, proposition 2.1 means that $\partial^\gamma h(y_o)$ is the set of all $v\in\mathcal{L}(R^n)$ which separate $A_{h(y_o)}(h)$ from y_o.

Proposition 2.2. If $A_{h(y_o)}(h)$ is convex, then $\partial^\gamma h(y_o)\neq\emptyset$.

Proof. Since $A_{h(y_o)}(h)$ is convex and $y_o\notin A_{h(y_o)}(h)$, there exists (by theorem 0.1) $v\in\mathcal{L}(R^n)$ satisfying (2.4), whence, by proposition 2.1, $v\in\partial^\gamma h(y_o)$.

Remark 2.2. a) We have the following result of converse type: If $m=1$ and $\partial^\gamma h(y_o)\neq\emptyset$ for all $y_o\in R^n$, then h is quasi-convex (i.e., $A_c(h)$ is convex for each $c\in R$). Indeed, if $c\in R$ and $y_o\in R^n\setminus A_c(h)$, then $h(y_o)\geqslant c$, whence $A_c(h)\subseteq A_{h(y_o)}(h)$. Then, by (2.2), for any $v\in\partial^\gamma h(y_o)$ we have

$$v(y)<_L v(y_o) \qquad (y\in A_c(h)), \quad (2.5)$$

so $A_c(h)$ can be separated from any outside point y_o by some $v \in \mathcal{L}(R^n)$, and hence $A_c(h)$ is an intersection of generalized half-spaces (by [18], proposition 1.3); thus, $A_c(h)$ is convex. Note that this result also follows from [9], propositions 2.9 (implication a) $\Rightarrow$ b)), 2.3, 2.4 and corollary 2.5.

b) One can define another surrogate subdifferential of h at y_o, by

$$\partial_o^\gamma h(y_o)=\{v\in \mathcal{U}(R^n)\,|\,y_o\in \mathcal{V}_{\Delta_{\{y_o\},v}}(h)\}. \qquad (2.6)$$

Then, clearly, $\partial_o^\gamma h(y_o) \subseteq \partial^\gamma h(y_o)$, and the preceding results remain valid for $\partial^\gamma h(y_o)$ replaced by $\partial_o^\gamma h(y_o)$. Similar remarks can also made for $\mathcal{U}(R^n)$ replaced by $\mathcal{O}(R^n)$. In the sequel, for simplicity, we shall consider only $v\in\mathcal{L}(R^n)$.

We recall (see e.g. [1], Ch.II, §1) that the indicator operator $\chi_G: R^n \to \overline{R}^m$ of a set $G \subseteq R^n$ is defined by

$$\chi_G(y)=\begin{cases} 0 & \text{if } y\in G \\ +\infty & \text{if } y\notin G, \end{cases} \qquad (2.7)$$

where $+\infty$ denotes the element $(+\infty \ldots +\infty)^T \in \overline{R}^m$.

<u>Proposition 2.3</u>. <u>For any</u> $G \subset R^n$ <u>and</u> $h: R^n \to \overline{R}^m$, <u>we have</u>

$$\mathcal{V}_G(h)=G\cap \mathcal{V}_{R^n}(h\dotplus\chi_G), \qquad (2.8)$$

<u>where</u> $x\dotplus(+\infty)=+\infty$ $(x\in\overline{R}^m)$.

<u>Proof</u>. If $g_o\in G\setminus \mathcal{V}_{R^n}(h\dotplus\chi_G)$, then there exists $y\in R^n$ such that $h(y)\dotplus\chi_G(y)<h(g_o)\dotplus\chi_G(g_o)=h(g_o)$. Hence, $\chi_G(y)=0$, so $y\in G$, and $h(y)<h(g_o)$. Thus, $g_o\notin\mathcal{V}_G(h)$.

Conversely, if $g_o\in G\setminus\mathcal{V}_G(h)$, then there exists $g\in G$ such that $h(g)<h(g_o)$, whence $\chi_G(g)=\chi_G(g_o)=0$ and $h(g)\dotplus$

$\dot{+}\chi_G(g) < h(g_o) \dot{+} \chi_G(g_o)$. Thus, $g_o \notin \mathcal{V}_{R^n}(h \dot{+} \chi_G)$.

Theorem 2.1. *Under the assumptions of theorem* 1.1, *the statements* 1^o-7^o *of theorem* 1.1 *are equivalent to the following ones:*

8^o. $0 \in \partial^\gamma(h \dot{+} \chi_G)(g_o)$.

9^o. $\partial^\gamma(h \dot{+} \chi_G)(g_o) = \mathcal{L}(R^n)$.

Proof. $8^o \Rightarrow 1^o$. Note that

$$\Delta_{\{y_o\},0} = \{y \in R^n \mid 0(y) \geqslant_L 0(y_o)\} = R^n \quad . \tag{2.9}$$

Hence, by (2.1), we have 8^o if and only if

$$(h \dot{+} \chi_G)(y) \nless (h \dot{+} \chi_G)(g_o) \qquad (y \in R^n), \tag{2.10}$$

or, equivalently, $h(y) \nless h(g_o)$ $(y \in G)$.

$1^o \Rightarrow 9^o$. If 1^o holds, then we have (2.10) for all $y \in R^n$, whence also for all $y \in R^n$ such that $v(y) \geqslant_L v(g_o)$, where $v \in \mathcal{L}(R^n)$ is arbitrary. Hence, by (2.1), we get $v \in \partial^\gamma(h \dot{+} \chi_G)(g_o)$.

Finally, the implication $9^o \Rightarrow 8^o$ is obvious.

3. SURROGATE LAGRANGIANS

Definition 3.1. We shall call *surrogate Lagrangian* for problem (P) the operator $\ell : R^n \times \mathcal{L}(R^n) \to \overline{R}^m$ defined by

$$\ell(y,v) = h(y) \dot{+} \chi_{\Delta_{G,v}}(y) \qquad (y \in R^n), \tag{3.1}$$

where $\Delta_{G,v} \subseteq R^n$ is the set (1.1).

Theorem 3.1. *Under the assumptions of theorem* 1.1, *the above statements* 1^o-9^o *are equivalent to each of*

10^o-12^o. *There exists* $v \in \mathcal{L}(R^n)$, $v \in \mathcal{U}(R^n)$, $v \in \mathcal{O}(R^n)$, *respectively, such that*

$$g_o \in G \cap \mathcal{V}_{R^n}(\ell(.,v)) . \tag{3.2}$$

Proof. It is enough to show that, for any $v \in \mathcal{L}(R^n)$, (1.9) is equivalent to (3.2).

If $g_o \in G \setminus \mathcal{V}_{R^n}(\ell(.,v))$, then there exists $y \in R^n$ such that $\ell(y,v) < \ell(g_o,v)$, i.e. $h(y) \dot{+} \chi_{\Delta_{G,v}}(y) < h(g_o) \dot{+} \chi_{\Delta_{G,v}}(g_o) = h(g_o)$ (since $g_o \in G \subsetneq \Delta_{G,v}$). Hence, $\chi_{\Delta_{G,v}}(y)=0$, so $y \in \Delta_{G,v}$, and $h(y) < h(g_o)$. Thus, $g_o \notin \mathcal{V}_{\Delta_{G,v}}(h)$.

Conversely, if $g_o \in G \setminus \mathcal{V}_{\Delta_{G,v}}(h)$, then there exists $y \in \Delta_{G,v}$ such that $h(y) < h(g_o)$, whence $\chi_{\Delta_{G,v}}(y) = \chi_{\Delta_{G,v}}(g_o) = 0$ (by $g_o \in G \subseteq \Delta_{G,v}$) and $\ell(y,v) < \ell(g_o,v)$. Thus, $g_o \notin \mathcal{V}_{R^n}(\ell(.,v))$.

4. SURROGATE DUAL PROBLEMS. SURROGATE LAGRANGE MULTIPLIERS

Definition 4.1. Given the vector optimization problem (P) of (0.5), let

$$W = \{z \in R^m \mid \exists v \in \mathcal{L}(R^n),\ \ell(y,v) \nleq z \quad (y \in R^n)\}, \tag{4.1}$$

where ℓ is the surrogate Lagrangian (3.1). We shall call surrogate dual problem to (P), the problem

(Q) find the Pareto maximal elements of W. (4.2)

Any element $z \in W$ will be called a feasible element for (Q); any Pareto maximal $z \in W$ will be called a solution of (Q), and any $v \in \mathcal{L}(R^n)$ as in (4.1) will be called a surrogate Lagrange multiplier associated to the solution z.

Theorem 4.1. Under the assumptions of theorem 1.1, the above statements 1^o-12^o are equivalent to each of

13^o. $h(g_o) \in W$.

14^o. $h(g_o)$ is a solution of (Q).

Proof. $5^\circ \Longrightarrow 13^\circ$. If $v \in \mathscr{L}(R^n)$ is as in 5°, then

$$\ell(y,v) = h(y) \dot{+} \chi_{\Delta_{G,v}}(y) = \begin{cases} h(y) \not< h(g_o) & (y \in \Delta_{G,v}) \\ +\infty \not< h(g_o) & (y \notin \Delta_{G,v}) \end{cases},$$

so $h(g_o) \in W$.

$13^\circ \Longrightarrow 14^\circ$. Assume that $h(g_o) \in W$ and let $z' \in W$. Then there exists $v' \in \mathscr{L}(R^n)$ such that

$$h(y) \dot{+} \chi_{\Delta_{G,v'}}(y) = \ell(y,v') \not< z' \qquad (y \in R^n), \qquad (4.3)$$

whence, since $g_o \in G \subseteq \Delta_{G,v'}$ (by (1.2)), we obtain

$$h(g_o) = h(g_o) \dot{+} \chi_{\Delta_{G,v'}}(g_o) = \ell(g_o, v') \not< z' \ ;$$

hence, since $z' \in W$ was arbitrary, $h(g_o)$ is a Pareto maximal element of W.

The implication $14^\circ \Longrightarrow 13^\circ$ is obvious.

$13^\circ \Longrightarrow 5^\circ$. If $h(g_o) \in W$, then there exists $v \in \mathscr{L}(R^n)$ such that

$$h(y) \dot{+} \chi_{\Delta_{G,v}}(y) \not< h(g_o) \qquad (y \in R^n), \qquad (4.4)$$

whence $h(y) \not< h(g_o)$ for all $y \in \Delta_{G,v}$, so 5° holds.

Remark 4.1. As shown by the above proof, the equivalences $5^\circ \Longleftrightarrow 13^\circ \Longleftrightarrow 14^\circ$ remain valid if we replace everywhere (including the definition of the Lagrangian ℓ) the surrogate constraint set $\Delta_{G,v}$ by any set $\Delta \subseteq R^n$ such that $G \subseteq \Delta$.

5. SADDLE-POINTS OF SURROGATE LAGRANGIANS

We recall (see [20], [13]) that if F, V are two sets and $K: F \times V \to \overline{R}^m$, a point $(y,v) \in F \times V$ is said to be *a saddle-point of* K, if

$$K(y',v) \not< K(y,v) \not< K(y,v') \qquad (y' \in F,\ v' \in V). \qquad (5.1)$$

We shall consider now the case when $F=R^n$, $V=\mathcal{L}(R^n)$ and $K=\ell$, the surrogate Lagrangian for problem (P) (see definition 3.1).

Theorem 5.1. Let G be a convex subset of R^n, let $h:R^n \to \bar{R}^m$ be such that $A_{h(g_o)}(h)$ is convex for each $g_o \in G$ and

$$G \cap \operatorname{dom} h \neq \emptyset, \tag{5.2}$$

where $\operatorname{dom} h=\{y' \in R^n \mid h(y')<+\infty\}$, and let $y \in R^n$. The following statements are equivalent:

1^o. $y \in \mathcal{V}_G(h)$.

2^o. There exists $v \in \mathcal{L}(R^n)$ such that (y,v) is a saddle-point of ℓ, i.e., such that

$$\ell(y',v) \not> \ell(y,v) \not> \ell(y,v') \qquad (y' \in R^n,\ v' \in \mathcal{L}(R^n)). \tag{5.3}$$

Proof. $1^o \Rightarrow 2^o$. Assume 1^o. Then, by theorem 4.1, implication $1^o \Rightarrow 13^o$, there exists $v \in \mathcal{L}(R^n)$ such that (using also $y \in G$ and (1.2))

$$\ell(y',v) \not> h(y)=h(y) \dotplus \chi_{\Delta_{G,v}}(y)=\ell(y,v) \qquad (y' \in R^n).$$

Furthermore, using again $y \in G$ and (1.2),

$$\ell(y,v)=h(y) \not> h(y)=h(y) \dotplus \chi_{\Delta_{G,v'}}(y)=\ell(y,v') \qquad (v' \in \mathcal{L}(R^n)).$$

$2^o \Rightarrow 1^o$. Let $v \in \mathcal{L}(R^n)$ be as in 2^o. If $y \notin G$, then, since G is convex, there exists (by theorem 0.1) $v' \in \mathcal{L}(R^n)$ such that

$$v'(y) <_L v'(g) \qquad (g \in G). \tag{5.4}$$

Then, by (1.1), we have $y \notin \Delta_{G,v'}$, whence

$$\ell(y,v')=h(y) \dotplus \chi_{\Delta_{G,v'}}(y)=+\infty , \tag{5.5}$$

and hence, by the second part of (5.3), $\ell(y,v)=+\infty$. But then, for any $y' \in G \cap \operatorname{dom} h \subseteq \Delta_{G,v} \cap \operatorname{dom} h$, we have

$$\ell(y',v)=h(y') \dotplus \chi_{\Delta_{G,v}}(y')<+\infty=\ell(y,v),$$

in contradiction with the first part of (5.3). Thus, $y \in G$.

Finally, since $y \in G \subseteq \Delta_{G,v}$, for each $y' \in G \subseteq \Delta_{G,v}$ we have, by the first part of (5.3),

$$h(y') = \ell(y', v) \nless \ell(y, v) = h(y),$$

and thus $y \in \mathcal{V}_G(h)$.

Remark 5.1. a) We have used the assumption (5.2) only in the proof of the implication $2^o \Rightarrow 1^o$. Note that, essentially, (5.2) is no restriction of the generality, since if $G \cap \text{dom}\, h = \emptyset$, i.e., $h(y') = +\infty$ $(y' \in G)$, then, obviously, $G = \mathcal{V}_G(h)$.

b) One can give a corollary of theorem 5.1, similar to corollary 1.2.

Combining theorems 5.1 and 4.1, we obtain

Corollary 5.1. *Under the assumptions of theorem 5.1, the statements* 1^o, 2^o *of theorem* 5.1 *are equivalent to the following ones:*

3^o. $y \in G$ *and* $h(y) \in W$.

4^o. $y \in \mathcal{V}_G(h)$ *and* $h(y)$ *is a solution of* (Q).

6. INEQUALITY CONSTRAINTS

We shall consider now the case when G is given by (0.6), and hence $\mathcal{V}_G(h)$ is the set (0.7). Of course, for this case one can consider the surrogate constraint sets $\Delta_{G,v}$ $(v \in \mathcal{L}(R^n))$ defined by (1.1), and one can apply the results of §§1-5. However, exploiting (0.6) and (0.7), we shall introduce some other surrogate constraint sets $\Delta_{u,v} \subseteq R^n$ $(v \in \mathcal{L}(R^k))$ and we shall prove corresponding results for them; moreover, we shall show that it is sufficient to consider $v \in \mathcal{L}(R^k)$ such that $v \geqslant_L 0$, or even such that $v \geqslant 0$.

Namely, in view of G of (0.6), where $u: R^n \to R^k$, let us define, for each $v\in\mathscr{L}(R^k)$, the set

$$\Delta_{u,v}=\{y\in R^n \mid v(u(y)) \leqslant_L 0\}. \tag{6.1}$$

Remark 6.1. a) For $v\geqslant 0$, the sets $\Delta_{u,v}$ correspond to those given, for scalar optimization, in [7].

b) In the particular case when $k=n$ and $u=I$, the identity operator, (0.6) and (6.1) become, respectively,

$$G=\{g\in R^n \mid g\leqslant 0\}=-R^n_+ , \tag{6.2}$$

$$\Delta_{I,v}=\{y\in R^n \mid v(y) \leqslant_L 0\}. \tag{6.3}$$

Note that $\Delta_{I,v}$ is not a particular case of $\Delta_{G,v}$ of (1.1); indeed, e.g., for $v=I$, we have $\Delta_{I,v}=-C_L$, but

$$\Delta_{G,v}=\{y\in R^n \mid \exists g\in -R^n_+ , \ y \geqslant_L g\}=R^n .$$

We shall prove now, for G of (0.6) and $\Delta_{u,v}$ of (6.1), where $v\in\mathscr{L}(R^k)$, $v\geqslant 0$, some results corresponding to those of §1 on any $G\subset R^n$ and $\Delta_{G,v}$ of (1.1), where $v\in\mathscr{L}(R^n)$.

Proposition 6.1. For G of (0.6) and $v\in\mathscr{L}(R^k)$, $v\geqslant 0$, we have

$$G\subseteq \Delta_{u,v} . \tag{6.4}$$

Proof. Obviously, $v\geqslant 0$ and $u(y)\leqslant 0$ imply $v(u(y))\leqslant 0$.

Theorem 6.1. Let $u:R^n\to R^k$ be a convex mapping, let $g_o\in G$, where G is defined by (0.6), and let $h:R^n\to \overline{R}^m$ be such that $A_{h(g_o)}(h)$ is convex. The following statements are equivalent:

1°. $g_o\in \mathcal{V}_G(h)$.

2°-4°. There exists $v\in\mathscr{L}(R^k)$, $v\in\mathcal{U}(R^k)$, $v\in\mathcal{O}(R^k)$, respectively, such that

$$v(x) >_L 0 \qquad (x\in R^k, \exists\, y\in A_{h(g_o)}(h),\ u(y)\leqslant x). \tag{6.5}$$

5°-7°. Same as 2°, 3° and 4° respectively, with

$v \geq_L 0$.

8^o–9^o. Same as 2^o and 3^o respectively, with $v \geq 0$.

10^o–12^o. There exists $v \in \mathcal{L}(R^k)$, $v \in \mathcal{U}(R^k)$, $v \in \mathcal{O}(R^k)$, respectively, satisfying $v \geq_L 0$ and

$$g_o \in \mathcal{V}_{\Delta_{u,v}}(h). \tag{6.6}$$

13^o–14^o. Same as 10^o and 11^o respectively, with $v \geq 0$.

Proof. $1^o \Leftrightarrow \ldots \Leftrightarrow 9^o$. Since $A_{h(g_o)}(h)$ and u are convex, so is $u(A_{h(g_o)}(h)) + R^k_+$; indeed, if $y_1, y_2 \in A_{h(g_o)}(h)$, $z_1, z_2 \in R^k_+$, $0 \leq \lambda \leq 1$, then $\lambda y_1 + (1-\lambda) y_2 \in A_{h(g_o)}(h)$ and $\lambda u(y_1) + (1-\lambda) u(y_2) = u(\lambda y_1 + (1-\lambda) y_2) + z$ for some $z \in R^k_+$, whence $\lambda(u(y_1) + z_1) + (1-\lambda)(u(y_2) + z_2) \in u(A_{h(g_o)}(h)) + R^k_+$. Also, obviously,

$$u(A_{h(g_o)}(h)) + R^k_+ = \{x \in R^k \mid \exists y \in A_{h(g_o)}(h),\ u(y) \leq x\}, \tag{6.7}$$

and hence (6.5) is equivalent to

$$v(u(y)) >_L v(z) \qquad (y \in A_{h(g_o)}(h),\ z \in -R^k_k). \tag{6.8}$$

Hence, the equivalences $1^o \Leftrightarrow \ldots \Leftrightarrow 9^o$ follow from (0.7) and theorem 0.2.

$7^o \Rightarrow 12^o$. If $v \in \mathcal{O}(R^k)$ of 7^o does not satisfy (6.6), then there exists $y' \in \Delta_{u,v} \cap A_{h(g_o)}(h)$. Then, by $y' \in \Delta_{u,v}$ we have $v(u(y')) \leq_L 0$. On the other hand, for $x = u(y') \in R^k$ we obtain, by $y' \in A_{h(g_o)}(h)$ and (6.5), $v(u(y')) = v(x) >_L 0$, a contradiction. The proof of the implication $9^o \Rightarrow 14^o$ is similar.

The implications $12^o \Rightarrow 11^o \Rightarrow 10^o$ and $14^o \Rightarrow 13^o \Rightarrow 10^o$ are obvious.

$10^o \Rightarrow 5^o$. If $v \in \mathcal{L}(R^k)$ of 10^o does not satisfy (6.5), then for some $x \in R^k$ there exists $y \in A_{h(g_o)}(h)$ with $u(y) \leq x$, such that $v(x) \leq_L 0$. Then, since $v \geq_L 0$, we obtain, by

(0.10), $v(u(y)) \leq_L v(x) \leq_L 0$, so $y \in \Delta_{u,v}$, and thus $g_0 \notin \mathcal{V}_{\Delta_{u,v}}(h)$.

Remark 6.2. The above proof of the implication $7^\circ \Rightarrow 12^\circ$ shows that, for any $v \in \mathcal{L}(R^k)$ (not necessarily in $\mathcal{O}(R^k)$ and not necessarily $\geq_L 0$), (6.5) implies (6.6). Thus, 2°-4° imply, respectively, that there exists $v \in \mathcal{L}(R^k)$, $v \in \mathcal{U}(R^k)$, $v \in \mathcal{O}(R^k)$, satisfying (6.6). However, for implications of the converse type (see the above proof of $10^\circ \Rightarrow 5^\circ$) we need the additional assumption $v \geq_L 0$.

Some concepts and results of the preceding sections can be carried over, mutatis mutandis, to the case of constraint sets G of the form (0.6), replacing $v \in \mathcal{L}(R^n)$ (or $\mathcal{U}(R^n)$) by $v \in \mathcal{L}(R^k)$ (or $\mathcal{U}(R^k)$), $v \geq_L 0$ (or $v \geq 0$) and $\Delta_{G,v}$ by $\Delta_{u,v}$ of (6.1); in some cases, $v \in \mathcal{O}(R^k)$ also works. For example, one can consider the surrogate Lagrangian

$$\ell(y,v) = h(y) + \chi_{\{y' \in R^n | v(u(y')) \leq_L 0\}}(y) \quad (y \in R^n, v \in \mathcal{L}(R^k), v \geq_L 0), \qquad (6.9)$$

and the surrogate dual problem

(Q) find the Pareto maximal elements of

$$W = \{z \in R^m | \exists v \in \mathcal{L}(R^k),\ v \geq_L 0,\ \ell(y,v) \nleq z \quad (y \in R^n)\}, \qquad (6.10)$$

where $\ell(y,v)$ is defined by (6.9). We omit the details.

REFERENCES

[1] G.P.Akilov and S.S.Kutateladze, Ordered vector spaces. Nauka, Novosibirsk, 1978 (in Russian).

[2] J.-P.Crouzeix, Contributions à l'étude des fonctions quasi-convexes. Thèse, Université de Clermont, 1977.

[3] F.Glover, A multiphase-dual algorithm for the zero--one integer programming problem. Oper. Res. 13 (1965), 879-919.

[4] H.J.Greenberg and W.P.Pierskalla, Surrogate mathematical programming. Oper. Res. 18(1970), 924-939.

[5] P.C.Hammer, Maximal convex sets. Duke Math. J. 22 (1955), 103-106.

[6] J.Jahn, Duality in vector optimization, Math.Progr. 25(1983), 343-353.

[7] D.G.Luenberger, Quasi-convex programming. SIAM J. Appl. Math. 16(1968), 1090-1095.

[8] J.-E.Martínez-Legaz, A generalized concept of conjugation. In: Optimization: Theory and algorithms (Eds. J.-B.Hiriart-Urruty, W.Oettli and J.Stoer). Lecture Notes in Pure and Appl. Math. 86, pp.45-59. M.Dekker, New York and Basel, 1983.

[9] J.-E.Martínez-Legaz, Exact quasiconvex conjugation. Z. Oper. Res. 27(1983), 257-266.

[10] J.-E.Martínez-Legaz, Lexicographical order, inequality systems and optimization. In: System modelling and optimization (Ed. P. Thoft-Christensen). Lecture Notes in Control and Information Sciences 59, pp.203-212. Springer-Verlag, Berlin-Heidelberg-New York-Tokyo, 1984.

[11] J.-E.Martínez-Legaz, Lexicographical order and duality in multiobjective programming. European J. Oper. Res. (to appear).

[12] J.-E.Martínez-Legaz and I.Singer, Lexicographical separation in R^n (to appear). Preprint INCREST 73/1985.

[13] V.V.Podinovskii and V.D.Nogin, Pareto-optimal solutions of multicriterial problems. Nauka, Moscow, 1982 (in Russian).

[14] I.Singer, Optimization by level set methods. I: Duality formulae. In: Optimization: Theory and algorithms (Eds. J.-B.Hiriart-Urruty, W.Oettli and J.Stoer). Lecture Notes in Pure and Appl.Math. 86, pp.13-43. M.Dekker, New York and Basel, 1983.

[15] I.Singer, Optimization by level set methods. III: Characterizations of solutions in the presence of duality. Numer.Funct.Anal.Optim. 4(1981-1982), 151-170.

[16] I.Singer, Optimization by level set methods. IV: Generalizations and complements. Numer. Funct. Anal. Optim. 4(1981-1982), 279-310.

[17] I.Singer, Surrogate dual problems and surrogate Lagrangians. J.Math.Anal.Appl. 98(1984), 31-71.

[18] I.Singer, Generalized convexity, functional hulls and applications to conjugate duality in optimization. In: Selected topics in operations research

and mathematical economics (Eds. G.Hammer and D. Pallaschke). Lecture Notes in Econ. and Math. Systems 226, pp.49-79. Springer-Verlag, Berlin-Heidelberg-New York-Tokyo, 1984.

[19] I.Singer, A general theory of surrogate dual and perturbational extended surrogate dual optimization problems. J. Math. Anal. Appl. 104 (1984), 351-389.

[20] T.Tanino, Saddle-points and duality in multi-objective programming. Int. J. Systems Sci. 13 (1982), 323-335.

[21] R.M.Van Slyke and R.J.-B.Wets, A duality theory for abstract mathematical programs with applications to optimal control theory. J.Math.Anal.Appl. 22 (1968), 679-706.