OPTIMIZATION WITH RESPECT TO PARTIAL ORDERINGS

by

J. M. Borwein

Submitted for the degree of Doctor of Philosophy
at the University of Oxford

Jesus College

Trinity Term 1974

TABLE OF CONTENTS

		Page number
Introduction		1
Chapter One :	Convex type functions	3
Chapter Two :	Tangent and pseudotangent cones	41
Chapter Three:	Farkas Lemmas and transposition theorems	61
Chapter Four:	Subgradients, tangent cones and local supportability	77
Chapter Five:	First order conditions	91
Chapter Six:	Multiplier theorems and minimax theorems	123
Chapter Seven:	Second order conditions	153
Chapter Eight:	Optimization in Hilbert space and variational inequalities for optimization	171
Bibliography		201

ABSTRACT

The thesis is primarily concerned with optimization problems which have objective functions which do not take values on the real line.

In chapter one convexity properties are investigated for functions in partially ordered vector spaces.

In chapter two the concept of a tangent cone is introduced and the previously little used notion of a weak tangent cone is defined. Properties of these cones are investigated and various differential relationships are proved.

Chapter three establishes several transposition theorems and Farkas lemmas both for linear and non-linear systems. These results have some applications in later chapters.

In chapter four the concept of a subgradient is extended and related to the tangent cone results of chapter two.

The fifth chapter establishes Kuhn-Tucker and Fritz John type necessary and sufficient conditions for general non-linear programming problems to have solutions with respect to various notions of minimization. These conditions are given in tangent cone terms. They include one sided derivative and subgradient results.

Chapter six includes a variety of multiplier theorems for convex and quasiconvex programmes in partially

ordered spaces. A minimax theorem is included.

Chapter seven uses tangent cones to generalize known second order conditions to more general problems and spaces.

Finally, in chapter eight, the results of chapter five are specialized to Hilbert space using pseudoinverses and projections on convex sets. The chapter also contains a section on variational inequalities which centres around the non-linear complementarity problem.

INTRODUCTION

In the thesis I have attempted to extend many of the standard results of nonlinear programming theory in the following directions:

(1) The objective functions are not generally supposed real valued,

(2) The various convexity assumptions usually associated with sufficiency conditions have been generalized and weakened,

(3) The notion of a tangent cone, and of constraint qualifications given in tangent cone terms, has been extended.

Using these three extensions I have attempted to both unify and enlarge what theory does exist for optimization with respect to partial orderings. I have found it possible to state and prove a number of fairly general theorems.

The two notions of minimization that are used have both been studied before, but not within the mainstream of abstract optimization theory as represented by the tangent cone investigations of Varaiya, Guignard and others. I hope that I have partially rectified this situation.

Some remarks on format seem in order. The chapters are arranged so that almost all the preliminary results are proved in chapters one through four and are applied and investigated further in chapters five through eight. The chapters are divided internally into numbered paragraphs, which are also used for purposes of cross reference. I have tried to keep the notation as uniform as possible. For example bilinear forms are used only in chapter eight where it seemed unavoidable. Otherwise the value of a linear functional x^+ at a point x has been denoted $x^+(x)$ not (x^+,x).

Finally, the bibliography is arranged chronologically within each author's listing and all references are referred to simply by name and date.

Chapter One

CONVEX TYPE FUNCTIONS

This chapter is devoted to a survey and extension of results on convex type functions. In the literature these functions are, except in the case of constraint functions, usually taken to be real valued. As will be seen this restriction is often unnecessary.

Certain preliminary definitions are necessary. For the sake of convenience all spaces are assumed to be real vector spaces. All general topological and functional analytic definitions are used as in Robertson and Robertson (1964) if they are not defined explicitly.

[1] A set C in X is <u>convex</u> if $\lambda x_1 + (1 - \lambda)x_2 \in C$ whenever $x_1, x_2 \in C$ and $0 \leq \lambda \leq 1$.

[2] A <u>topological vector space</u> is a vector space with an associated topology in which the vector operations are continuous. Such a topology is said to be <u>compatible</u>.

[3] A (locally) <u>convex</u> (topological vector) <u>space</u> is a topological vector space in which there is a base of neighbourhoods of the origin consistent of convex sets.

[4] A cone S in X is a set with the property that $\lambda x \in S$ whenever $x \in S$ and $\lambda \geq 0$. S is a convex cone if it is convex in addition.

[5] Partial orderings

Each convex cone S in X determines a relation '$\geq_S$' which is transitive and reflexive and which is given by

$$x \geq_S y \quad \text{iff} \quad x - y \in S.$$

When there is no ambiguity $\geq_S$ will be denoted simply by $\geq$. This relationship is compatible with the vector structure. That is

(1) $x \geq o$ and $y \geq o$ implies $x + y \geq o$.

(2) $x \geq o$ and $\lambda \geq o$ implies $\lambda x \geq o$.

The relation determined by the cone S is called the vector (partial) ordering of X and the said (X,S) or $(X,\geq)$ is a partially ordered vector space.

[6] Conversely if $\geq$ is a symmetric and transitive relation satisfying (1) and (2) then $S = \{x | x \in X \text{ and } x \geq 0\}$ is a convex cone in S and $\geq$ is exactly the ordering on X induced by S.

[7] In some cases only (2) is required of the cone which need not then be convex. If $S \cap -S = \{o\}$ then S is said to be pointed or proper. It is clear that S is pointed if and only if the induced ordering is anti-symmetric, that is if $x \geq o$ and $x \leq o$ $(- \geq o)$ then $x = o$.

With these definitions it is possible to turn to an investigation of the elementary properties of convex type functions. It will be useful to list the eight function types of primary interest.

[8] $f : X \to Y$ is <u>convex</u> with respect to S a convex cone in Y if

$$f(\lambda x + (1-\lambda) y) \leq_s \lambda f(x) + (1-\lambda) f(y)$$
$$x, y \in X; \; 0 \leq \lambda \leq 1.$$

If $-f$ is convex f is said to be <u>concave</u>. Similar remarks apply to the following definitions.

[9] $f: X \to Y$ is <u>quasiconvex</u> with respect to S, if for $0 \leq \lambda \leq 1$.

$$f(\lambda x + (1-\lambda) y) \leq f(y) \text{ whenever } f(x) \leq f(y) \; x, y \in X.$$

[10] $f : X \to Y$ is <u>strongly quasiconvex</u> with respect to S if $S(z) = \{x \mid f(x) \leq_s z\}$ is a convex set for each z in Y. $S(z)$ is called a <u>level set</u> of f with respect to S.

[11] $f : X \to Y$ is <u>absolutely quasiconvex</u> with respect to S if whenever

$$f(\lambda x + (1-\lambda) y) \geq f(y)$$

for some $0 \leq \lambda \leq 1$, one has

$$f(x) \geq f(y).$$

[12] If S is a cone in a topological vector space and $S^o \neq \emptyset$ it is possible to define a <u>strict inequality</u> (denoted by $>_s$) by $x >_s y$ if and only if $x - y \in S^o$. Clearly $x > y$ implies $x \geq y$. Equally clearly if $S = R^+$, the nonnegative real axis, then $x >_s 0$ has the usual meaning.

With this extra definition to sharpen the previous ones:

[13] $f : X \rightarrow Y$ is <u>strictly convex</u> with respect to S if

$$f(\lambda x + (1-\lambda)\, y) < \lambda f(x) + (1-\lambda)\, f(y) \quad x, y \in X;\ o \leq \lambda \leq 1.$$

[14] $f : X \rightarrow Y$ is <u>strictly quasiconvex</u> with respect to S if

$$f(\lambda x + (1-\lambda)\, y) < f(y) \text{ whenever } f(x) \leq f(y) \quad x \neq y \text{ and } o < \lambda < 1.$$

[15] $f : X \rightarrow Y$ is <u>strongly strictly quasiconvex</u> with respect to S if whenever $o < \lambda < 1$, $x \neq y$, and $f(x) \leq z$, $f(y) \leq z$ then $f(\lambda x + (1-\lambda)y) < z$.

[16] Ponstein (1967) has introduced, in the real case, a property which will be called (P) <u>strict quasiconvexity</u> and which is weaker than [14] or [15] but suffices for some basic propositions. $f : X \rightarrow Y$ is (P) strictly quasiconvex with respect to S if

$$f(\lambda x + (1-\lambda)\, y) < f(y) \text{ whenever } f(x) < f(y) \text{ and } o < \lambda \leq 1.$$

[17] <u>Relationships between the above definitions</u>

The properties will be referred to by their respective numbers since these are directly above. All the relationships follow directly from the definitions.

(1) [13] → [8] → [10] → [9]

(2) [13] → [15] → [14] → [16]

(3) [13] → [15] → [10] → [9]

(4) [8] → [11], [8] → [16]

The next proposition will be useful in the sequel.

[18] Proposition: If S is closed and $S^o \neq \emptyset$, strong quasiconvexity is equivalent to $T(z) = \{x \mid f(x) < z\}$ being convex $\forall z \in Y$.

Proof: $\Rightarrow$ Let N be a convex neighbourhood in S. Suppose that $x, y \in T(z)$. Then $f(x) - z \in -S^o$; $f(y) - z \in -S^o$ or equivalently, for any $a \in N$ one has

$$f(x) \leq z - a, \; f(y) \leq z - a.$$

Since f is strongly quasiconvex $f(\lambda x + (1 - \lambda) y) \leq z - a$ for all $a \in N$ and this is the same as

$$f(\lambda x + (1 - \lambda) y) < z.$$

$\Leftarrow$ If $T(z)$ is convex and $f(x) \leq z$, $f(y) \leq z$ then when $a \in S^o$ $a/_n \in S^o$ for any $n \in N$ and

$$f(x) < z + a/_n, \; f(y) < z + a/_n.$$

By the convexity of $T(z)$ $\forall z$,

$$f(\lambda x + (1 - \lambda) y) < z + a/_n$$

and taking limits

$$f(\lambda x + (1 - \lambda) y) - z \in - \overline{S^o} = S. \blacksquare$$

When $(Y, S) = (R, R^+)$ or more generally any linearly ordered space more can be said about the definitions.

[19] Proposition: Strong quasiconvexity and quasiconvexity coincide. (That is [9] $\Leftrightarrow$ [10]).

Proof: [10] $\Rightarrow$ [9] is general. Conversely if $f(x) \leq z$ and $f(y) \leq z$ then $z \geq z_1 = \max(f(x), f(y))$. By [9] one has

$$f(\lambda x + (1 - \lambda) y) \leq z_1 \leq z \text{ as required. } \blacksquare$$

[20] In a linearily ordered space the cone generally has no interior. One can, however, define x <u>less than but not equal</u> y by

$x \lneq y$ if $x \leq y$ and $x \neq y$.

This definition is good for any cone and in the case of a linear ordering it can be used to give meaning to strict inequality. Replacing $<$ by $\lneq$ one has:

[21] <u>Proposition</u>: If (Y, S) is linearly ordered

(1) Strict quasiconvexity [14] and strong strict quasiconvexity [15] coincide.

(2) Absolute quasiconvexity [11] and (P) strict quasiconvexity [16] agree.

<u>Proof</u>: [15] $\Rightarrow$ [14] and [16] $\Rightarrow$ [11] follow directly from the relevant definitions and [20] . [14] $\Rightarrow$ [15] follows as does [9] $\Rightarrow$ [10] in [19]. ∎

In general, the strong quasiconvex types are actually stronger than the corresponding quasiconvex types. As an example one has the following.

[22] <u>Example</u>: Let $f: R \to R^2$ with the ordering in R^2 the <u>orthant ordering</u> $(x_1, x_2) \geq (y_1, y_2)$ if $x_1 \geq y_1$ and $x_2 \geq y_2$. Let

$$f(x) = \begin{cases} (1,0) & x \geq 1 \\ (2,2) & |x| < 1 \\ (0,1) & x \leq -1 \end{cases}$$

Suppose $f(x) \nleq f(y)$.

(1) $y \geq 1$, then $f(x) \leq (1,0)$ and $x \in \{r | 1 \leq r < \infty\}$

(2) $|y| < 1$, then $f(x) \leq (2,2)$ and $x \in \mathbb{R}$

(3) $y \leq -1$, then $f(x) \leq (0,1)$ and $x \in \{r | -\infty < r \leq -1\}$.

In each case $\{x \mid f(x) \leq f(y)\}$ is convex and thus f is quasiconvex with respect to the orthant ordering. However, if $z = (1,1)$, $S(z) = \{x \mid f(x) \leq (1,1)\}$ is not convex and f is not strongly quasiconvex. ∎

More sophisticated examples could be given but the above serves to indicate the reason for the divergence of definitions in general.

Although (P) strict quasiconvexity does not in general imply quasiconvexity, as is shown by $f : R \to R$ with $f(o) = 1$ and $f(x) = o$, $x \neq o$, only the mildest of continuity conditions is necessary for the implication to hold when $(Y,S) = (R,R^+)$.

[23] Definition: $f\ X \to Y$ has the <u>one point exclusion property</u> with respect to S if $f\ (x_o) \leq z$ whenever $f(x) \leq z$ for all x in a line segment $[x_1, x_2]$ containing x_o and not equal to x_o.

A continuity condition which is clearly stronger is:

[24] Definition: $f : X \to Y$ is <u>lower semicontinuous</u> with respect to S if $S(z) = \{x \mid f(x) \leq z\}$ is closed $\forall z \in Y$.

Upper semicontinuity is defined similarily. On the line this reduces to the usual definition. In general, however, one also has:

[25] Definition: $f : X \to Y$ is <u>fully lower semicontinuous</u> with respect to S, a cone with interior, if for all z in y

$$F(z) = \{x \mid f(x) > z\} \text{ is an open set.}$$

These notions are related by the following proposition.

[26] Proposition: Suppose S is a closed convex cone with interior in a convex space Y. Then $f : X \to Y$ is lower semicontinuous whenever f is fully lower semicontinuous.

Proof: Suppose $x_o \in \overline{S(z)}$. Let $\{x_t \mid t \in T\}$ be a net in S(z) with limit x_o. Let $s \in S^o$ and let $m \in N$. Since $f(x_o) > f(x_o) - s/_m$ and [25] holds there is some $t_o \in t$ such that $f(x_t) > f(x_o) - s/_m$ when $t \geq t_o$. Since Y is a convex space $s/_m$ tends to zero in Y as m tends to infinity and one has

$$z \geq f(x_{t_o}) > f(x_o) - s/_m$$

and $f(x_o) \leq z$. Thus S(z) is closed.▌

It becomes apparent that some mechanism is needed for relating the order-bounded sets, those $\{x \mid y \leq x \leq z\}$, and the original topology. The following which is taken from Kelley and Namioka (1963) is sufficient for present purposes.

[27] Definition: A convex cone S in a pseudo normed space with pseudonorm p, is said to be normal if whenever x and y belong to S and have pseudo norms greater than or equal to 1 one has $p\ (x + y) \geq e$ where e is some fixed positive number.

Equivalently one has the requirement that the set $(B + S) \cap (B - S)$ is bounded where $B = \{x \mid p(x) \leq 1\}$. The following theorem holds in a pseudo normed space.

[28] Theorem: (Kelley and Namioka 23.7). If S is normal each order-bounded set is bounded. If $x_o \in S^o$ then S is normal if and only if $\{y \mid -x_o \leq y \leq x_o\}$ is bounded.▌

One has for closed cones and finite dimensional spaces the following:

[29] Proposition: If X is a finite dimensional pseudonormed Hausdorff space (and hence normable) S is pointed i.f.f. S is normal.

Proof: $\Leftarrow$ If S is not pointed there is some x with $x \in S \cap -S$ and $\| x \| = 1$ and, hence, $\| x + (-x) \| = o$ and S is not normal. $\Rightarrow$ Let $\{x_n\}$, $\{y_n\}$ be sequences in S with $\| x_n \| \geq 1$ $\| y_n \| \geq 1$ and, in contradiction of normality, with $\| x_n + y_n \| = 1/_n$. Let $k_n = x_n / \| x_n \|$ then since X is finite dimensional one can suppose k_n is convergent to k_o which will be non zero. Now, since $\|x_n\| \geq 1$,

$$\| k_n + y_n \| x_n \|^{-1} \| \leq 1/_{n \|x_n\|} \leq 1/_n$$

and this means that $k_n + \|x_n\|^{-1} y_n \rightarrow o$. Since $k_n \rightarrow k_o \neq o$ $- \|x_n\|^{-1} y_n \rightarrow - k_o \neq 0$. Since the cone is closed and $x_n, y_n \in S$ both $-k_o \in S$ and $k_o \in S$ contradicting pointedness. ▮

This result does not hold true in general Banach spaces.

[30] Example: Define the cone S in l_1 by

$$S = \{ \{ x_k \} \mid x_1 \geq o, x_k + x_{k+1} \geq o, k \geq o \}$$

S is pointed since $x \in S \cap -S$ implies that $x_1 = o$ and $x_k + x_{k+1} = o$. That is $\{x_k\} = o$.

S is not normal. Suppose e is the constant of definition [27]. Let $\{x_k\}$, $\{y_k\}$ have

$$x_k = \begin{cases} (-1)^{k+1}/_n & k = 3,\ldots,2n+3 \\ o & \text{otherwise} \end{cases}$$

and

$$y_k = \begin{cases} (-1)^k/_n & k = 2, \ldots, 2n+2 \\ 0 & \text{otherwise.} \end{cases}$$

Then $\|\{x_k\}\| = \sum_{k=3}^{2n+3} \left| \frac{(-1)^{k+1}}{n} \right| \geq 1$ and $\{x_k\} \in S$

$\|\{y_k\}\| = \sum_{k=2}^{2n+2} \left| \frac{(-1)^k}{n} \right| \geq 1$ and $\{y_k\} \in S$

$\|\{x_k + y_k\}\| = 1/_n$

Taking n sufficiently large $1/_n < e$. ∎

[31] Proposition: If both f and -f are fully lower semicontinuous with respect to a normal cone S with S closed and with interior then f is continuous.

Proof: Let $z_1, z_2 \in Y$. By hypothesis the set

$$\{x \mid z_1 < f(x) < z_2\}$$

is open. Since S is normal and has interior the norm bounded sets in Y generate the norm topology. Thus the inverse images of open sets are open and f is continuous. ∎ Since any finite dimensional pointed cone is normal this result is a true extension of the standard relationship.

Returning now to the discussion of quasiconvexity in [22]:

[32] Proposition: If $f : X \to R$ is (P) strict quasiconvex [16] and satisfies [23] then f is quasiconvex [9].

Proof: Suppose $f(x) \leq f(y)$. If $f(x) < f(y)$ then

$$f(\lambda x + (1 - \lambda) y) \leq f(y) \quad \text{by [16].}$$

If $f(x) = f(y)$ suppose that $x_o, x_1 \in [x,y]$ with $f(x_o) > f(x)$ and $f(x_1) > f(x)$. Then either $x_1 \in [x_o, y]$ or $x_1 \in [x, x_o]$. In either case by [16] one has

$$f(x_1) < f(x_o).$$

Symmetrically $f(x_o) < f(x_1)$ since $x_o \in [x_1, y]$ or $[x, x_1]$ which gives $f(x_o) < f(x_o)$. Thus there can only be one exceptional point on the line segment $[x,y]$. This is excluded by hypothesis. Thus $f(\lambda x + (1-\lambda)y) \leq f(x)$ if $o \leq \lambda \leq 1$. ▌

In similar vein to Proposition [32] is:

[33] Proposition: If $f : X \to R$ is (P) strictly quasiconvex and satisfies the one point exclusion property then if $f(x) \leq c$ and $f(y) < c$; $f(\lambda x + (1-\lambda)y) < c \quad o \leq \lambda < 1$.

Proof: Suppose $f(y) < c$ and $f(x) \leq c$. If $f(y) < f(x)$ then $f(\lambda x + (1-\lambda)y) < f(x) \leq c$ by [16], while if $f(x) \leq f(y)$ one has $f(\lambda x + (1-\lambda)y) \leq f(y) < c$ if $o \leq \lambda < 1$ since, by [32], f is quasiconvex. ▌

The next few results are concerned with properties of convex functions which for the most part do not generalise to quasiconvex ones.

Since the sums of quasiconvex functions are generally not quasiconvex (even if one is convex) it seems worth noting some of the quasiconvexity maintaining operations before turning to convex functions.

[34] Proposition: If $f : X \rightarrow R$ is quasiconvex and $g : X \rightarrow R$ is the indicator function of a convex set C (that is $g(x) = o$ if $x \in C$ and $g(x) = \infty$ if $x \notin C$) then $f + g$ is quasiconvex.

Proof: $S(r) = \{x \mid f(x) + g(x) \leq r\} = \{x \mid f(x) \leq r\} \cap C$. Since both C and $\{x \mid f(x) \leq r\}$ are convex $f + g$ is quasiconvex. ∎

[35] The pointwise supremum of a family of real valued quasiconvex functions is quasiconvex.

[36] A function M mapping (Y,S) into (Z,T) such that $x - y \in S$ if and only if $M(x) - M(y) \in T$ will be called cone monotone.

Proposition: (1) If $f : X \rightarrow Y$ is quasiconvex with respect to $S \subset Y$ and $M : Y \rightarrow Z$ is cone monotone then $g = Mf$ is quasiconvex with respect to $T \subset Z$.

(2) If f is strongly quasiconvex with respect to S and M is cone monotone and surjective then Mf is strongly quasiconvex with respect to T.

Proof: Let $g = Mf$.

(1) $\{x \mid g(x) \leq_T g(y)\} = \{x \mid M(f(x)) \leq_T M(f(y))\}$
$= \{x \mid f(x) \leq_S f(y)\}$ since M is cone monotone. Since f is quasiconvex with respect to S this last set is convex.

(2) $\{x \mid g(x) \leq_T z\} = \{x \mid M(f(x)) \leq_T M(y)\}$ since M is assumed surjective. Then as in (1) $\{x \mid M(f(x) \leq_T M(y)\}$ is equal to $\{x \mid f(x) \leq y\}$ which is convex $\forall y \in Y$ since f is strongly quasiconvex. ∎

[37] If $A : X \to Y$ is linear and $f : Y \to Z$ is quasiconvex with respect to S then $(f\,A)\,(x) = f\,(Ax)$ defines a function $fA : X \to Z$ which is quasiconvex with respect to S. If in addition A is surjective and f is strongly quasiconvex so is fA.

[38] If $A : X \to Y$ is linear and $f : X \to R$ is quasiconvex then $Af : X \to R$ defined by $(Af)(x) = \inf\{f(x) \mid Ax = y\}$ is quasiconvex.

Rockafellar (1970a) gives an exhaustive list of convexity preserving operations for real valued functions. Note that if f is taken to be convex in [37] or [38], the composite mapping is also convex.

The following characterization of convex type functions is very useful.

[39] <u>Proposition</u>: $f : X \to Y$ is strongly quasiconvex with respect to a closed cone S if f is lower semicontinuous and

$$f\left(\tfrac{1}{2}x + \tfrac{1}{2}y\right) \leq z \text{ whenever } f(x) \leq z \text{ and } f(y) \leq z.$$

<u>Proof</u>: Assume inductively that for $o \leq m \leq 2^{n-1}$ and $n \leq n_o - 1$ one has

$$(1)\ f\left(\frac{mx}{2^n} + \left(1 - \frac{m}{2^n}\right)y\right) \leq z \text{ when } f(x) \leq z;\ f(y) \leq z.$$

For $n_o = 2$ this is true by hypothesis. Now

$$(2)\ f\left(\frac{m + 2^{n_o - 1}}{2^{n_o}}x + \left[1 - \frac{m + 2^{n_o-1}}{2^{n_o}}\right]y\right)$$

$$= f\left(\tfrac{1}{2}x + \tfrac{1}{2}\left[\left(1 - \frac{m}{2^{n_o-1}}\right)y + \frac{m}{2^{n_o-1}}x\right]\right)$$

By the hypothesis of the theorem this last term is $\leq z$ if (1) holds. Exchanging x and y in (2) one sees that for $o \leq k \leq 2^n$

(3) $f\left(\frac{k}{2^n}x + \left(1 - \frac{k}{2^n}\right)y\right) \leq z$ if $f(x) \leq z$, $f(y) \leq z$.

Using the semicontinu-ity of f with respect to S and the fact that the diadic rationals are dense in $[o,1]$ one obtains the desired result. ‖ Similarly:

[40] <u>Proposition:</u> $f : X \to Y$ is convex with respect to S if f is lower semicontinuous with respect to S and satisfies

$$f\left(\tfrac{1}{2}x + \tfrac{1}{2}y\right) \leq \tfrac{1}{2}f(x) + \tfrac{1}{2}f(y). \,‖$$

The analagous results hold for strict convex type functions.

[41] If X' denotes the topological dual of X then the <u>dual cone</u> S^+ is defined by

$$S^+ = \{x^+ \in X' \mid x^+(x) \geq o \ \forall x \in S\}.$$

S^+ is a closed convex cone even if S is an arbitrary set.

The <u>second dual</u> $(S^+)^+$ is defined by

$$(S^+)^+ = \{x \in X \mid x^+(x) \geq o \ \forall x^+ \in S^+\}.$$

If S is a closed convex cone a standard separation arguement (for cones) shows that $(S^+)^+ = S$. The dual cone can be used to give the following important characterization of convexity.

[42] <u>Proposition:</u> $f : X \to Y$ is convex with respect to a closed convex cone S if and only if $u^+ f : X \to R$ is convex for every $u^+ \in S^+$.

Proof: $\Rightarrow$ If f is convex with respect to S, then for $0 \leq \lambda \leq 1$, and for $u^+ \in S^+$

$$u^+ \left[\lambda f(x) + (1-\lambda) f(y)\right] \geq u^+ \left[f(\lambda x + (1-\lambda) y)\right]$$

which asserts the convexity of $u^+ f$.

$\Leftarrow$ If $u^+ f$ is convex $\forall u^+ \in S^+$ then

$$u^+ \left[\lambda f(x) + (1-\lambda) f(y) - f(\lambda x + (1-\lambda) y)\right] \geq 0$$

and $f(\lambda x + (1-\lambda) y) - [\lambda f(x) + (1-\lambda) f(y)] \in (-S^+)^+$ which since S is a closed convex cone means that f is convex with respect to $S = (S^+)^+$. ∎ Unfortunately the same characterization of strongly quasiconvex functions breaks down.

[43] Example: Let $f : R \to R^2$ be defined by $f(x) = (x, -x^3)$ with the orthant ordering P. Then f is strongly quasiconvex but if $u^+ = (1,1) \in P^+ = P$.

$$u^+ f = x - x^3$$

which is clearly not quasiconvex.

[44] Another property of convex functions (or sets) is that of <u>local convexity</u>. A set $A \subset X$ is said to be <u>locally convex</u> if for each $x \in A$ there is a neighbourhood N of x with $A \cap N$ convex. A result of Kelley and Namioka (1963) says that in a Hausdorff topological vector space a closed connected locally convex set is convex.

[45] Defining the <u>epigraph</u> of $f : X \to Y$ <u>with respect to B</u> by $\mathrm{Epi}_B f = \{(x,y) \mid y - f(x) \in B\}$ one has:

[46] Proposition: $f : X \to Y$ is convex with respect to S if and only if $Epi_s f$ is a convex set.

Proof: $\Rightarrow$ Suppose (x_1, y_1) and $(x_2, y_2) \in Epi_s f$.

Then $f(x_1) \leq y_1$, $f(x_2) \leq y_2$ and since f is convex

$$f(\lambda x_1 + (1-\lambda) x_2) \leq \lambda f(x_1) + (1-\lambda) f(y_1) \leq \lambda y_1 + 1 - \lambda y_2$$

and $(\lambda x_1 + (1-\lambda) x_2, \lambda y_1 + (1-\lambda) y_2) \in Epi_s f$ which must be convex.

$\Leftarrow$ $(x_1, f(x_1))$ and $(x_2, f(x_2)$ belong to $Epi_s f$ which being convex means that

$$(\lambda x_1 + (1-\lambda) x_2, \lambda f(x_1) + (1-\lambda) f(x_2)) \in Epi_s f.$$

By the definition of $Epi_s f$

$$f(\lambda x_1 + (1-\lambda) x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2). \blacksquare$$

[47] Using the above proposition and the result of Kelley and Namioka quoted in [44] one derives:

Proposition: $f : X \to Y$ is convex with respect to S if and only if f satisfies

$$f(\lambda x_1 + (1-\lambda) x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2)$$

whenever x_1 and x_2 belong to some neighbourhood N.

Proof: The condition is equivalent to the local convexity of $Epi_s f$. $\blacksquare$

[48] This result is clearly not true for quasiconvex functions as is seen by $f(x) = \begin{cases} 1 - x^2 & |x| > 1 \\ 0 & |x| \leq 1 \end{cases}$

In fact with $N = \{x \mid |x| \leq 1\}$ f is locally quasiconvex but it is not quasiconvex.

[49] The definition in [45] could be rephrased to include a restricted domain C. Since this can equally well be done by redefining f, it seems simpler to leave it as it is.

[50] Proposition: $f : X \to Y$ is strongly quasiconvex if and only if $\mathrm{Epi}_s f$ has the following property:

If (y, x_1) and $(y, x_2) \in \mathrm{Epi}_s f$ then $(y, \lambda x_1 + (1 - \lambda) x_2)$ does for $o \leq \lambda \leq 1$.

Proof: This is immediate from the definitions of the epigraph and quasiconvexity. ▌

[51] A set $B \subset Y$ is said to minimizable with respect to a cone S if there is some $z \in Y$ with $b - z \in S \ \forall b \in B$. Maximizability is defined dually.

[52] Proposition: If S is a cone with non empty interior in a convex space then all bounded sets are minimizable and maximizable with respect to S. The maximizing and minimizing points can be taken in $\pm S$.

Proof: Since B is bounded it is absorbed by all neighbourhoods. In particular if $s \in S^o$, there is an open set N with $\lambda_o > o$ and

$$s + \lambda_o B \subset s + N \subset S$$

where the first containment follows from the boundedness of B and the second from $s \in S^o$. Since S is a cone

$$\lambda_o^{-1} s + B \subset S$$

which rewritten says $b \geq -\lambda_o^{-1} s$ whenever $b \in B$. Applying the same argument to $-B$ one sees that B is maximizable. ▌

[53] The results of [27] and [52] can be used to generalise the standard result (Luenberger (1969)) that a convex function on R is continuous throughout the relative interior of its domain if it is continuous anywhere. For simplicity X and Y are assumed normed. By the <u>relative interior</u> of a set C (denoted riC) one means the interior with respect to the smallest closed affine subspace containing C. It is a standard result that for a finite dimensional convex set $C, riC \neq \emptyset$.

<u>Proposition</u>: Let $f : X \to Y$ be convex with respect to S which is normal with interior. Then if C is a convex set in domf and f is continuous at $x_0 \in riC$ f is continuous throughout riC.

<u>Proof</u>: Suppose without loss that $f(0) = 0$, $x_0 = 0$ and $riC = C^o$. Let $\epsilon_0 > 0$ be given. Then $\exists \delta_0$ such that $\| x \| < \delta_0$ implies that $\| f(x) \| < \epsilon_0$ since f is continuous at o.

Let $y_0 \in C^o$. $\exists \beta > 1$ with $\beta y_0 \in C^o$.
If $\| z - y_0 \| < (1 - \beta^{-1}) \delta_0$ then $z = y_0 + (1 - \beta^{-1})\, x$ with $\|x\| < \delta_0$, $x \in C$ and $f(z) \leq \beta^{-1} f(\beta y_0) + (1 - \beta^{-1}) f(x)$.
Since,

$\| f(x) \| \leq \epsilon_0$ when $\|x\| < \delta_0$ there is, using [52], some d such that $f(x) \leq d \;\forall x$ with $\|x\| < \delta_0$ and

$$f(z) \leq \beta^{-1} f(\beta y_0) + (1 - \beta^{-1})\, d = d(y_0)$$

if $\| y_0 - z \| < (1 - \beta^{-1}) \delta_0$. For $0 < \epsilon < 1$ one has

$$(z + y_0) = \epsilon(y_0 + \epsilon^{-1} z) + (1 - \epsilon)\, y_0$$

and by convexity

$$f(z + y_o) - f(y_o) \leq \epsilon \left[f(\epsilon^{-1} z + y_o) - f(y_o)\right].$$

So if $\| z \| < \epsilon (1 - \beta^{-1}) \delta_o$

(1) $f(z + y_o) - f(y_o) \leq \epsilon \left[d(y_o) - f(y_o)\right] = \epsilon a$.

Moreover, since $y_o = \frac{1}{1+\epsilon} (z + y_o) + (1 - \frac{1}{1+\epsilon}) (-\epsilon^{-1} z + y_o)$

$$f(y_o + z) - f(y_o) \geq - \epsilon\left[f(y_o - \epsilon^{-1} z) - f(y_o)\right]$$

which as before gives

(2) $f(y_o + z) - f(y_o) \geq - \epsilon a$

Combining (1) and (2) gives $- a\epsilon \leq f(y_o + z) - f(y_o) \leq a\epsilon$

whenever $\| z \| \leq \epsilon(1 - B^{-1})\delta_o$. By the result of [28]

the set $A = \{x \mid - a \leq x \leq a\}$ is bounded because S is normal and one has

$$f(y_o + z) - f(y_o) \in \epsilon A \text{ when } \| z \| < \epsilon (1 - \beta^{-1}) \delta_o$$

and f is continuous throughout riC. ∎

[54] By the result in [29] any pointed cone in R^n is normal and in particular [53] generalises Luenberger's proposition. Also the condition that $S^o \neq \emptyset$ is needed for [52] and it plus the boundedness of A are, by [28] equivalent to normality of S.

[55] The next proposition relates continuity off and interior points of $Epi_s f$ when X and Y are normed spaces.

Proposition: If S is a normal cone with interior and $ri(\text{dom } f) \neq \emptyset$ then f is continuous at x_o if and only if $(x_o, y_o) \in ri\ Epi_s f$ for some y_o.

Proof: $\Leftarrow$ Without loss $x_o = o = f(o)$ and replacing X by V(C) the variety (affine subspace) spanning domf one can assume that $(\text{domf})^o \neq \phi$. Suppose $(0, y_o) \in \text{ri Epi}_s f$, then since

$$V(\text{Epi}_s f) = V(C) \times Y$$

(o, y_o) may be supposed interior to $\text{Epi}_s f$.

Thus, δ_1 and $\delta_2 > o$ exist with $(x,y) \in \text{Epi}_s f$ when

$$\|x - o\| < \delta_1 \text{ and } \|y_o - y\| < \delta_2.$$

As in [53]

$$-\lambda f(-\lambda^{-1} x) \leq f(x) \leq \lambda f(\lambda^{-1} x).$$

Also

$$f(x) \leq y \text{ when } \|y - y_o\| < \delta_2 \text{ and } \|x\| < \delta_1.$$

Since neighbourhoods in normed spaces are bounded and $S^o \neq \phi$, $N = \{y \mid \| y - y_o\| \leq \delta_2\}$ is maximizable and there is some a_o with $f(x) \leq a_o$ if $\|x\| < \delta_1$.

Thus if $\|x\| < \lambda\delta_1$

$$-\lambda a_o \leq f(x) \leq \lambda a_o .$$

Since S is normal the same argument as in [53] shows that f is continuous at o. $\Rightarrow$ Let $\epsilon > o$ be given and suppose f is continuous at $o \in (\text{domf})^o$. Then $\|x\| < \delta$ implies that $\|f(x)\| < \epsilon/2$ and by [52] there is some $s_2 \in S$ with

$$-s_2 \leq f(x) \leq s_2.$$

Let $s_1 \in S^o$ with N_ϵ such that $s_1 + N_\epsilon \subset S$ then $f(x) < s_1 + s_2 = s_o \in S$ when $\|x\| < \delta$.

If $\|y - s_o\| \leq \epsilon/2$ then

$$f(x) - y = (f(x) - s_o) + (s_o - y)$$

$$\subset N_{\epsilon/2} - s_o + N_{\epsilon/2} \subset N_\epsilon - s_o \subset -S$$

Thus if $\|x - x_0\| \leq \delta$ and $\|y - s_0\| < \epsilon/_2$ $(x,y) \in \text{Epi}_s f$ and $(x_0, s_0) \in \text{ri Epi}_s f$ as desired.∎

[56] The continuity results above can be used to generalise a theorem of Rockafellar's (1970).

<u>Theorem</u>: Let $f : X \to Y$ be convex with respect to a convex cone S with $S^o \neq \emptyset$. Let X, Y be normed spaces with X reflexive and let B be a weakly compact subset of the relative interior of domf (W) in the weak topology. If f is weakly continuous on W then f is Lipshitzian relative to B. In particular if X is finite dimensional and f is continuous at some point of ri(domf) the result holds.

<u>Proof</u>: By restricting attention to the variety spanned by domf one can assume that W is the weak interior of domf. Let $U \subset X$ be the unit ball; then U is weakly compact since X is reflexive and hence so is

$$\frac{1}{n} U + B \quad \forall n \in N \text{. Moreover}$$

$$\bigcap_{n \in N} \left(\frac{1}{n} U + B\right) \cap \widetilde{W} = \emptyset .$$

This is the intersection of weakly compact sets so there is some n with $B + \frac{1}{n} U \subset W$.

Since f is assumed continuous on domf and $B + \frac{1}{n} U$ is weakly compact $f(B + \frac{1}{n} U)$ is weakly bounded and thus both minimizable and maximizable, since $S^o \neq \emptyset$.

Let $b_1 \leq f(B + \frac{1}{n} U) \leq b_2$ and let $x, y \in B$ $x \neq y$. Then $z = y + \frac{1}{n} \|y - x\|^{-1} (y - x) \in B + \frac{1}{n} U$ and $y = \lambda z + (1 - \lambda) x$ with $\lambda = \|y - x\| / \frac{1}{n} + \|y - x\|$. Since f is convex with respect to S

$$f(y) \leq (1-\lambda)\ f(x) + \lambda f(z)$$

and

$$f(y) - f(x) \leq \lambda(b_2 - b_1).$$

Since x, y are interchangable

$$f(y) - f(x) \in \{w \mid -\lambda(b_2 - b_1) \leq w \leq \lambda(b_2 - b_1)\}$$

This last set is contained in D where D is

$$D = \{w \mid -\|x - y\|k \leq w \leq \|x - y\|k\}$$

and $k = n(b_2 - b_1)$.

By a result of Kelley and Namioka (1963) there is $L > o$ (since S is normal) with

$$\alpha\|w\| \leq k\|x - y\| \quad \forall w \in W.$$

In particular for $x, y \in B$

$$\|f(x) - f(y)\| \leq L^{-1}\ k\|x - y\|.$$

The second conclusion follows from [53] and [55].▌

Rockafellar's initial result was the case $X = R^m$ $(Y,S) = (R,R^+)$. If $X = R^m$ and $Y = R^n$ with S any pointed cone with interior then it is clear that f is Lipshitz on any closed bounded set in ri(domf).

[57] <u>Examples of discontinuous behaviour</u>

(1) Let $D = \{x \in C\ [0,1] \mid dx/_{dt} \in C\ [0,1]\}$ and let $A : D \to C\ [0,1]$ be defined by $Ax = dx/_{dt}$. A is discontinuous but is convex with respect to any cone S.

(2) $f : R \to l_\infty$ (with the orthant ordering) given by $f(r) = \{r^{2n}\}$ is convex with $\text{domf} = \{r \mid |r| \leq 1.\}$. In this case f is continuous if $|r| < 1$ but is discontinuous at 1 since $\|f(1 - \frac{1}{n}) - f(1)\| = \sup_k \|(1 - \frac{1}{n})^{2k} - 1\|$ which is 1.

[58] Although convex functions share the property of [53] with linear functions a convex function can be continuous at a point and not weakly continuous.

Let $f; l_2 \to l_2$ (with the orhant orderings) be given by $f(\{x_n\}) = \{x_n^{2n}\}$. f is clearly convex with dom$f = l_2$. It is reasonably simple to verify continuity especially at the origin. f is not weakly continuous since $\{\{x_{nk}\}\} = \{\{2\delta_{nk}\}\}$ is weakly convergent but $\{f(\{x_{nk}\})\} = \{\{(2\delta_{nk})^{2n}\}\}$ is not since $u^+ = \{\frac{1}{n}\} \in l_2^+ = l_2$ and

$$u^+ \cdot f(\{x_{nk}\}) = \sum_n \frac{1}{n}(2\delta_{nk})^{2n} = 2^{2k}/k \longrightarrow \infty .$$

Note that S is normal but has no interior.

[59] The following partial analogues of the linear situation do hold:

Proposition: If $f = X \to Y$ is strongly quasiconvex with respect to S and lower semicontinuous over C, a closed convex set then f is weakly lower semicontinuous over C.

Proof: $\{x | f(x) \leqslant z\} \cap C$ is a closed convex set and thus weakly closed. ∎

[60] More interestingly one has

Proposition: If $f = X \to Y$ is strongly quasiconvex with respect to S and lower semicontinuous sequentially on any convex set C then f is weakly lower semicontinuous on C.

Proof: Let $\{x_n\} \subset C$ be a sequence with limit $x_o \in C$. Since x_o belongs to the weak closure of the convex hull of $\{x_k\}_{k=n}^{\infty}$ for any n, x_o actually belongs to the closure of the convex hull. Thus there is for each n a point z_n and scalars $\lambda_{nk} \geqslant 0$ with

$\sum_{k}^{m_n} \lambda_{nk} = 1$ and

$$z_n = \sum_{k=n}^{m_n} \lambda_{nk} \; x_k \to x_o.$$

Since f is strongly quasiconvex over C

$$f(z_n) \leqslant \max_{n \leqslant k \leqslant m_n} f(x_k) .$$

Hence if $f(x_k) \leqslant_S z_o$ and $x_k \to x_o$ one has $f(z_n) \leqslant_S z_o$ when $n \geqslant n_o$.

Because $z_n \to x_o$ and f is sequentially semicontinuous one has $f(x_o) \leqslant z$ and f is weakly sequentially lower semicontinuous on C.∎ This last result generalizes Daniel (1971) and provides at least a partial analogue to the equivalence of weak and strong continuity of linear maps.

[61] Since in the result of [53] it is only to prove that bounded sets are maximizable that $S^o \neq \phi$ is used, one could have required only the former condition. This condition would not be much weaker because the following partial converse to [52] holds.

<u>Proposition</u>: If $S \subset Y$ is a <u>generating</u> cone, that is $S - S = Y$, and Y is normed then $S^o \neq \phi$.

<u>Proof</u>: Since Y is normed the unit ball U is bounded. Let $y_o \in Y$ be a maximizer for U. Then

$$y_o \geqslant u \quad \forall u \in U.$$

Since $S - S = Y$ $y_o = s_1 - s_2$ $s_1, s_2 \in S$ and

$$s_1 + U \subset s_2 + S \subset S$$

so that $s_1 \in S^o$.∎

[62] In further reference to [53], [55] it is apparent that if S is not pointed it is unreasonable to expect continuity since in the nonpointed case there is at least one direction in which the behaviour of f is not restricted at all. It is also apparent that

no sort of simple continuity result should be expected for quasiconvex mappings even on the line. Certainly dense countable sets of discontinuities can exist because any monotone real valued mapping is quasiconvex. On the line this is the worst that can happen because:

Proposition: (Stoer and Witzgall (1970)) A function $f: C \to R$, where C is convex, is (quasi)convex if and only if its restriction to any line segment is (quasi)convex. Moreover, $f: R \to R$ is quasiconvex if and only if there is a partition of R into two (possibly disjoint) intervals (I_1, I_2) with f non-increasing on I_1 and non-decreasing on I_2.|

In particular this means that if $f: R \to R$ is quasiconvex it is of bounded variation and has at worst countable discontinuities. This also implies that the set of discontinuities of any real valued quasiconvex map has no interior.

[63] Boundedness of level sets

It is a well known result in R^n that any closed unbounded convex C contains all half lines of the form $x + th$, $t \geq o$ where h is some fixed non zero point and x is any member of C. h is called a direction of recession. An unbounded convex set in general need not have any such directions.

Example: Let X be l_∞ and let C be the convex set defined by $C = \{\{a_k\} \mid |a_k| \leq k, k = 1, 2 \ldots\}$ which is clearly unbounded and closed. Suppose that $a + tb$, $t \geq o$ was a line segment in C, $a \in C$, $b \neq o$. Then for some b_k, b_k is non zero and for $t > t_o$ $|a_k + tb_k| > k$ which means $a + tb$ is not in C.|

[64] The next result, which generalizes a proof of Stoer and

Witzgall (1970) is phrased in R^n because it relies on the existence of half lines in unbounded sets.

Proposition: Let $f: R^n \to Y$ be lower semicontinuous and convex with respect to a generating closed convex cone S. Either all nonvoid level sets are bounded or they are all unbounded.

Proof: Suppose the level sets $S(z) = \{x \mid f(x) \leq z\}$ are such that $S(z_1)$ is bounded and $S(z_2)$ is unbounded. Since f is supposed convex and lower semicontinuous all the level sets are closed and convex in R^n. Hence there is $h_2 \neq o$ with $x + th_2 \in S(z_2)$ $\forall\, x \in S(z_2)$ and $t \geq o$.

If z_3 is chosen such that $z_3 \geq z_2$, $z_3 \geq z_1$, which can be done since S is generating, $S(z_3)$ is unbounded because it contains $S(z_2)$.

Let $x \in S(z_1)$. Then $x \in S(z_3)$ and since h_2 is a direction of recession for $S(z_2)$ it is also one for $S(z_3)$ and $x + th_2 \in S(z_3)$ $\forall t \geq o$. Thus

$$f(x + th_2) \leq \lambda f(x) + (1-\lambda) f(x + \frac{t}{1-\lambda} h_2)$$
$$\leq \lambda z_1 + (1-\lambda) z_3 \quad \text{if } o < \lambda < 1.$$

Letting $\lambda \to 1$

$$f(x + th_2) \leq z_1 \quad \forall t \geq o, \text{ since S is closed.}$$

This shows that $S(z_1)$ is unbounded and a contradiction has been established.▮

[65] For quasiconvex functions the result in [64] is not true. The next proposition which generalizes Lemma 4.9.7. of Stoer and Witzgall (1970) is a partial clarification.

Definition: A chain (linear ordering) is order complete if any subset B of A which can be maximized has a supremum in B. That is, there exists z_1 such that if $z \geq b$ $\forall b \in B$ then $z \geq z_1$ and

$z_1 \geqslant b \quad \forall b \in B$.

[66] Proposition: Let $f: R^n \to Y$ be lower semicontinuous and strongly quasiconvex with respect to S. Let A be an order complete chain in a sequential topological space Y. Suppose that any descending convergent sequence in A has limit in A. If the level sets $\{S(a) \mid a \in A\}$ contain both bounded and unbounded sets there is an $\bar{a} \in A$ with $S(a)$ unbounded exactly when $a \geqslant \bar{a}$.

Proof: Suppose $S(a_2)$ is unbounded and $S(a_1)$ is bounded. Let $\bar{a} = \inf\{a \mid S(a) \text{ is unbounded}\}$. Then $a_1 \lneqq \bar{a} \leqslant a_2$. It remains to show $S(a)$ is unbounded. Let $T(a)$ denote the directions of recession of $S(a)$ with norm one. These sets are compact and, since $T(a) \subset T(b)$ if $a \leqslant b$, have the finite intersection property for $a > \bar{a}$. Hence, $\bigcap_{a > \bar{a}} S(a) \neq \phi$. Let h be in this intersection.

For each $a > \bar{a}$ and for $x \in S(\bar{a})$ one has (1) $f(x + th) \leqslant a \quad \forall t \geqslant o$. The assumptions on A imply that there is a sequence $\{a_n\}$, $a_n \geqslant \bar{a}$, with $a_n \to \bar{a}$. This with (1) shows $S(\bar{a})$ is unbounded. ▮

Taking $f: R \to R$, $f(x) = \begin{cases} 1 & |x| \geqslant 1 \\ x^2 & |x| < 1 \end{cases}$ one sees that $S(r)$ is unbounded if and only if $r \geqslant 1$. Setting $A = \{o\} \cup \{x \mid x > 2\}$, which is an order complete chain which does not contain its limit points one has an example in which the proposition does not hold.

[67] Differential characterizations of convex type functions

For the most part differential conditions will be introduced as they prove necessary. The following few propositions are given for the sake of completeness.

Definition: If $f: X \to Y$ is a mapping between two convex spaces then f is said to be β-*differentiable at x* *with respect to a family* β of sets in X if there is a continuous linear transformation $f'(x_o)$: X → Y with

$$t^{-1}\left[f(x+th) - f(x)\right] - f'(x_o)(h) \to o \text{ as } t \to o$$

uniformly in the topology defined by β.

f is said to be (1) *compactly differentiable* if β is all sequentially compact sets and (2) *boundedly differentiable* if β is all bounded sets. In particular these notions agree in *sequential Montel spaces*. In normed spaces (2) is just *Fréchet* differentiation.

[68] *Proposition*: If f: X → R is boundedly or compactly differentiable then f is quasiconvex if and only if

$$f'(x_o)(x - x_o) \leq o \text{ when } f(x) \leq f(x_o).$$

Proof: This is proved in Ponstein (1967) for X = R^n and Fréchet differentiation. The differences are entirely technical since any sensible β-derivative will suffice in his proof. ▌ From now on when any reasonable derivative will do it will just be called differentiable.

[69] *Proposition*: f: X → Y is convex with respect to a closed cone S if and only if

$$f'(x_o)(x - x_o) \leq f(x) - f(x_o) \quad \forall x, x_o \in X.$$

Proof: The result when Y = R is standard. In the general case by [42] $u^+ f$ is convex (and differentiable) $\forall u^+ \in S^+$.

By the linear result this is equivalent to

$$u^+(f(x) - f(x_o)) \geq u^+ f'(x_o)(x - x_o) \quad \forall u^+ \in S^+.$$

Since S is closed this last inequality gives as equivalent

$$f(x) - f(x_o) \geq_S f'(x_o)(x - x_o). ▌$$

It does not appear that the condition

$f'(x_o)(x - x_o) \leqslant_S o$ whenever $f(x) \leqslant_S f(x_o)$ is equivalent to (strong) quasiconvexity although by a direct derivative argument it is implied by quasiconvexity.

[70] A condition which will be of primary importance in optimization is

$$f'(x_o)(x - x_o) \in -S \text{ implies } f(x) - f(x_o) \in S \quad \forall x \in A$$

which is called <u>pseudo-convexity</u> over A at x_o. It is possessed by convex functions ([69]). The function $f: R \to R$ given by $f(x) = \begin{cases} -x^2 & x \in [o,1] \\ \infty & x \notin [o,1] \end{cases}$ is quasiconvex but not pseudo-convex when the derivatives are taken to be one sided at o and 1.

[71] A useful relationship which simplifies proofs given in Guignard (1969) and Cottle and Ferland (1970) is:

<u>Proposition</u>: Let $f: X \to R$ be quasiconvex on a convex set C and differentiable at x_o. Suppose that for some $y \in C$ $f'(x_o)(y-x_o) > o$: then f is pseudo-convex at x_o on C.

<u>Proof</u>: Suppose $f'(x_o)(x - x_o) \geqslant o \quad \forall x \in C$.

Then $(1 - \lambda) f'(x_o)(x - x_o) + \lambda f'(x_o)(y - x_o) > o \quad \forall x \in C$.

Since $x, y \in C$ $x_\lambda = \lambda y + (1 - \lambda) x_o \in C$ and

$$f'(x_o)(x_\lambda - x_o) > o.$$

Since f is supposed quasiconvex on C

$$f(x_\lambda) > f(x_o).$$

The continuity of f implies that $f(x) \geqslant f(x_o)$. ∎

Guignard's case was $C = X$ and $f'(x_o) \neq o$ while Cottle and Ferland had $X = R^n$, $C = R^{n+}$, $f'(x_o) \neq o$. It is easy to show that these are both subsumed.

[72] The natural condition

$$f'(x_o)(x - x_o) \leqslant f(x) - f(x_o) \text{ if } f(x) \leqslant f(x_o)$$ is strong enough to imply convexity in most cases. Precisely one has:

Proposition: Let $f: X \to R$ be twice Fréchet differentiable with $f''(x)$ continuous in x then the above condition implies that f is convex.

Proof: Consider $X = R$. By Taylor's theorem

$$f(\lambda y + (1 - \lambda)x) - f(x) = \lambda f'(x)(y - x) + \tfrac{1}{2}\lambda^2 f''(x_\lambda)(y - x)^2$$

with $x_\lambda \to x$ as $\lambda \to o$.

The condition of the hypothesis implies quasiconvexity ([68]) and thus for $o < \lambda < 1$ and $f(y) \leqslant f(x)$ one has

$$f(\lambda y + (1 - \lambda)x) \leqslant f(x) \text{ and } f(\lambda y + (1 - \lambda)x) - f(x) \geqslant f'(x)\lambda(y-x).$$

Thus

$$\tfrac{1}{2}\lambda^2 f''(x_\lambda)(y - x)^2 \geqslant o.$$

On dividing by λ^2 and letting $\lambda \to o$ one gets $f''(x)(y - x)^2 \geqslant o$. Now, if $f(x) \neq \inf f(y)$ there is some $y \neq x$ with $f(y) \leqslant f(x)$. In this case $(y - x)^2 > o$ and $f''(x) \geqslant o$.

Otherwise, let $\{x_n\}$ be a sequence of points with $x_n \neq x$ $x_n \to x$. By the definition of x, $f(x) \leqslant f(x_n)$ and thus $f''(x_n)(x_n - x)^2 \geqslant o$. As before $f''(x_n) \geqslant o$. Letting $x_n \to x$ $f''(x) \geqslant o$ because f'' is continuous.

Consider now $g(\lambda) = f(x + \lambda y)$ for fixed $x, y \in X$. It is immediately verifiable that g satisfies the conditions and hence $g(\lambda) = f(x + \lambda y)$ is convex for any x, y. This, using the first part of the proposition in [62], implies that f is convex. ∎

Concepts of minimization with respect to cones

There is a profusion of possible extensions to the notion of the minimum of a real valued function over a set A. The two most useful and possibly most natural are defined below. S is always assumed to be a closed convex cone.

[73] Definition: $f: X \to Y$ is said to have a strong minimum (with respect to S) over A at x_o if $f(x) - f(x_o) \in S \quad \forall x \in A$.

[74] Definition: $f: X \to Y$ is said to have a weak mimimum (with respect to S)over A at x_o if $f(x) - f(x_o) \notin -S^o$ when $x \in A$.

If $S^o = \phi$ any point is a weak minimum so from now on the interior of S will be assumed nonvoid when weak minima are being discussed.

[75] Proposition (1) Any strong minimum is a weak minimum.

(2) If S is pointed any two strong minima agree in value.

(3) If C is convex and x_o is a strong minimum f or f over C and if f is quasiconvex with respect to S, a pointed cone, then $M = \{x \mid f(x) = f(x_o) = \text{strong min} \{f(x) \mid x \in C\}\}$ is convex.

Proof: Only (3) is not immediate. If $x_1, x_2 \in M$ and $o \leq \lambda \leq 1$ then $\lambda x_1 + (1 - \lambda)x_2 \in C$ and $f(\lambda x_1 + (1 - \lambda)x_2) \leq f(x_1)$ by quasiconvexity. By (2) and the definitions

$$f(\lambda x_1 + (1 - \lambda)x_2) = f(x_o)$$

and $\lambda x_1 + (1 - \lambda)x_2) \in M$. ▌

[76] Proposition: Let $f: X \to Y$ be fully lower semicontinuous with respect to S and let A be closed then the set of weak minima for f over A with respect to S is closed.

Proof: Suppose $\{x_n \mid n \in N\}$ is a net of weak minima with

$x_n \to x_o$. Since A is closed $x_n \in A$ implies $x_o \in A$. Suppose for some x in A $f(x) < f(x_o)$. By full lower semicontinuity $f(x) < f(x_n)$ if $n \geq n_o$ since S^o is an open set. This contradicts the minimality of x_n. ∎

Similarily one has:

77 Proposition. If f: $X \to Y$ is lower semicontinuous with respect to S and A is closed the set of strong minima for f over A is closed. ∎

78 If one wishes to guarantee the convexity of the set of weak minima over a convex set quasiconvexity is too weak. The condition stated below seems artificial but it is equivalent to quasiconvexity when $(Y,S) = (R,R^+)$.

Proposition: The set of weak minima over C is convex if f: $X \to Y$ satisfies: Whenever $x, y, x_o \in C$ and for some $o < \lambda < 1$ $f(\lambda x + (1 - \lambda)y) < f(x_o)$ then $f(x) < f(x_o)$ or $f(y) < f(x_o)$. ∎

79 If X is a topological space then x_o is called a local minimum for f over A with respect to S if x_o is a minimum over $A \cap N$ for some neighbourhood N. If A = X = N the minimum is called global. Ponstein's (1967) result that every local minimum of a real valued (P) strictly quasiconvex mapping is global has the following extensions.

80 Theorem: If f: $X \to Y$ is (P) strictly quasiconvex with respect to S then every local weak minimum with respect to S is global.

Proof: Suppose x_1 is a local non global minimum. Then there is some $x_2 \in X$ with $f(x_2) - f(x_1) \in -S^o$. Let $x_\lambda = \lambda x_2 + (1 - \lambda)x_2$.

For λ sufficiently small x_λ will belong to the neighbourhood over which x_1 is a weak minimum.

Thus $f(x_\lambda) - f(x_1) \notin -S^o$ for $o < \lambda < \lambda_o$.

But by (P) strict quasiconvexity ([16])

$$f(x_\lambda) - f(x_1) \in -S^o$$

since

$$f(x_2) - f(x_1) \in -S^o$$

and a contradiction has been derived. ▮

For strong local minima one has (dually)

[81] Theorem: If f is absolutely quasiconvex ([11]) with respect to S then every strong local minimum is global.

Proof: Let x_1 be a local minimum and let $x_o \in X$. For $o < \lambda < \lambda_o$ $f(\lambda x_1 + (1-\lambda)x_o) \geqslant f(x_1)$.

By [11] $f(x_o) \geqslant f(x_1)$ and x_1 is a global minimum. ▮

If $(Y,S) = (R,R^+)$ then by [21] both [80] and [81] agree with Ponstein's result. When this is the case one can in fact show that if every local minimum of a quasiconvex function is global the f is (P) strictly quasiconvex.

More generally for any convex space X one has

[82] Proposition: If $f: X \rightarrow Y$ is quasiconvex with respect to a pointed cone S, and x_2 is any nonglobal but local strong minimum then $f(x)$ is constant on $L(\epsilon) = \{x \mid x = \lambda x_1 + (1-\lambda)x_2, o \leqslant \lambda \leqslant \epsilon\}$ where x_1 is any global minimum and ϵ is some positive number.

Proof: Since x_2 is a local minimum there is $\epsilon > o$ with $f(\lambda x_1 + (1-\lambda)x_2) \geqslant f(x_2)$ if $o \leqslant \lambda \leqslant \epsilon$. Since $f(x_1) \leqslant f(x_2)$ and f is quasiconvex

$$f(\lambda x_1 + (1-\lambda)x_2) \leqslant f(x_2) \text{ if } o \leqslant \lambda \leqslant 1.$$

Thus for $x \in L(\epsilon)$ $f(x) = f(x_2)$ since S is pointed. ▮

Weak and strong maximization are defined dually to minimization. Any convex minimization result yields a dual maximization result for concave type functions. Thus it is usually unnecessary to consider both maximization and minimization problems. The next results, however, give information about maxima of quasiconvex functions.

[83] Theorem: If $f: X \to Y$ satisfies

(1) If $f(x) \lneq f(z)$ and $f(y) \leq f(z)$ then

$$f(\lambda x + (1-\lambda)y) \lneq f(z) \quad \text{for } o < \lambda \leq 1,$$

then if f achieves its strong maximum over C, a convex set contained in domf, at $x_o \in$ riC then f is constant on C.

Proof: Let $z \in$ riC with $f(z) \geq f(y)$ $\forall y \in C$ and let $x \in C$. Then there is some $y \in C$ and $o < \lambda_o < 1$ with $z = \lambda_o x + (1-\lambda_o)y$. (Otherwise z would be a boundary point). Now, if $f(y) \leq f(z)$ and $f(x) \neq f(z)$ one has, since $f(x) \leq f(z)$, that $f(\lambda x + (1-\lambda)y) \lneq f(z)$. This is impossible for $\lambda = \lambda_o$ and $f(x)$ must, therefore be equal to $f(z)$. Thus f is constant on C. ▮

[84] Corollary: If $f: X \to Y$ is such that either

(1) f is convex

or

(2) $(Y,S) = (R,R^+)$ and f is (P) strictly quasiconvex and satisfies the one point exclusion property then the result holds.

Proof: For (1) it is easily verified that any convex f satisfies (1) of [83], while for (2) the proposition of [33] proves that the property is satisfied. ▮

The proof method of [83] is derived from a result of Rockafellar (1970a) which is in fact the corollary of [84] in case (1) with $(Y,S) = (R,R^+)$.

For any quasiconvex function it is simple to show the following result.

[85] Proposition: If $f: X \to Y$ is quasiconvex with respect to S and if A is any set over which f has a strong maximum at x_o then x_o is a strong maximum for f over the convex hull, C, of A.

Proof: $A \subset \{x \mid f(x) \leq f(x_o)\} = S(f(x_o))$.

Since f is quasiconvex $S(f(x_o))$ is convex and thus

$C \subset S(f(x_o))$.

Equivalently $f(x) \leq f(x_o) \quad \forall x \in C$. ∎

[86] Definitions have been made of convex like conditions which do not require that one be in a vector space. In particular, a function $f: X \to R$ is called pathwise connected if whenever $x, y \in X$ there is an arc $p(t)$ with $p(o) = x$ and $p(1) = y$ and with $f(p(t)) \leq \max(f(x),f(y))$. Strict pathwise connectedness is defined similarily. Many of the previous results hold for strict pathwise connectness. For example local minima are still global. The properties defined in [8] through [16] could all be extended analogously but the difficulty in verifying the connectedness of a function and its relative inutility because of this suggest that the effort is not worthwhile.

[87] If f is both quasiconvex and quasiconcave with respect to S f is called quasiaffine while if it is both (P) strictly quasiconcave and quasiconvex as well it is called strictly quasiaffine.

If $(Y,S) = (R,R^+)$ then Stoer and Witzgall (1970) have shown that the set of maxima for f over a convex set C consists exclusively of extreme points. The result can be seen to hold for any strictly quasiaffine function with respect to a pointed cone. Note that on the line if f is (P) strictly quasiconvex and P strictly quasiconcave it is automatically strictly quasiaffine.

Multivalued convex and quasiconvex functions

In another direction to the notions previously discussed lies the idea of multivalued quasiconvexity and convexity. As will be seen later many standard multiplier theorems can be painlessly extended to cover multivalued functions.

The sequel gives the definitions which will be used and some propositions.

[88] Definition: $F: X \to Y$ is said to be a multivalued convex function with respect to S if $F(x)$ is a subset of Y for each $x \in X (F: X \to 2^Y)$ and if whenever $y_1 \in F(x_1)$, $y_2 \in F(x_2)$ and $0 \leqslant \lambda \leqslant 1$ there is some $y_\lambda \in F(\lambda x_1 + (1-\lambda)x_2)$ with $y_\lambda - [\lambda y_1 + (1-\lambda)y_2] \in -S$.

[89] Definition: $F: X \to Y$ is multivalued (P) strictly quasiconvex with respect to S if whenever $0 < \lambda < 1$, $y_1 \in F(x_1)$, $y_2 \in F(x_2)$ with $y_1 < y_2$ there is some $y_\lambda \in F(\lambda x_1 + (1-\lambda)x_2)$ and with $y_\lambda < y_2$.

[90] Definition: $F: X \to Y$ is multivalued quasiconvex with respect to S if whenever $y_1 \in F(x_1)$, $y_2 \in F(x_2)$ with $y_1 \leqslant y_2$ there is a $y_\lambda \in F(\lambda x_1 + (1-\lambda)x_2)$ with $y_\lambda \leqslant y_2$.

The following list of facts, collected as a theorem, follow from the definitions or from the same type of arguments as in the single valued cases.

[91] Theorem: (1) If $f: X \to Y$ is quasiconvex, (P) strictly quasiconvex, or convex with respect to S it is multivalued of the same type.

(2) F is multivalued convex with respect to S if and only if $\mathrm{Epi}_S F = \{(x,z) | \exists \ y \in F(x) \ z \geq y\}$ is convex.

(3) Convexity in the multivalued sense implies both multivalued quasiconvexity and (P) strict quasiconvexity. ∎

[92] Definition: x_o is a weak (local) minimum for F over A if there is some $y_o \in F(x_o)$ such that whenever $y \in F(x)$ and $y - y_o \in -S^o$ $x \notin A$ $(x \notin N \cap A)$.

[93] Proposition: If F is multivalued (P) strict quasiconvex with respect to S every local minimum is global.

Proof: Suppose x_o is a local minimum, then there is a neighbourhood N and $y_o \in F(x_o)$ such that when $x \in N(x_o)$ and $y \in F(x)$ $y - y_o \notin - S^o$.

Suppose that $y_1 - y_o \in - S^o$ and $y_1 \in F(x_1)$. Then for $o < \lambda < \lambda_o$ $\lambda x_1 + (1 - \lambda)x_o \in N(x_o)$. Since F is multivalued (P) strictly quasiconvex there is some $y_\lambda \in F(\lambda x_1 + (1 - \lambda)x_o)$ with $y_\lambda < y_o$ because $y_1 < y_o$. This contradicts the local minimality of y_o which asserts that no such y_λ can exist for $\lambda < \lambda_o$. ∎

[94] If $F: X \to R$ is multivalued (quasi)convex and for each x $F(x)$ is a nonempty compact set then

$$f(x) = \min \{ r | r \in F(x) \}$$

is a (quasi)convex retraction $f \subset F$. However, even on the line examples exist of multivalued convex and quasiconvex functions with no single valued retractions.

[95] Examples: (1) $F: R \to R$ defined by

$$F(r) = \left\{ \left[\tfrac{1}{2}r\right] + n \,\middle|\, n \in N \right\},$$

where $[r]$ is the greatest integer less than r, is a multivalued convex function which has no everywhere defined single valued convex restriction because the graph of F is not connected.

(2) Any multivalued maximal monotone mapping (see the final chapter) f mapping R into R is multivalued quasiconvex but won't necessarily contain a maximal monotone single valued restriction.

Chapter 2

TANGENT AND PSEUDOTANGENT CONES

Tangent cones and pseudotangent cones

The elementary observation in calculus that a function has derivative o at an extreme value is in many ways the keystone of all optimization results. The following standard proposition from Luenberger (1969) is the motivation for the developments in this section.

[1] Proposition: Let f be the real valued functional defined on a vector space X. Suppose that x_o minimizes f on a convex set $C \subset X$ and that f is Gateaux differentiable at x_o. Then

$$f'(x_o)(x - x_o) \geqslant o \qquad \forall\, x \in C. \,|$$

Essentially it was in order to generalize this result to nonconvex sets, with the corresponding implications for more general minimization problems, that the notions of tangent cones were introduced by Varaiya (1967), Guignard (1969) and others.

[2] The set T(E,x) consisting of all limits of the form $h = \lim \lambda_n(x_n - x)$ with $x_n \in E \subset X$, $\lambda_n \geqslant o$ and $x_n \rightarrow x$ in the topology on X is called the tangent cone to E at x. It is largely irrelevant whether the convergence is defined in terms of nets or sequences. For the purposes of simplicity all spaces will be assumed to be Hausdorff and locally convex from now on. Convergence in the given topology will be denoted by $\rightarrow$ while convergence in the weak topology, denoted $\sigma(X,X')$, will be denoted by $\rightharpoonup$ or by "wlim".

[3] The set wT(E,x) consisting of all h which are limits in the weak

topology of nets of the form $\lambda_n(x_n - x)$ with $\lambda_n \geq 0, x_n \in E$ and $x_n \to x$ will be called the <u>weak tangent cone</u> to E at x. This definition, which was announced by Nashed (1971), but which does not seem to have appeared in published articles, is extremely useful in optimization.

[4] T(E,x) and wT(E,x) are, respectively, closed and weakly closed cones but need not be convex. This motivates the next definitions.

The closures of the convex hulls of T(E,x) and wT(E,x) are called the <u>pseudotangent</u> and weak <u>pseudotangent cones</u>, respectively, and are denoted by P(E,x) and wP(E,x).

[5] A set $E \subset X$ will be said to <u>pseudoconvex at x_o with respect to a set</u> $F \subset X$ if $E - x_o \subset P(F,x_o)$. Weak <u>pseudoconvexity</u> is defined analogously. Guignard (1969) defined a set to be pseudoconvex at x_o if $E - x_o \subset P(E,x_o)$ which coincides with this definition in the case E = F. When this is so E will merely be called <u>pseudoconvex at x_o</u>.

This chapter is devoted to a development of the properties of tangent cones which are useful for framing optimization conditions but includes some results whose interest is intrinsic.

The next theorem lists some properties of tangent cones given by Guignard (1969).

[6] <u>Theorem</u>: If I is any index set then

(1) $T(\bigcap_{i \in I} A_i, x) \subset \bigcap_{i \in I} T(A_i, x)$

(2) $P(\bigcap_{i\in I} A_i,x) \subset \bigcap_{i\in I} P(A_i,x)$

(3) $\bigcup_{i\in I} T(A_i,x) \subset T(\bigcup_{i\in I} A_i,x)$

(4) $\bigcup_{i\in I} P(A_i,x) \subset P(\bigcup_{i\in I} A_i,x)$

Proof: These all follow from the definitions and the properties of closed convex cones.▌

The result holds also for weak cones. For the rest of the chapter results which hold by the same argument for both tangent cones and weak tangent cones will be marked simply by ' (W) also ' at the conclusion.

[7] A set A is said to be starshaped at $x \in A$ if whenever $y \in A$ and $0 \leq \lambda \leq 1$ $\lambda x + (1-\lambda)y \in A$. Clearly any convex set is starshaped at all its members.

[8] Proposition: (1) $A \subset B \Rightarrow T(A,x_o) \subset T(B,x_o)$; if $x_o \in \bar{A}$, $o \in T(A,x_o)$.

(2) $T(A,x_o) \subset wT(A,x_o)$; $P(A,x_o) \subset wP(A,x_o)$.

(3) The union of sets each pseudoconvex at x_o is pseudoconvex at x_o. (Guignard (1969))

(4) If $x_o \in A^o$, $T(A,x_o) = X$.

(5) If A is starshaped at x_o, A is pseudoconvex at x_o.

Proof: (1), (2), (3) follow directly from the definitions.

(4) Let $x \in X$. Since $x_o \in A^o$ there is an $n_o \in N$ such that for $n \geq n_o$ $x_n = x/n + x_o \in A$. Setting $\lambda_n = n$, $\lambda_n(x_n - x_o) = x$ and since $x \to x$ $x \in T(A,x_o)$.

(5) Let $x \in A$ then $x_n = \frac{1}{n}x + (1-\frac{1}{n})x_o \in A$ since A is starshaped at x_o. Setting $\lambda_n = n$ again

$$\lim n(x_n - x_o) = x - x_o \in T(A,x_o) \text{ and } A - x_o \subset T(A,x_o).$$

Thus A is pseudoconvex at x_o.▌

(W) also. ▮

[9] Proposition: If X is a metrizable convex space $T(A,x_o) = T(\bar{A},x_o)$.

Proof: If suffices to show that $T(\bar{A},x_o) \subset T(A,x_o)$. Hence let $\lambda_n(x_n - x_o) \to h$ with $x_n \in \bar{A}$, $\lambda_n \geqslant o$. If $h = o$ then $h \in T(A,x_o)$. If $h \neq o$ then $\{\lambda_n\}$ is unbounded and without loss of generality can be taken to be positive with limit ∞. Since $x_n \in \bar{A}$ there is a point $y_n \in A$ with $y_n - x_n \in \lambda_n^{-1} N_n$ (where N_n is a nested countable base for the topology). Then

$$\lambda_n(y_n - x_o) = \lambda_n(x_n - x_o) + \lambda_n(y_n - x_n)$$

and $\lambda_n(y_n - x_o) - \lambda_n(x_n - x_o) \subset N_n$. Thus

$$\lim \lambda_n(y_n - x_o) = h. ▮$$

(W) also. ▮

This proposition allows one for the most part to examine only closed sets.

[10] Proposition: If there is some set E with $A \subset E \subset B$ and E pseudoconvex at x_o then A is pseudoconvex at x_o with respect to B.

Proof: $A - x_o \subset E - x_o \subset P(E,x_o) \subset P(B,x_o)$. ▮

(W) also. ▮

In particular this holds if there is some convex set C with $A \subset C \subset B$.

[11] An example of a set with a trivial tangent cone and a nontrivial weak tangent cone is given below.

Example: Let $X = l_2$ and let $A \subset X$ be the set comprising of the o element in l_2 and $\left\{ \{\bar{x}_n\} \mid x_{nk} = \frac{1}{n} \text{ if } k = 1 \text{ or } n \text{ and } x_{nk} = o \text{ otherwise} \right\}$.

(1) $T(A,o) = o$. Clearly $\bar{x}_n \to o$. Suppose $\lambda_n \bar{x}_n \to a \neq o$.

Then $\|\lambda_n \bar{x}_n\| \geqslant \epsilon > o$ if $n \geqslant n_o$.

$\| \lambda_n \bar{x}_n \| = \frac{1}{n}\sqrt{2}\,\lambda_n$ so $\frac{\lambda_n}{n}$ does not tend to o.

Now

$$\| \lambda_n \bar{x}_n - \lambda_m \bar{x}_m \| = \left[\left(\frac{\lambda n}{n} - \frac{\lambda m}{m}\right)^2 + \left(\frac{\lambda n}{n}\right)^2 + \left(\frac{\lambda m}{m}\right)^2\right]^{\frac{1}{2}} \geqslant \frac{\lambda_n}{n}$$

which means that $\lambda_n \bar{x}_n$ does not converge and $T(A,o) = o$.

(2) $\underline{(1,o,o \ldots) \in w\,T(A,o)}$ since if $\lambda_n = n$

$\lambda_n \bar{x}_n = (1,o,o \ldots, o,1,o \ldots) \rightarrow (1,o,o, \ldots)$. ▌

[12] When S is a convex set alternative characterizations of $T(S,x_o)$ are possible.

<u>Proposition</u>: (<u>Varaiya</u> (1967)) When S is convex

$$\overline{\bigcup_{\lambda > o} \lambda(S - x_o)} = T(S,x_o) = P(S,x_o) \qquad x_o \in S.$$

<u>Proof</u>: If $h \in \overline{\bigcup_{\lambda > o} \lambda(S - x_o)}$ $h = \lim \lambda_n(x_n - x_o)$, $\lambda_n \geqslant o$, $x_n \in S$. The proof of [8] (5) shows that $S - x_o \subset T(S,x_o)$ so that $h_n = \lambda_n(x_n - x_o) \in T(S,x_o)$. Since $T(S,x_o)$ is a closed cone $h \in T(S,x_o)$. The converse containment is immediate. It is also clear that when S is convex $T(S,x_o)$ is convex and thus equals $P(S,x_o)$. ▌

(W) also. ▌

[13] <u>Proposition</u>: If S is convex $T(S,x_o) = w\,T(S,x_o)$.

<u>Proof</u>: By [12] $T(S,x_o) = \overline{\bigcup \lambda(S-x_o)}$ and

$w\,T(S,x_o) = \overline{\bigcup \lambda(S - x_o)}$ (the closure in the weak topology).

However, $\bigcup \lambda(S - x_o)$ is a convex set and has the same weak and initial closures (Taylor (1958)). ▌

[14] The above proposition allows another characterization of pseudoconvexity. $\underline{[A]}$ will be used to denote the convex hull of A.

Proposition: A is pseudoconvex at x_o if and only if

$$T([A],x_o) = P(A,x_o).$$

Proof: $\Rightarrow$ If A is pseudoconvex at x_o one has

$$A - x_o \subset P(A,x_o)$$

and since $P(A,x_o)$ is a closed convex cone

$$T([A],x_o) = \overline{\bigcup_{\lambda > o} \lambda([A] - x_o)} \subset P(A,x_o).$$

Since $A \subset [A]$ the other containment follows from

$$P(A,x_o) \subset P([A],x_o) = T([A],x_o) \text{ by } [12].$$

$\Leftarrow$ If $T([A],x_o) = P(A,x_o)$

$A - x_o \subset [A] - x_o \subset T([A],x_o) \subset P(A,x_o)$ using [8] (5) and [12].|

[15] The previous results also give the bound

$$P(A,x_o) \subset w\, P(A,x_o) \subset P([A],x_o)$$

and

Proposition: If A is pseudoconvex at x_o then $wP(A,x_o) = P(A,x_o)$.|

[16] The next theorem lists various translation properties of cones.

Theorem: (1) $T(A,x_o) = T(A - x_o, o)$.

Let $g(x) = f(x + x_o) - f(x)$ then

(2) $T(f(A),f(x_o)) = T(g(A - x_o), g(o))$

(3) $T(g^{-1}(B),o) = T(f^{-1}(B + f(x_o)), x_o)$.

Proof: These all follow by calculation from the definitions.|

(W) also.|

[17] Proposition: $\overline{P(A,x_o) + P(B,y_o)} \subset P(A + B, x_o + y_o)$ with equality if A and B are pseudoconvex at $x_o \in A$ and $y_o \in B$ respectively.

Proof: $P(A,x_o) = P(A + y_o, x_o + y_o) \subset P(A + B, x_o + y_o)$ if $y_o \in B$.

$P(B,y_o) = P(B + x_o, x_o + y_o) \subset P(A + B, x_o + y_o)$ if $x_o \in A$.

Since $P(A + B, x_o + y_o)$ is a closed convex cone

$$\overline{P(A,x_o) + P(B,y_o)} \subset P(A + B, x_o + y_o).$$

If A is pseudoconvex at x_o and B is at y_o then

$$[A] - x_o \subset P(A,x_o) \text{ and } [B] - y_o \subset P(B,y_o).$$

Thus $[A] + [B] - (x_o + y_o) \subset P(A,x_o) + P(B,y_o)$.

Since $[A] + [B]$ is a convex set one derives

$$P(A + B, x_o + y_o) \subset T([A] + [B], x_o + y_o) \subset \overline{P(A,x_o) + P(B,y_o)}.\blacksquare$$

(W) also.∎

A more interesting and much more useful result is proved next. It gives conditions under which [6] (1) can be replaced by equality.

[18] Theorem: Suppose A and B satisfy the following conditions.

(1) $[A] \cap [B] = [A \cap B]$.

(2) $[A]^o \cap [B]^o \neq \phi$. (or in fact $ri[A] \cap ri[B] \neq \phi$).

(3) $A \cap B$ is pseudoconvex at x_o.

Then $P(A,x_o) \cap P(B,x_o) = P(A \cap B, x_o)$.

Proof: By [6] (1) it suffices to show $P(A \cap B, x_o) \supset P(A,x_o) \cap P(B,x_o)$.

Since $A \subset [A]$, $B \subset [B]$

$$P(A,x_o) \cap P(B,x_o) \subset P([A],x_o) \cap P([B],x_o).$$

By a theorem of Rockafellar's (1970a) which is proved in R^n but holds in any convex space

$$P([A],x_o) \cap P([B],x_o) = P([A] \cap [B],x_o) \text{ when (2) holds.}$$

By (1), therefore,

$$P(A,x_o) \cap P(B,x_o) \subset P([A] \cap [B],x_o) = P([A \cap B],x_o)$$

and since (3) holds [14] shows that

$$P(A,x_o) \cap P(B,x_o) \subset P([A \cap B],x_o) = P(A \cap B,x_o).\blacksquare$$

(W) also.∎

[19] A set K is said to be <u>polyhedrally convex</u> if

$$K = \{x \in X \mid x_i^+(x) \geqslant a_i \quad i = 1,\dots,n \quad x_i^+ \in X'\}.$$

The next result lists various properties of polyhedral sets in R^n. Many of these are valid more generally.

<u>Theorem</u>: (1) If K is polyhedrally convex and M is a subspace of R^n then

$$P(K,x_o) \cap M = P(K \cap M,x_o).$$

(2) If K is polyhedrally convex $P(K,x_o)$ is polyhedral.

(3) The sums and duals of polyhedral sets are polyhedral.

(4) If B is polyhedrally convex and A is convex and $ri(A) \cap B \neq o$ then $P(A,x_o) \cap P(B,x_o) = P(A \cap B,x_o)$.

(5) If B and A are polyhedrally convex then

$$P(A,x_o) \cap P(B,x_o) = P(A \cap B,x_o).$$

<u>Proof</u>: These results, in R^n, all follow from the definitions with the exception of (4) which is proved in Rockafellar (1970a).▌

[20] This paragraph gives various examples of tangent cones which in particular show reasons why the conditions in [17] and [18] are imposed.

(1) <u>X = R</u>; $A = \left\{\begin{matrix}+1\\ -n\end{matrix}\right\}_1^{\infty} \cup \{o\}$, $x_o = \frac{1}{2}$

$B = (-\infty,\frac{1}{2}]$ $\qquad y_o = -\frac{1}{2}$.

$P(A,x_o) = o$ since $\frac{1}{2}$ is an isolated point.

$P(B,x_o) = (-\infty,o] = P(A,x_o) + P(B,x_o)$ while

$P(A + B,x_o + y_o) \supset P(A + o,x_o + y_o) = P(A,o) = R.$▌

A is an example of a closed set which is pseudoconvex at o but is not starshaped there.

(2) <u>X = R</u>; A is the rationals in $[o,1]$; B is the irrationals in $[o,1]$ plus $\{o,1\}$, $x_o = o$; then A, B fail only to satisfy [17] (3) and $P(A,o) = P(B,o) = [o,\infty)$; $P(A \cap B,o) = o.$▌

(3) $\underline{X = R^2}$; $A = \{(x,y) \mid x^2 \leqslant y\}$, $B = \{(x,y) \mid x^2 \geqslant -y\}$, $(x_o,y_o) = (o,o)$.

$$P(A,(x_o,y_o) = \{(x,y) \mid y \geqslant o\} \; ; \; P(B,(x_o,y_o) = \{(x,y) \mid y \leqslant o\}$$

but $P(A \cap B,(x_o,y_o) = (o,o)$. ▮

In this case only [18](2) is violated.

(4) $\underline{X = R^2}$; $A = \{(x,y) \mid x \geqslant o, y \geqslant o\} \cup \{(x,y) \mid x = o \text{ or } y = o\}$

$$B = \{(x,y) \mid x \geqslant o,\ y \geqslant o\} \cup \{(x,y) \mid x = y \text{ or } x = -y\}$$

and $(x_o,y_o) = (o,o)$.

$$P(A,(x_o,y_o)) = P(B,(x_o,y_o)) = R^2$$

while $P(A \cap B,\ (x_o,y_o)) = \{(x,y) \mid x \geqslant o,\ y \geqslant o\}$. ▮

In this case [18](1) is violated.

(5) $\underline{X = R}$; $A = C$, the Cantor set in $[o,1]$, $x_o \in C$.

Then $P(C,x_o)$ is either R^+, R^- or R dependent on whether the ternary expression of x_o contains finitely many twos or finitely many zeros or infinitely many of both. ▮

[21] As a partial converse to [18] one has:

Proposition: If $P(A \cap B,x_o) = P(A,x_o) \cap P(B,x_o)$ and both A and B are pseudoconvex at x_o so is $A \cap B$.

Proof: $(B \cap A) - x_o = (B - x_o) \cap (A - x_o)$

$\subset P(B,x_o) \cap P(A,x_o)$ (by pseudoconvexity)

$\subset P(A \cap B,x_o)$ (by hypothesis)

and $B \cap A$ is pseudoconvex at x_o. ▮

[22] An apparently open question is whether the condition that a closed set be pseudoconvex at all its points is equivalent to convexity. A partial answer to this is provided by the next propositions.

Proposition: Suppose $f\colon X \to R$ is Fréchet differentiable then Epi$f$ is pseudoconvex at all its points exactly when f is convex.

Proof: => The proof relies on the geometrically evident assertion that $P(\text{Epi}f,(x_o,f(x_o))) = \{(x,y) \mid y \geqslant f'(x_o)(x\ )\}$ which is straightforward but slightly tedious to prove and is thus omitted. Suppose that Epif is pseudoconvex at $(x_o,f(x_o))$ for all x_o. Then

$$\text{Epi}f \subset (x_o,f(x_o) + P(\text{Epi}f,(x_o,f(x_o)))$$

or $\quad (x - x_o, f(x) - f(x_o)) \in P(\text{Epi}f,(x_o,f(x_o)) \quad \forall x \in X$

which, using the assertion above, shows that

$$f(x) - f(x_o) \geqslant f'(x_o)(x - x_o) \quad \forall x, x_o \in X.$$

This is equivalent, using [1.69], to convexity of f.

<= This follows immediately from [8] and [1.46].|

[23] The next result links tangent cones directly with the results of chapter 1 for the first time.

Theorem: (1) If f is (P) strictly quasiconvex and fully upper semicontinuous with respect to S then any weak minimum x_o for f over A is a weak minimum over $x_o + T(A,x_o)$.

(2) If f is absolutely quasiconvex and upper semicontinuous with respect to S then any strong minimum x_o for f over A is a strong minimum over $x_o + T(A,x_o)$.

Proof: Let $y_o \in x_o + T(A,x_o)$. Then $y_o = \lim y_n + x_o$

$y_n = \lambda_n(x_n - x_o)$, $\lambda_n \geqslant 0$ $\quad x_n \in A$ and $x_n \to x_o$.

(1) Suppose $f(y_o) - f(x_o) \in -S^o$.

By [1.25] $\quad f(x_o + y_n) < f(x_o)$ for n sufficiently large.

Also $x_n = (1 - \lambda_n^{-1})\, x_o + \lambda_n^{-1}(x_o + \lambda_n(x_n - x_o))$ so that

$$f(x_n) = f((1 - \lambda_n^{-1})\, x_o + \lambda_n^{-1}(x_o + y_n)) < f(x_o)$$

since f is assumed (P) strictly quasiconvex and since $f(x_o + y_n) < f(x_o)$.

(Note that one may assume $0 < \lambda_n^{-1} < 1$ for n sufficiently large since $x_o \neq y_o$, $f(x_n) < f(x_o)$ contradicts the assumed minimality of x_o over A so that no such y_o can exist and the result is true.

(2) The second case follows similarily but under the weaker continuity assumption of [1.24].|

Corollary: If in (2) the hypothesis that f is quasiconcave with respect to S is added, $x_o + T([A], x_o)$ can be used to replace $x_o + T(A, x_o)$.

Proof: The quasiconcave version of [1.85] allows A to be extended to [A].|

These results extend, from continuous real valued convex mappings, a theorem of Norris (1971).

Derivatives and tangent cones

The notions previously discussed are now related to differential properties since it is in this form that they are of most use in optimization.

[24] Theorem: Let $f: X \to Y$ be compactly differentiable and suppose x_o is a strong minimum for f over A with respect to a closed convex cone S then

$$f'(x_o)(h) \in S \qquad \forall h \in P(A, x_o).$$

Proof: Zlobec (1973) proves the result when $(Y,S) = (R,R^+)$. There is no difficulty in using the properties of cones and the definition of a strong minimum ([1.73]) to extend it.|

[25] For a weak minimum the following theorem holds.

Theorem: Let $f: X \to Y$ be compactly differentiable at x_o and suppose x_o is a weak minimum for f over A with respect to a closed convex cone with interior, then there is some $u^+ \in S^+ / \{o\}$ with

$$u^+(f'(x_o)) \in P^+(A,x_o).$$

<u>Proof</u>: Let $E = \{ y \mid \exists\, h \in T(A,x_o) \text{ such that } (f'(x_o))(h) \leqslant y \}$

Suppose $h \in T(A,x_o)$ with $(f'(x_o))(h) < o$. Now

$$h = \lim \lambda_n(x_n - x_o) \quad x_n \in A, x_n \to x_o, \ \lambda_n \geqslant o.$$

Let $h_n = \lambda_n(x_n - x_o)$

$$f'(x_o)(h) = \lim_n \lambda_n\left[f(x_o + \lambda_n^{-1}h_n) - f(x_o)\right]$$

since f is compactly differentiable. Hence for $n \geqslant n_o$

$$\lambda_n\left[f(x_o + \lambda_n^{-1}h_n) - f(x_o)\right] < o$$

and since $-S^o$ is a cone

$$f(x_n) = f(x_o + \lambda_n^{-1}h_n) < f(x_o).$$

This contradicts the minimality of x_o as $x_n \in A$. Thus $E \cap -S^o = \phi$

E is clearly a convex set and the Hahn-Banach theorem implies the existence of a non-zero linear functional with

$$u^+(y) \geqslant o \quad \forall y \in E; \quad u^+(s) \leqslant o \quad \forall s \in -S.$$

So

$$o \neq u^+ \in S^+ \quad \text{and} \quad u^+(f'(x_o))(h) \geqslant o \quad \forall h \in T(A,x_o).$$

By continuity and linearity of $f'(x_o)$ $\quad u^+(f'(x_o)) \in P^+(A,x_o)$. |

While, as Zlobec, remarks the compact derivative is the natural derivative to use when dealing with tangent cones it is the bounded derivative which plays the comparable role for weak tangent cones.

[26] <u>Theorem</u>: If f is taken to be boundedly differentiable the results of [24] and [25] remain true with $wP(A,x_o)$ replacing $P(A,x_o)$. In [25] it is necessary for $f'(x_o)$ to be completely continuous. (See the remarks on complete continuity at the end of the chapter). X is assumed to be sequential.

Proof: To mirror the previous proofs it is only necessary to observe that when $h_n = \lambda_n(x_n - x_o), x_n \to x_o, \lambda_n \geqslant o, x_n \in A$ and $h_n \to h_o$ one can assume that $\{h_n\}$ is weakly bounded since $\mathfrak{X}$ is sequential. Since the same sets are bounded in any topology of the dual pair (Robertson and Robertson (1964)) it is bounded. The definition of the bounded derivative enables one to assert that

$$\lambda_n[f(x_o + \lambda_n^{-1}h_n) - f(x_o)] \to f'(x_o)(h)$$

and the proofs proceed as before. |

To see that [26] is a genuine sharpening of [24] it is only necessary to consider the example in [11]. With A and x_o as in [11] and $f: l_2 \to R$, Zlobec's theorem says that $f'(x_o) \in P^+(A,x_o) = l_2$ while [26] requires $f'(x_o)$ to have first coordinate non negative. Note that since l_2 is a Banach space the bounded derivative is in fact Fréchet.

[27] Guignard's (1969) sufficiency condition also has extensions to weak cones and (Y,S).

Theorem: Let $f: X \to Y$ be β - differentiable at x_o. Suppose that (1) $f'(x_o)(h) \in S \quad \forall h \in P(A,x_o)$, (2) A is pseudoconvex at x_o, (3) f is pseudoconvex at x_o with respect to S; then x_o is a strong minimum for f over A.

Proof: Since A is pseudoconvex at x_o, $x - x_o \in P(A,x_o) \quad \forall x \in A$.

Thus $\quad f'(x_o)(x - x_o) \in S \quad \forall x \in A,$

and since f is supposed pseudoconvex at x_o

$$f(x) \geqslant_s f(x_o) \quad \forall x \in A. \,|$$

[28] Theorem: Suppose in [27] that (1) becomes (1)′ $f'(x_o)(h) \in S$ $\forall h \in w\,P(A,x_o)$ and (2) becomes (2)′ A is weakly pseudoconvex at

x_o then the result still holds.

Proof: The proof is independent of the nature of the cones.|

[29] A sufficiency condition of a sort can be proven for weak minima. The result guarantees a real valued equivalent problem.

Theorem: Suppose that for some $u^+ \in S^+/\{o\}$

$$u^+(f'(x_o)(h)) \geqslant o \quad \forall h \in P(A, x_o)$$

and that u^+f is pseudoconvex at x_o and A is pseudoconvex at x_o then

$$u^+f(x) \geqslant u^+f(x_o) \quad \forall x \in A.$$

Proof: This is contained in [27].|

The analogous result for weak cones is contained in [28].|

The pseudoconvexity of u^+f is not implied by the pseudoconvexity of f with respect to S it is however implied by the convexity of f with respect to S ([1.42], [1.43]).

[30] That [29] actually adds to ones knowledge can be seen from the following example of a function which is quasiconvex with respect to the orthant in R^2 but has no equivalent real valued map.

Example: Let $f: R \to R^2$, with S the coordinate cone, be given by $f(x) = (x^3 + 1, -x)$. f is quasiconvex and every x is a weak minimum for f over S since

$$f(x) \leqslant f(y) \quad \Rightarrow \quad x = y.$$

Let $u^+ = (r_1, r_2) \in S$, that is $r_1 \geqslant o, r_2 \geqslant o$ $(r_1, r_2) \neq o$.

Then

$$u^+ f(x) = r\,x^3 - r_2 x + r_1$$

which is unbounded on R. For any x_o, therefore, there is some x with $u^+f(x) < u^+f(x_o)$ and f has no real valued equivalent map.

[31] In keeping with the definition of Zlobec (1973), the set

(i) $\underline{\mathcal{E}} = \{E \mid E \in B(X,Y) : E(P(g^{-1}(B),x_o)) \subset P(B,g(x_o))\}$

is called the <u>local cone</u> (<u>of derivatives</u>) where $g: X \to Y$.

(ii) $\underline{w\,\mathcal{E}} = \{E \mid E \in B(X,Y) : E(wP(g^{-1}(B),x_o)) \subset wP(B,g(x_o))\}$

will correspondingly be called the <u>weak local cone</u>.

Zlobec introduced the terminology of [31](i) to allow the formulation of optimality conditions when g is not differentiable. Massam (1973) has shown that $\mathcal{E}$ is a closed convex cone in the topology of convergence on bounded sets. It is clear that this holds for $w\,\mathcal{E}$. More information on cone containments can be found in Ritter (1969a, b).

[32] <u>Proposition</u>: [Zlobec (1973)] If g is compactly differentiable at x_o then $g'(x_o) \in \mathcal{E}$.

<u>Proof</u>: The proof is similar to that of the next proposition. |

[33] <u>Proposition</u>: Let $g: X \to Y$ be boundedly differentiable at x_o and let X be sequential: then $g'(x_o) \in w\,\mathcal{E}$.

<u>Proof</u>: Set $\epsilon > o$. Let $h \in w\,T(g^{-1}(B),x_o)$. Then

$h_n = \lambda_n(x_n - x_o), g(x_n) \in B, x_n \to x_o, \lambda_n \geqslant o$ and $h_n \rightharpoonup h_o$.

Let $u_1^+, \ldots u_k^+ \in X'$ and let N be any neighbourhood in X with $u_i^+(N) < \epsilon$ for $i = 1,\ldots,k$. Because $h_n \rightharpoonup h_o$ in a sequential space one can assume that $\{h_n\}$ is in fact a sequence (relabelling if necessary). Then $\{h_n\}$ is a bounded set because it weakly bounded. By the definition of the bounded derivative

(1) $\lambda_n\left[g(x_o + \lambda_n^{-1}h_n) - g(x_o)\right] - g'(x_o)(h_n) \to o$ as in [26].

Note that λ_n^{-1} can be assumed convergent to zero by choosing a subsequence if necessary since otherwise $h = o$. From (1) one has for $n \geqslant n_1$

(2) $\lambda_n[g(x_n) - g(x_o)] - g'(x_o)(h_n) \in N$.

By the definition of differentiation g' is a continuous map and is, therefore, $\sigma(X,X') - \sigma(Y,Y')$ continuous (Robertson and Robertson (1964)). There is then some n_2 with

(3) $|u_i^+(g'(x_o))(h_n) - u_i^+(g'(x_o))(h)| < \epsilon \qquad i = 1,2,..k.$

for $n \geqslant n_2$. From (2),(3) and the choice of N

(4) $| u_i^+ [\lambda_n[g(x_n) - g(x_o)] - (g'(x_o))(h)] | < 2\epsilon$

when $n \geqslant \max(n_1,n_2,n_3)$ and (4) implies the weak convergence of $\lambda_n[g(x_n) - g(x_o)]$ to $(g'(x_o))(h)$.

Since g is continuous and $x_n \to x_o$, $g(x_n) \to g(x_o)$ and $g'(x_o)(h) \in wT(B,g(x_o))$.

The linearity and continuity of $g'(x_o)$ suffice to derive

$g'(x_o)(wP(g^{-1}(B),x_o)) \subset wP(B,g(x_o))$.|

If the continuity assumptions on g are strengthened the conclusion can be improved.

[34] <u>Definition</u>: $f: X \to Y$ is said to be <u>completely continuous</u> if it is continuous from the weak topology on X to the given topology on Y.

[35] <u>Proposition</u>. If $g'(x_o)$ is completely continuous then the conclusion of [33] can be strengthened to

$(g'(x_o))(wP(g^{-1}(B),x_o)) \subset P(B,g(x_o))$.

<u>Proof</u>: Examining the proof of [33] one sees that (3) can be replaced by (3)' $(g'(x_o))(h_n) - (g'(x_o))(h) \in N$ for $n \geqslant n_4$ where N is now <u>any</u> neighbourhood. One then combines (3)' and (2) and proceeds as in the proposition.|

Since $g'(x_o)$ is automatically continuous it suffices in addition to require that $g'(x_o)$ map bounded sets into compact sets. In normed spaces this last condition coincides with compactness of $g'(x_o)$. More generally compactness is stronger. Nashed (1971) has that a sufficient condition in a normed space for $g'(x_o)$ to be completely continuous is the complete continuity of g itself.

[36] For a mapping $E \in \mathcal{E}$ Zlobec (1973) has defined the set

$$K(E) = \{h \mid E(h) \in P(B, g(x_o))\}$$

For a mapping $E \in w\mathcal{E}$ $wK(E)$ is defined analogously by

$$wK(E) = \{h \mid E(h) \in wP(B, g(x_o))\}$$

The following theorems give some conditions for $P(g^{-1}(B), x_o)$ to equal $K = K(g'(x_o)))$. Clearly, by [32] and [33], this reduces to showing that

$$g'(x_o)^{-1}\left[P(B, g(x_o))\right] \subset P(g^{-1}(B), x_o).$$

The results are all framed in normed spaces since they all rely at some level on the implicit function theorem (Luenberger (1969)) or on similar results.

[37] Theorem: (Halkin (1972b)) Let X and Y be Banach spaces and $T \subset X$ a closed subspace. Suppose

(1) f is continuously differentiable in a neighbourhood U of x_o.

(2) $f'(x_o)$ is a bijection of T onto Y.

Then there is some neighbourhood N of x_o with a continuously differentiable mapping j of N into X with

(i) $f(x + j(x)) = f(x_o) + f'(x_o)(x - x_o) \quad \forall x \in N$

(ii) $\lim_{p \to 0} \sup_{\|x - x_o\| < p} \|j(x)\|/p = 0$. ∎

[38] Theorem: (Halkin (1972a)) Let X be normed and Y finite dimensional.

Suppose

(1) f is differentiable in a neighbourhood U of x_o.

(2) $f'(x_o)$ is surjective.

Then the conclusions of [37] hold except that j need not be continuously differentiable.|

These two results can be used to give sufficient conditions for $K = P(g^{-1}(B),x_o)$ when B is a closed subspace.

[39] Proposition: Suppose $g: X \to Y$ satisfies the hypotheses of [37] or [38] and that $B \subset Y$ is a closed subspace. Then

$$K = P(g^{-1}(B),x_o) = T(g^{-1}(B),x_o).$$

Proof: Suppose $g'(x_o)(h) \in P(B,g(x_o)) = B$. Set $x_o = o$.

For $n \geq n_o$ $h/n \in N$ and thus by (i)

$$g(x_o + h/n + j(h/n)) = \tfrac{1}{n}g'(x_o)(h) + g(x_o) \in B.$$

Thus

$$x_o + h/n + j(h/n) \in g^{-1}(B) \text{ and}$$

$$\lim_{n\to\infty} \frac{h/n - j(h/n)}{\frac{1}{n}} \in T(g^{-1}(B),x_o).$$

Using (ii) this limit is just h. The theorem now follows from the fact that K always contains $T(g^{-1}(B),x_o)$.|

Flett (1966) has proved [39] by a rather complicated implicit function theorem only using $g'(x_o)$ is onto Y. He does not need T to exist in [37].

[40] Definition: When $K = P(g^{-1}(B),x_o)$ g will be said to be regular at x_o. If $wP(g^{-1}(B),x_o) = wK$ g will be called weakly regular at x_o.

These definitions generalize the standard notion of regularity

(Luenberger (1969)). The subject will be discussed again later.

Remarks on complete continuity of derivatives ([26], [34])

[41] The complete continuity of $f'(x_o)$ is necessary in any argument such as a separation argument in which one wishes to deduce from $h \in wP(A,x_o)$ and

$$h_n = \lambda_n(x_n - x_o) \rightharpoonup h_o$$

that (1) $(f'(x_o))(h_o) \in -S^o$ implies $f(x_n) - f(x_o) \in -S^o$, $n \geqslant n_o$. This relies on the fact that

$$\lambda_n \left[f(x_n) - f(x_o)\right] - (f'(x_o))(h_n) \to 0$$

which in conjunction with $h_n \rightharpoonup h_o$ and $f'(x_o)$ completely continuous gives $\lambda_n \left[f(x_n) - f(x_o)\right] \to (f'(x_o))(h_o) \in -S^o$. Then $n \geqslant n_o$ implies, since S^o is open in the topology in which one has convergence, that the implication (1) holds.

[42] Alternatively one can define minimization in the weak topology by requiring that S has an interior S^{wo} in the weak topology, and by defining f to have a weak minimum in the weak topology at x_o over A if

$$f(x) - f(x_o) \notin -S^{wo} \quad \text{when } x \in A.$$

With this notion complete continuity is unnecessary since now $\lambda_n\left[f(x_n) - f(x_o)\right] \rightharpoonup f'(x_o)(h_o) \in -S^{wo}$ implies that $f(x_n) - f(x_o) \in -S^{wo}$ for $n \geqslant n_o$ as the cone now has interior in the topology in which $\lambda_n\left[f(x_n) - f(x_o)\right]$ converges.

[43] These considerations lead to the following general proposition.

Proposition: If S is supposed to have weak interior then any separation argument which holds for (strong) tangent cones and compact derivatives remains valid for weak tangent cones and bounded derivatives without complete continuity: provided that x_o is required to be a minimum in the weak topology.▌

It must be emphasised that most applications of weak tangent cones will rely on [26] and [33] which require X to be sequential. For this reason any bounded derivative will from now on be assumed to have a sequential domain space and the hypothesis will not be listed separately.

Chapter Three

FARKAS LEMMAS AND TRANSPOSITION THEOREMS

Farkas Lemmas and Transposition Theorems.

Two of the most useful optimization methods (other than direct arguments which usually involve the Hahn-Banach theorem) are the application of Farkas Lemmas and Transposition or Alternative Theorems. This chapter proves various results of this nature some of which will be applied subsequently. The following definitions are needed. They are taken from Berge (1959) which is a good general reference text for multivalued maps. Recall that $F : X \to Y$ is multivalued if F maps points in X onto subsets of Y.

[1] Definitions: (1) $F : X \to Y$ is said to be upper semicontinuous as a multivalued mapping between topological spaces if for any neighbourhood $V \subset Y$ with $F(x_o) \subset V$ there is a neighbourhood $U \subset X$ of x_o with $F(x) \subset V$ when $x \in U$. (2) F is called lower semicontinuous as a multivalued mapping if for any neighbourhood $V \subset Y$ with $F(x_o) \cap V \neq \emptyset$ there is a neighbourhood $U \subset X$ of x_o with $F(x) \cap V \neq \emptyset$ when $x \in U$.

It is apparent that these concepts coincide when f is single valued.

[2] Theorem: Let $f : X \to Y$ $g : X \to Z$ be arbitrary (single valued) mappings. Let $U \subset Y$, $V \subset Z$ be any sets such that $U \subset R(f)$. Then the following are equivalent:

(1) $f(x) \in U \Rightarrow g(x) \in V$

(2) $g(x) \subset h \circ f\ (x)$ $\quad \{h(U)\} \subset V$ where h is a multivalued mapping of Y into Z.

Proof: $\Rightarrow$ Let $h(u) = \{g(x) \mid f(x) = u\}$.

In multivalued terms $h = gf^{-1}$. Suppose $u \in U$ then, since $U \subset R(f)$, $u = f(x)$ for some x and $x \in f^{-1}(u)$.

$$h(u) = g(f^{-1}(u)) \subset V \text{ by (1)} .$$

It is clear that $g(x) \subset h \circ f\ (x)$ and that (2) $\Rightarrow$ (1). |

[3] Proposition: h is singlevalued if and only if $f(x) = f(y)$ implies $g(x) = g(y)$.

Proof : $h\ (f(x)) = \{g(y) \mid f(x) = f(y)\}$. |

[4] Theorem: (1) If g is continuous and f^{-1} is upper (lower) semicontinuous h is upper (lower) semicontinuous.

(2) If g is convex with respect to S and f is linear then h is multivalued convex with respect to S([1.88])

(3) If g and f are linear h satisfies $th(x) + h(y) \subset h(t\,x + y)$ for $t \in R$ $x, y \in Y$. (This property is called multivalued linearity).

Proof : (1) Follows by a result of Berge (1969) or directly. (2), (3) follow from the observation that the inverse of a linear map is multivalued linear and from $h = gf^{-1}$. |

It is a property of multivalued upper semicontinuous mappings F that $\{F(x)\}$ is compact. (Berge (1959)). It is apparent that the previous results could have been rephrased for f and g multivalued. In this case the continuity conditions in [4](1) would ask only for g to be semicontinuous.

Craven (1972) has proved that h is continuous when g is continuous and f is continuous, open and surjective. He was only considering maps satisfying [3]. The next proposition shows that even in this case [4] (1) generalises his result.

[5] Proposition: If $f : X \to Y$ is an open map then f^{-1} is lower semicontinuous as a multivalued map from Y into X.

Proof: Suppose there is a neighbourhood $V \subset X$ such that

$$f^{-1}(y_o) \cap V \neq \emptyset .$$

Let $U = f(V)$. U is open in Y since f is open. Also, since $f^{-1}(y_o) \cap V \neq \emptyset$, there is some $x_1 \in V$ with $f(x_1) = y_o$ and $y_o \in U$. If $y \in U$ then $y = f(x)$ for some $x \in V$ since $U = f(V)$. For this x one has

$$x \in f^{-1}(y) \cap V.$$

Thus for any V there is a neighbourhood U of y_o such that $f^{-1}(y) \cap V \neq \emptyset$ if $y \in U$. This is just definition [1] (2).∎

Combining [5] and [4] (1) one has that h is lower semicontinuous when f is open and g is continuous. This in turn implies that h is continuous when it is singlevalued and everywhere defined. This does not need Craven's hypothesis that f is continuous. As a corollary one has the following linear Farkas lemma.

[6] Theorem: Let X, Y, Z be convex spaces with X fully complete and Y separated and barrelled. Let $S \subset Y$ and $Q \subset Z$ be cones with Q pointed. Suppose $A : X \to Y$ and $B : X \to Y$ are continuous linear mappings with $R(A) = Y$. The following are equivalent:

(1) $Ax \in S \Rightarrow Bx \in Q$

(2) $B = T \cdot A$ for some continuous linear $T : Y \to Z$ with $T(S) \subset Q$.

Proof: $\Rightarrow$ $Ax = Ay \Rightarrow A(x-y) = o \in S \cap -S$. By (1) $B(y-x) \in Q \cap -Q = \{0\}$

Thus the condition of [3] is satisfied and by [2] there is a single valued map T with $B = T \cdot A$ and $T(S) \subset Q$. [4] (3) shows that T is linear. Finally T is continuous. This is true because A is open (the hypothesis of the open mapping theorem are satisfied) and one may apply the preceeding remark. $\Leftarrow$ This is immediate. ∎

With $R(A) \cap S$ replacing S the theorem can be proven for $R(A)$ a fully complete subspace since A will be open onto $R(A)$. It is also worth noting that if A is assumed open the theorem holds in general (convex) spaces.

[7] If R(A) is not assumed closed there appears to be no satisfactory continuity result in the literature. The best one might hope for is the equivalence of [6] (1) with (2)' $B = \lim_n T_n A$, $T_n(S) \subset Q$, T_n continuous.

Since the only results I have obtained in this direction are restricted and weaker, they are not included herein. Such results would be useful in optimization.

[8] The next result gives a convex extension to [6].

<u>Theorem</u>: Let $f : X \to Y$ and $g : X \to Z$ be differentiable at a point x_o with $R(f'(x_o)) = Y$. Suppose further that f is convex with respect to S and g is convex with respect to Q with Q pointed. Suppose that X is fully complete and Y is barrelled, if (1) $f(x) \geq_s f(x_o) \Rightarrow g(x) \leq_q g(x_o)$ there is some continuous T mapping Y into Z with $T(S) \subset Q$ and such that

$$(2)\ g(x) + T\, f(x) \geq_q g(x_o) + T\, f(x_o)\ .$$

Proof: Since f and g are differentiable and convex

(3) $f(x) - f(x_o) \geq_s f'(x_o)\,(x-x_o)$

(4) $g(x) - g(x_o) \geq_q g'(x_o)\,(x-x_o)$ by [1.69].

Using (1), (3) and (4) one sees that

$$f'(x_o)\,(x-x_o) \in S \text{ implies } g'(x_o)\,(x-x_o) \in -Q .$$

Since $R(f'(x_o)) = Y$ the Farkas Lemma, [6], can be applied to $f'(x_o)$ and $g'(x_o)$ giving $T : Y \to Z$, T continuous with $T(S) \subset Q$ and

(5) $T f'(x_o) + g'(x_o) = o$.

Using (3) and $T(S) \subset Q$ one sees that

(7) $T(f'(x_o)\,(x-x_o)) \leq_q T(f(x) - f(x_o))$.

Combining (7) with (4) and (5) one has

$$g(x) - g(x_o) \geq_q g'(x_o)\,(x-x_o) = -Tf'(x_o)\,(x-x_o) \geq_q -T\,(f(x) - f(x_o))$$ as desired.∎

If f and g are linear this reduces to [6] since continuous linear mappings are their own derivatives. Setting $x_o = o$

$$g(x) + T f(x) \in Q \cap -Q = \{o\}$$ as before.

[9] If one wishes to extend the result of [6] to

$$A_1x \in S_1,\ A_2x \in S_2,\ \ldots,\ A_nx \in S_n \Rightarrow Bx \in Q$$

it is necessary to impose further restrictions on the cones Ritter (1969 ,b) has proved the following result which is quoted for future use.

Theorem: Suppose $A_i : X \to Y_i$ $i = 1,\ldots,n$, $B : X \to Z$ are continuous linear mappings between Banach spaces and that Z is reflexive. Suppose also that

(1) $\exists\, \bar{x} \ni A_i \bar{x} \in S_i^o \quad i = 1,\ldots,n$.

(2) $Q \subset Z$ is a closed normal cone

then the following are equivalent:

(3) $A_1 x \in S_1, \ldots, A_n x \in S_n \Rightarrow Bx \in Q$.

(4) There are continuous linear mappings $T_i : X_i \to Z \quad i = 1,\ldots,n$ with $T_i\,(S_i) \subset Q$ and $B = T_1 A_1 + \ldots + T_n A_n$.▮

Craven (1972) seems to use this result without proving it or imposing conditions (1) and (2), in that he tries to deduce the result directly from the basic lemma by looking at $A = (A_1, \ldots, A_n)$.

<u>Transposition Theorems for convex and linear mappings</u>

The first result of this section generalises a Transposition theorem proved by Craven and Mond (1973) to multivalued convex functions.

[10] <u>Theorem</u>: Let X, Y, Z be convex spaces. Let $F : X \to Y$ be a multivalued convex function with respect to S, a closed convex cone with interior. Let $h : X \to Z$ be affine and open and suppose that F is multivalued lower semicontinuous on $C \subset X$, a convex set with interior. If there is no solution to

(1) $h(x) = o$, $x \in C$ and $F(x) \cap -S^o \neq \emptyset$

then

(2) there is some $p^+ \in S^+$, $q^+ \in Z'$ with $(p^+, q^+) \neq o$ and $p^+ \,(F(x)) + q^+ \,(h(x)) \geq o \;\; \forall\, x \in C$.

<u>Proof</u>: Let $A = \{(w,z) \mid \exists x \in C,\ h(x) = z \text{ and } \exists y \in F(x) \ni w >_S y\}$

Then (i) A is convex since C is convex, h is affine and F is multivalued convex.

(ii) Let $s \in S^o$. There is some balanced neighbourhood N_o with $s + N_o \subset S^o$. Let $x_o \in C^o$ and $z_o = h(x_o)$ then there is some $y_o \in S^o$ with $a = y_o - y_1 \in S^o$ where y_1 is any given point in $F(x_o)$. (This could be done for any y_1).

Choose a balanced open set N_1 with $a + 2N_1 \subset S^o$ and another open set $N_2 \subset X$ with $x_o + N_2 \subset C$. Since F is lower semicontinuous and $F(x_o) \cap (N_1 + y_1) \neq \emptyset$ there is some open $N \subset N_2$ with

$$F(x) \cap (N_1 + y_1) \neq \emptyset$$

when $x \in N$.

Choose $\bar{y} \in F(x)$ and $\bar{y} \in N_1 + y_1$ for any $x \in N + x_o$.

Let $y \in N_1 + y_o$. Then

$$y - \bar{y} = (y - y_o) + (y_1 - \bar{y}) + a \in 2N_1 + a \subset S^o.$$

Moreover, since h is affine and open, $h(N + x_o)$ is open and if $(y, z) \in (N_1 + y_o,\ h(N + x_o))$ one has

$$z = h(x),\ x \in C \text{ and } y > \bar{y} \in F(x).$$

Thus $(y, z) \in A$ which implies that $(y_o, h(x_o)) \in A^o$.

(iii) Let $M = \{(w, o) \mid w \in -S\}$

M is clearly convex. If there is no solution to (1) then $M \cap A = \emptyset$.

An application of the Hahn Banach theorem allows one to assert the existence of a linear functional $a^+ = (p^+, q^+) \neq o$ with

$$p^+(w) + q^+(z) \geq o \quad \text{if } (w,z) \in A.$$

If one lets $y_n = y + \frac{1}{n} s$ for a fixed $s \in S^o$ and any $y \in F(x)$, $x \in C$; one has $(y_n, x) \in A$ and

(iv) $p^+(y_n) + q^+(h(x)) \geq o \quad \forall x \in C.$

Taking limits as $n \to \infty$

$$p^+(y) + q^+(h(x)) \geq o \quad \forall x \in C, \; y \in F(x).$$

Further $p^+ \in S^+$, since if $y \in S^o$ $(ty, h(x)) \in A$ when t is sufficiently large (as in (ii)) and if $p^+(y) < o$ (iv) could not hold.|

By [1.53] the continuity assumptions are fulfilled when S is a normal cone and F is single valued and continuous at some point of $C \subset (\text{dom} f)^o$. A further generalisation is given below.

[11] Theorem: Let B be a closed convex cone in Z and suppose the hypotheses of [10] hold. Then:

(1) There is a solution to $F(x) \cap -S^o \neq o$, $h(x) \in -B$, $x \in C$.

or (2) There are $p^+ \in S^+$ $q^+ \in Q^+$, not both zero, with

$$p^+(F(x)) + q^+(h(x)) \geq o \quad \forall x \in C.$$

Proof: Let $\hat{A} = \{(w,z) \mid x \in C_1 \; z - h(x) \in B, \; w - y \in S^o$ for some $y \in F(x)\}$. Since A of [10] is a subset of $\hat{A}$, $\hat{A}$ has interior. $\hat{A}$ is convex in much the same way as A was.

Let $\hat{M} = \{(w,z) \mid \; w \in -S, \; z \in -B\}$.

Suppose (1) has no solution then $\hat{A} \cap \hat{M} = \emptyset$. Again, using the Hahn Banach theorem, there are p^+, q^+ not both zero with

(3) $p^+(w) + q^+(z) \geq p^+(s) + q^+(b)$

when $(w,z) \in A$ and $s \in -S$ $b \in -B$. In particular, as before, $p^+(y) + q^+(h(x)) \geq o$ if $x \in C$ and $y \in F(x)$.

$p^+ \in S^+$ and $q^+ \in B^+$ or (3) is impossible.|

This includes [10] as the case B = o. The next two results are corollaries to [11].

[12] Corollary: Let X, Y, Z, F, h, S and B be as in [11]. The following strict alternative holds.

(1) There is some $p^+ \neq o \in S^+$, $q^+ \in B^+$ with

$$p^+(F(x)) + q^+(h(x)) \geq o \quad \forall x \in X$$

or (2) There is a solutimn x_o to

$$h(x) \in -B \quad F(x) \cap -S^o \neq \emptyset .$$

Proof: $\Leftarrow$ If (2) has no solution there is by [11] a solution to (1). It remains to verify that $p^+ \neq o$ If $p^+ = o$ then

$$q^+(h(x) \geq o \quad \forall x \in X.$$

Since h is open $R(h) = Y$ and $q^+ = o$. This forces p^+ to be non zero which is a contradiction.

$\Rightarrow$ Conversely if x_o solves (2) and $p^+ \in S^+ \backslash \{o\}$, $q^+ \in B^+$ there is some $y_o \in F(x_o) \cap -S^o$ and $h(x_o) \in -B$. Then

$$p^+(y_o) + q^+ h(x_o)) \leq p^+(y_o) < o$$

and there is no solution to (1). ∎

[13] Corollary: Suppose C is a closed convex cone with interior and that $f = Ax + d$ is continuous and affine with $d \in -S$. Suppose that $B = o$, h is linear and there is x_1 with $f(x_1) \in -S^o$. Then, when [11] (1) has no solution, p^+ and q^+ of [11] (2) are such that $p^+(d) = o$, and $p^+ \neq o$ implies $q^+ \neq o$.

Proof: Setting $x_o = o \in C$ and using

$$p^+(A.x + d) + q^+(h(x)) \geq o \quad x \in C$$

one has $p^+(d) \geq o$. This with $p^+ \in S^+$, $d \in -S$ gives $p^+(d) = 0$.

Since $f(x_1) < o$, if $p^+ \neq o$ one has $q^+ (h(x_1)) \geq -p^+(f(x_1)) > o$

and $q^+ \neq o$. ∎

[14] Remarks: (1) This last result is very much like Lemma [2:3] of Ritter (1969,b). Again there is no reason why A should not be multivalued linear [4] with the condition $f(x_1) \cap -S^o \neq \phi$.

(2) These results ([10] - [13]) extend many in Mangasarian (1969).

(3) Craven and Mond (1973) prove [12] with B = 0 and F single valued. They invoke an extra, and redundant hypothesis that h have an injective transpose h^*. This is implied by the hypothesis that h is open.

Proof: Suppose $h^*(y') = o \quad y' \in Y'$.

Since $h^*(y')(x) = h(x)(y') \quad \forall x \in X$

one sees that $h(x)(y') = o \ \forall x \in X$. In any topological vector spaces h open implies $R(h) = Y$ so that $h(x)(y') = o$ for all $x \in X$ means that $y' \in Y^+ = o$.∎

Linear transposition theorems

The convex transposition theorems of the previous section suffer from the fact that when linear problems such as

$$Ax > o, \quad Bx \geqslant o, \quad Cx = o$$

have no solution one can deduce that

$$a^+ A + b^+ B + c^+ C = o \qquad a^+ \geq o \quad b^+ \geq o$$

but not $a^+ \neq o$.

The following general linear alternative theorem deals with this problem and is used to derive as corollaries most of Mangasarian's (1969) alternative theorems in a more general setting. The next paragraphs are a necessary preliminary. Dual spaces will from now on always be supposed to be endowed with the strong topology $B(X',X)$ unless it is stated otherwise.

[15] Proposition: (1) If C and D are closed convex cones in a locally convex space X

$$(C \cap D)^+ = \overline{C^+ + D^+}.$$

(2) If in addition X is normed and $C^o \cap D \neq \emptyset$

$$(C \cap D)^+ = C^+ + D^+.$$

Proof: This is essentially proved in Ritter (1969,a).|

Another condition under which $(C \cap D)^+ = C^+ + D^+$ is given by the following result.

[16] Definition: A set $C \subset X$ is said to be radial to x_o ($x_o \in \text{rad } C$) if for each $x \in X$ there is some $t_o > o$ such that $(1-t)\, x_o + t\, x \in C$ when $o < t < t_o$. A convex set is radial at any interior point.

[17] Proposition: Let X be a barrelled, (sequential), Hausdorff, convex space. Let C, $D \subset X$ be closed convex cones such that $\bar{x} \in \text{rad}(C) \cap D$. Then

$$(C \cap D)^+ = C^+ + D^+.$$

Proof: It suffices to show that $C^+ + D^+$ is closed; suppose that $\{b_n^+\}$ is a sequence in $C^+ + D^+$ with $b_n^+ \rightarrow b^+$, $b_n^+ = c_n^+ + d_n^+$, $c_n^+ \in C^+$ and $d_n^+ \in D^+$. Since $\bar{x} \in C \cap D$

$$b_n^+(\bar{x}) \geq c_n^+(\bar{x}) \geq 0.$$

Also, since $b_n^+ \rightarrow b^+$, $\{b_n^+(\bar{x})\}$ is a bounded set. Suppose $x \in X$. Since C is a convex cone with $\bar{x} \in \text{rad } C$ there is some $t = t(x)$ with $\bar{x} + tx$ and $\bar{x} - tx$ in C. This means that

$$c_n^+(\bar{x}) \geq t\, c_n^+(x) \geq -\, c_n^+(\bar{x})$$

This shows that $E = \{c_n^+\}$ is $\sigma(X', X)$ bounded. The hypothesis of barrelledness means that E is equicontinuous which in turn implies that $\bar{E}$ is $\sigma(X', X)$ compact. (Robinson and Robinson (1964)).

Thus one has a subsequence (not distinguished notationally) with $c_n^+ \rightarrow c^+$. Since $\overline{E} \subset C^+$ and C^+ is weakly closed $c^+ \in C^+$.

Similarly $d_n^+ \rightarrow b^+ - c^+$ which must belong to D^+. Thus

$$b^+ = c^+ + (b^+ - c^+) \in C^+ + D^+$$

and $C^+ + D^+$ is a closed set. ∎

Ritter's proof of [15] (2) relies on the existence of a point in $C^o \cap D$. Cones can exist, in normed spaces, which have no interior but have a radial point. Since a normed space need not be barrelled when it is incomplete one sees that [15] (2) and [17] are not strictly comparable. In Banach spaces, though, [17] will include [15].

[18] The general linear theorem can now be proved. It is stated in normed spaces for simplicity.

<u>Theorem</u>: Let X_i $i=0,\ldots,4$ be normed spaces with $A_i \in B[X_o, X_i]$. Suppose X_o and each $R(A_i)$ is a Banach space and that there are closed convex cones $S_i \subset X_i$ $i=1,\ldots,4$ with $S_i^o \cap R(A_i) \neq \emptyset$ $i = 2, 3$ and $S_4 = 0$. Suppose S_2 is pointed and that

$$(*) \quad A_2^{-1}(S_2)^+ + A_3^{-1}(S_3)^+ + N(A_4)^{\perp} \text{ is closed.}$$

Then either there is a solution to

$$(1) \quad A_1 x \in S_1^o \;, \quad A_2 x \in S_2 / \{0\} \;, \quad A_3 x \in S_3, \; A_4 x = 0,$$

or to

$$(2) \quad a_1^+ A_1 + a_2^+ A_2 + a_3^+ A_3 + a_4^+ A_4 = 0 \,,$$

where $a_i^+ \in S_i^+$ $i=1,\ldots,4$ and either (a) $a_1^+ \neq 0$ or (b) for any fixed $s_2^+ \in S_2^+$ a_2^+ can be chosen with $a_2^+ - s_2^+ \in S_2^+$.

If $(S_2^+)^o \neq \emptyset$ the alternative is strict.

<u>Proof</u>: $\Rightarrow$ Suppose (1) has no solution.

Set $E_1 = \{x \mid A_1 x > 0\}$; $E_2 = \{x \mid A_2 x \gtrsim 0, \; A_3 x \geqslant 0, \; A_4 x = 0\}$.

<u>Case(i)</u> $E_1 = \emptyset$; Using a standard separation argument there is no difficulty in satisfying 2(a) with $a_2^+ = a_3^+ = a_4^+ = 0$.

Case (ii): $E_2 = \phi$; This can be written, for a fixed $s_2^+ \in S_2^+$, as

$$A_2 x \geq o, \; A_3 x \geq o, \; A_4 x = o \Rightarrow s_2^+ A_2 x \leq o.$$

Then $E_3 = A_2^{-1}(S_2) \cap A_3^{-1}(S_3) \cap N(A_4) \subset (s_2^+ A_2)^{-1}(-S_2)$. Using (*) and [15] (1) one has

$$(3) \; - s_2^+ A_2 \in A_2^{-1}(S_2)^+ + A_3^{-1}(S_3)^+ + N(A_4)^\perp = (E_3)^+.$$

Farkas Lemma [6] can be used, since each $R(A_i)$ is a Banach space, to derive that $u_i^+ \in A_i^{-1}(S_i)^+$ implies $u_i^+ = d_i^+ A_i$ and $d_i^+ \in (R(A_i)) \cap S_i)^+$. The condition $R(A_i) \cap S_i^o \neq \phi$ $i = 2,3$ enables one to use [15] (2) and write

$$d_i^+ = a_i^+ + b_i^+ \quad a_i^+ \in S_i^+, \quad b_i^+ \in R(A)^\perp \; i = 2,3$$

while d_4^+ can be replaced by a_4^+ by an application of the Hahn-Banach theorem. Clearly $a_i^+ A_i = d_i^+ A_i$ $i = 2,3,4$. Using (3) one has

$$(4) \; -s_2^+ A_2 = u_2^+ + u_3^+ + u_4^+ = a_2^+ A_2 + a_3^+ A_3 + a_4^+ A_4$$

Setting $a_1^+ = o$, one has a solution to (2)(b).

Case (iii): $E_1 \neq \phi$, $E_2 \neq \phi$.

By assumption $E_1 \cap E_2 = \phi$. E_1 is clearly convex with interior The convexity of E_2 follows from the pointedness of S_2. The Hahn-Banach theorem asserts the existence of $u^+ \in E_2^-$ and E_1^+. It is easy to show that $E_2^+ = E_3^+$ and then, as in (4), to write

$$(5) \; - u^+ = a_2^+ A_2 + a_3^+ A_3 + a_4^+ A_4.$$

Moreover, since $u^+ \in E_1^+$, $A_1 x > o \Rightarrow u^+(x) \geq o$. Since there is some x_1 with $A_1 x_1 > o$ one has in fact that $A_1 x \geq o$ implies $u^+(x) \geq o$. (To see this let $x_\lambda = x + \lambda x_o$, $\lambda > o$. Then $A_1(x_\lambda) > o$. Hence $u^+(x_\lambda) \geq o$ and $u^+(x) \geq o$). The same argument as before now shows that $u^+ = a_1^+ A_1$, $a_1^+ \in S_1^+ / \{o\}$ and (5) provides a solution to 2(a).

$\Leftarrow$ When $(S_2^{+})^{0} \neq \phi$ and both (1), (2) have solutions one has

(6) $o = a_1^{+} A_1x + a_2^{+} A_2x + a_3^{+} A_3x + a_4^{+} A_4x \geq a_i^{+} A_ix$, $i = 1,2$.

If $a_1^{+} = o$ the fact that $A_1x > o$ produces a contradiction in (6) while otherwise a_2^{+} can be chosen to belong to $S_2^{+\,o}$ and $a_2^{+} A_2x > o$ which again contradicts (6).▌

[19] Remark: Banach space hypotheses are only needed for the Farkas lemma while normed conditions are only used for $(R(A_i) \cap S_i)^{+} = R(A_i)^{+} + S_i^{+}$. It is therefore possible to prove [18] with the spaces sequential barrelled and fully complete with the operators with barrelled range. This merely uses [17] instead of [15] (2).

[20] Conditions for (*) to hold can easily be derived from [15] (2) or [17] Another condition is given by

Proposition: (*) is closed when S_2, S_3 are polyhedral cones ([2.19]) and A_4 has finite dimensional range.-

Proof: Since S_2 is polyhedrally convex

$$A_2^{-1} (A_2) = \{x \mid y_j^{+} (A_2 x) \geq o \ j = 1,\ldots,m\}$$
$$\{x \mid (A_2^{*} (y_j^{+}) (x) \geq o \ j = 1,\ldots,m\}$$

which is polyhedral. Similarly $A_3^{-1} (S_3)$ is. Using the simplest Farkas Lemma $A_2^{-1} (S_2)^{+}$, $A_3^{-1} (S_3)^{+}$ are finitely generated.

Since $N(A_4) = R(A_4^{*})$ and $\dim R(A_4^{*}) \leq \dim D(A_4^{*})$

$A_2^{-1} (S_2)^{+} + A_3^{-1} (S_3)^{+} + N(A_4)$ is finitely generated and hence closed.▌ Ritter's alternative theorem (1969, b) is included as the case of [18] in which X_2, X_3, X_4 are finite dimensional Euclidean spaces and S_2, S_3 are the orthant orderings. The proposition shows that (*) is satisfied since the orthants are polyhedral.

[21] <u>Theorem</u>: (Generalised Gale's Theorem) Let X, Y be normed spaces and let $A \in B[X,Y]$ have closed complete range. Suppose $S \subset X$ is a closed convex cone such that $N(A)^{\perp} + S^{+}$ is closed.

Then for any $q^{+} \in Y'$ either

(1) $Ay = o \quad q^{+}(y) = -1 \quad y \in S$

or

(2) $A^{*} y' \leq_{S+} q^{+}$ (A^{*} is the adjoint of A.)

has solution but not both.

<u>Proof</u>: $\Rightarrow$ Suppose (1) has no solution then there is no solution to

(1)' $Ay = o, \ -q^{+}(y) < o, \ y \in S$

Applying the theorem of [18] to (1)', and noting that $S^{o} \neq \phi$ is not needed since the operator $Iy = y$ is surjective, one has

$$- rq^{+} + s^{+} + q^{+} A = o$$

with $r > o$, $s^{+} \in S^{+}$, $a^{+} \in Y'$. This can be rewritten as

$A^{*} (r^{-1} a^{+}) = q^{+} - r^{-1} s^{+} \leq q^{+}$. ▮ This theorem includes Gale's equality theorem (Mangasarian (1969)) since S = Y is a perfectly good candidate. The closure condition is met by any finite dimensional map and any polyhedral cone.

[22] It seems worth noting that at least for some convex functions an anologue of [18] can be proved. Suppose $f: X \to Y$ is convex with respect to S_1 and that $f(o) = o$, $f'(o)$ exists and has closed range.

<u>Theorem</u>: With all terms as in 18 and f satisfying the condition above either there is a solution to

(1) $f(x) \in -S_1^{o}, \ A_2 x \in S_2 / \{o\}, \ A_3 x \in S_3, \ A_4 x = o$

or multipliers exist as described in [18] with

(2) $- a_1^{+} f(x) + a_2^{+} A_2 x + a_3^{+} A_3 x + a_4^{+} A_4 x \geq o \ \forall x \in X.$

Proof: If there is no solution to (1) there is also no solution to

$$(1)' \quad -f'(o)\, x > o,\ A_2 x \gneqq o,\ A_3 x \geqq o,\ A_4 x = o.$$

[18] can be applied to this yielding .

$$(3) \quad -a_1^+ f'(o) + a_2^+ A_2 + a_3^+ A_3 + a_4^+ A_4 = o$$

Since f is convex with respect to S_1 and $f(o) = o$

$$a_1^+ f(x) \geqq a_1^+ f'(o)(x)$$

which gives the desired conclusion when substituted in (3).∎

[23] A generalised form of Steimke's Lemma is an easy consequence of the transposition theorems.

Proposition: Let X,Y be normed spaces with $A \in B[X,Y]$ having closed, complete range. Let $S < X$ be a closed convex cone with interior. Then either (1) or (2) is solvable but not both.

$$(1) \quad A^* y' \in S^+/\{o\}$$

$$(2) \quad A\, x = o,\ x \in S^o.$$

Proof: If (2) has no solution then for some $s^+ \in S^+/\{o\}$

$$s^+ + a^+ A = o$$

or equivalently $A^*(-a^+) = s^+ \in S^+/\{o\}$. If (1) and (2) have solutions y' and x respectively one has

$$o = {}'Ax)\ (y') = (A^+ y')(x) > o.$$∎

In R^n a derivation of Tucker's Theorem (Mangasarian (1969)) is now an easy corollary.

Chapter Four

SUBGRADIENTS, TANGENT CONES AND LOCAL SUPPORTABILITY

Subgradients, Tangent Cones and Local Supportability

[1] Definition: If $f:X \to Y$ and S is a closed convex cone in Y then the subgradient $\partial_S f(x)$ at $x \in X$ is the set of linear continuous mappings $\bar{Z}:X \to Y$ satisfying

(*) $f(y) - f(x) \geqslant_S \bar{Z}(y-x) \quad \forall y \in X.$

When $\partial_S f(x)$ is not empty f is said to be subdifferentiable at x. If it is unambiguous $\partial_S f(x)$ will be written $\partial f(x)$. f is said to have supergradient $\bar{\partial}_S f(x)$ at x if $-f$ has subgradient $\partial_S -f(x)$.

Proposition: Let S be any closed convex cone and $f:X \to Y$. Then:

(1) $\partial_S f(x)$ is a convex set in $B[X,Y]$ and is closed in the weak operator topology.

(2) $T \subset S \Rightarrow \partial_T f \subset \partial_S f$.

(3) $u^+ \in S^+ \quad \Longrightarrow \quad u^+\partial_S f \subset \partial u^+ f$.

Proof: These all follow directly from [1].∎

The notion of subgradients originates in the study of convex functionals (Rockafellar (1970,a), Fenchel and others) and their importance arises largely from the following composite theorem. It is stated in R^n but retains much of its validity more generally.

[2] Theorem: Let $f:R^n \to R$ be a (proper) convex function.

(1) For $x \notin \text{dom}f$ $\partial f(x) = \phi$;

For $x \in \text{ri}(\text{dom}f)$ $\partial f(x) \neq \phi$.

(2) $x \in (\text{dom}f)^o \Longleftrightarrow \partial f(x)$ is closed and bounded and non empty.

(3) $\partial f(x)$ is everywhere single valued on $\operatorname{dom} f \Longleftrightarrow \partial f(x) = \{f'(x)\}$. ▮

In this chapter a few existence theorems for subgradients of convex functions with respect to S are given. These are followed by some tangent cone properties associated with subgradients.

[3] Proposition: If $f = X \to Y$ is convex with respect to S, a pointed convex closed cone and $f'(x)$ exists, then $\{f'(x)\} = \partial_S f(x)$.

Proof: Now

$$\lim_{t \to o} \frac{f(x + th) - f(x)}{t} = (f'(x_o))\ (h)$$

and since for $T \in \partial_S f(x)$ and $t > o$

$$\frac{f(x + th) - f(x)}{t} - T(h) \in S$$

one has $[f'(x) - T](h) \in S$ for all $h \in X$.
This means that $[f'(x)-T]\ (h) \in S \cap - S = \{0\}$ for all h and $f'(x) = T$. ▮

From now on $\{f'(x)\}$ and $f'(x)$ are identified.

When $X = R^n$ and $Y = R$ the existence of a single valued $\partial f(x)$ is also sufficient for $f'(x)$ to exist and equal $\partial f(x)$. This and most other properties of subgradients rely on relationships between support hyperplanes and convex sets. Since these do not appear to have natural extensions to planes of greater deficiency it seems unlikely that much can be said in general about subgradients of convex functions.

One case in which the situation is radically different is when Y is a real sequence space and S is the co-ordinate cone. In this case f is convex exactly when each co-ordinate is convex and (*) of [1] is satisfied by a function $T = \{Z_i\}$ exactly when each Z_i satisfies (*) for the corresponding real valued f_i. The situation can then be read off from [2].

Some properties which hold more generally are given below. A few preliminaries are necessary.

[4] Definition: If $A \subset R^n$ and $B \subset R^m$ are convex cones then A will be said to be polygonal with respect to B (denoted ApB) if there is some m by n matrix K with $x \in A \Longleftrightarrow Kx \in B$.

[5] Proposition: If ApB and K is surjective then

$$y^+ \in A^+ \Longleftrightarrow y^+ = K^T c^+,\ c^+ \in B^+.$$

Proof: If $y^+ \in A^+$ then $y^+(x) \geqslant o\ \forall x \in A$. Thus

$$Kx \in B \Rightarrow y^+(x) \geqslant o.$$

By the Farkas Lemma ([3.6]) $y^{+T} = c^{+T} K$ for some $c^+ \in B^+$. ∎

[6] Proposition: (1) ApB, BpC $\Rightarrow$ ApC.

(2) Suppose A and B and pointed and ApB, BpA.

Then, K of definition [4] is invertible.

Proof: There are matrices K and L with K $m \times n$, L $n \times m$ and

$$x \in A \Longleftrightarrow Kx \in B;\ x \in B \Longleftrightarrow Lx \in A.$$

Since A is pointed $Kx = o \in B \cap -B \Longleftrightarrow x \in A \cap -A = \{o\}$

Similarly since B is pointed $Lx = o \Longrightarrow x = o$. This in turn implies that

$$\text{rank } K = n \leqslant m = \text{rank } L \leqslant n$$

Thus $m = n$ and since a square matrix of full rank is invertible $K(L)$ is invertible.▮

[7] <u>Proposition</u>: If ApB and A and B are related by an invertible K

then $\quad \bar{x} \in \partial_A f(x) \Longleftrightarrow K\bar{x} \in \partial_B Kf(x)$.

<u>Proof</u>: $\bar{x}(h) \geqslant_A f(x + h) - f(x) \Longleftrightarrow$

$$K\bar{x}(h) \geqslant_B Kf(x + h) - Kf(x)$$

since ApB. Because K is invertible, when

$$\bar{y}(h) \geqslant_B Kf(x + h) - Kf(x)$$

then $\bar{y} = K\bar{x}$ for some $\bar{x} \in R^n$ and

$$\bar{x}(h) \geqslant_A f(x + h) - f(x).$$ ▮

Since, as was remarked above, the subgradient of a function $f:R^m \to R^n$ with the orthant ordering is completely determined by its co-ordinates' behaviour and subgradients [7] only extends this result to any closed convex cone in R^n generated by n distinct half spaces, i.e. - generated by n linearly independent constraints.

[8] <u>Proposition</u>: Let $f:R^m \to R^n$ be convex with respect to a cone S which is specified by $x \in S \Longleftrightarrow Kx$ has each column non negative for an invertible matrix $K \subset R^{n^2}$. Then all the properties of $\partial f(x)$ asserted in [2] hold for $\partial_S f(x)$.

Proof: $x \in ri(domf) \Leftrightarrow Kx \in ri(domKf)$ since K is invertible. Similarly $x \in (domf)^o \Leftrightarrow Kx \in (domKf)^o$. All the conclusions now follow from the above discussion except for the boundedness of $\partial_S f(x)$ which follows from

$$\| \bar{x} \| \leqslant |K^{-1}| \, \| K \bar{x} \|$$

and from [7] with $B = R^{n+}$. ∎

[9] Proposition: If all hypotheses are as in [8] and $\partial_S f(x)$ is single valued then $\partial_S f(x) = f'(x)$.

Proof: This also follows from the previous discussion. ∎

It is possible to extend the result in [8] so that the cone S is generated by n + 1 half-spaces.

[10] Proposition: Suppose all the hypotheses of Proposition [8] hold and that $S_1 = S \cap \{x \mid \bar{a}(x) \geqslant 0\}$ then $\partial_{S_1} f(x)$ is non empty on ri(domf).

Proof: f is supposed to be convex with respect to S_1 and hence with respect to S. Thus $\exists \bar{Z} = (\bar{z}_1, \ldots, \bar{z}_n) = K^{-1}(\bar{y}_1, \ldots, \bar{y}_n) \in \partial_S f(x)$ when $x \in ri(domf)$ and $(\bar{y}_1, \ldots, \bar{y}_n) \in \partial Kf$.

If the rows of K are denoted by $\bar{a}_1, \ldots, \bar{a}_n$ then the a_i are independent and $\bar{a} = r_1 \bar{a}_1 + \quad + r_n \bar{a}_n$ $r_1, \ldots, r_n \in R$.

Since, by a theorem of Rockafellar (1970,a),

(1) $\partial g(x) + \partial h(x) = \partial (g + h)(x)$ when $x \in ri(domg) \cap ri(domh)$

this also implies that

(2) $\partial(g - h)(x) \subset \partial g(x) - \partial h(x)$ when $g, h, g-h$, are convex.

(1) and (2) mean that

(3) $\partial \bar{a}f = \partial(r_1\bar{a}_1 f + \ldots + r_n\bar{a}_n f) \subset r_1\partial\bar{a}_1 f + \ldots + r_n\partial\bar{a}_n f$

since $\bar{a}f, \bar{a}_1 f, \ldots, \bar{a}_n f$ are all convex.

By the previous results $\partial\bar{a}f(x) \neq \emptyset$ when $x \in ri(domf)$. Let $\bar{y} \in \partial\bar{a}f(x)$. Using (3)

$$\bar{y} = r_1\bar{y}_1 + \ldots + r_n\bar{y}_n \quad \bar{y}_i \in \partial\,\bar{a}_i f$$

$$= r_1K\bar{z}_1 + \quad + r_nK\bar{z}_n = \bar{a}\bar{Z}\,.$$

Then $\quad \bar{a}\bar{Z}(h) \leqslant \bar{a}f(x + h) - \bar{a}f(x)$

and from above $\quad \bar{a}_k\bar{Z}(h) \leqslant \bar{a}_k f(x + h) - \bar{a}_k f(x)$.

Thus $\quad K_1\bar{Z} \in \partial K_1 f(x) \quad$ where K_1 is $\begin{bmatrix} K \\ \bar{a} \end{bmatrix}$

and $\quad \bar{Z} \in \partial_{S_1} f(x)$. ∎

This argument cannot be extended inductively since there is no guarantee that the vectors chosen from (3) would agree for two different $\bar{a}$ and $\bar{b}$. However, as an easy corollary one has:

[11] <u>Proposition</u>: If $S = \{x \mid \bar{a}_t(x) \geqslant 0,\ t \in T\}$ and there are $\bar{a}_{t_1}, \ldots, \bar{a}_{t_n}$ with

$$(2) \quad \partial(\bar{a}_t f) = \sum_{i=1}^{n} r_i\partial(\bar{a}_i f) \qquad t \neq t_1, \ldots, t_n \quad ,$$

then $\partial_S f(x) \neq \emptyset$ if $x \in ri(domf)$.

Proof: (2) guarantees that the extra functionals in S are irrelevant.▮

In particular (3) of the proposition in [1] guarantees that for any cone S_2 containing S in [8], [10] or [11]

$$\partial_{S_2} f(x) \neq \emptyset \text{ when } x \in \text{ri dom} f.$$

For the remainder of the Chapter it will be assumed that f has a non trivial subgradient at the relevant points. From the previous discussion when f is convex this is at worst a requirement that $f'(x)$ exists and if often weaker.

Tangent cones and subgradients.

[12] Proposition: Let $f:X \rightarrow Y$. Then x is a strong minimum for f with respect to S if and only if $0 \in \partial_S f(x)$.

Proof: $\Rightarrow f(y) - f(x) \in S \quad \forall x \in \text{dom} f$ if x is a minimum. This means $0 \in \partial_S f(x)$. $\Leftarrow$ If $0 \in \partial_S f(x)$

$$f(y) - f(x) \geqslant_S 0(y - x) = 0 \quad \forall y \in X$$

and x is a strong minimum.▮

[13] Proposition: Let $f:X \rightarrow Y$ and let x be a strong maximum for f over A with respect to S. Then for any $\overline{Z} \in \partial_S f(x)$

$$\overline{Z}\ [wP(A,x)] \subset -S.$$

Proof: Let $\lambda_n(x_n - x) \rightarrow h \in wT(A,x)$. By assumption

$$0 \geqslant f(x_n) - f(x) \geqslant \overline{Z}(x_n - x) \text{ since } x_n \in A.$$

Thus $\lambda_n \bar{Z}(x_n - x) = \bar{Z}\lambda_n(x_n - x) \in -S$ since $\lambda_n \geq 0$.

By definition $\bar{Z}$ is continuous, hence

$$\bar{Z}(h) = \lim \bar{Z}\lambda_n(x_n - x) \in -S,$$

since S is a closed convex cone and thus weakly closed. ∎

[14] Proposition: Let $g:X \to Y$, S be a closed convex cone in Y and $\bar{Z} \in \partial_S g(x)$. Then $\bar{Z}\,[wT(g^{-1}(-S),x)] \subset P(-S,g(x))$.

Proof: let $h \in wP(g^{-1}(-S),x)$ Then $h = w\lim h_n$ where $h_n = \lambda_n(x_n - x)$, $g(x_n) \in -S$, $x_n \to x$ and $\lambda_n \geq 0$. Since g is subdifferentiable

$$\lambda_n\,[g(x_n) - g(x)] \geq \bar{Z}(h_n).$$

For each n, the left hand side belongs to $\lambda_n(-S - g(x))$ and thus for some net $\{s_n\}$ in S

$$\bar{Z}(h) = w\lim(\lambda_n\,[g(x_n) - g(x) - s_n] \in \overline{\bigcup_{\lambda > 0} \lambda(-S - g(x_0))}$$

By [2.12] this last set is $P(-S,g(x))$. ∎

[15] Proposition: Let $g:X \to Y$. Suppose g has a supergradient at X with respect to a closed convex cone S with interior and suppose $g^{-1}(-S)$ is convex. Then if there is $\bar{Z} \in \partial_S g(x)$ and $h_0 \in X$ with

(1) $\bar{Z}(h_0) + g(x) \in -S^0$,

it follows that

(2) $h \in P(g^{-1}(-S),x)$ whenever $\bar{Z}(h) \in P(-S,g(x))$.

Proof: Suppose $\bar{Z}(h) \in P(-S,g(x))$ then there are $\lambda_n \geq 0, s_n \in S$ with $-s_n \to g(x)$ such that

$\lambda_n(-s_n-g(x)) \rightarrow \overline{Z}(h)$.

Let $h_t = h + th_o$ $0 < t < 1$. Then

$$\overline{Z}(h_t) + tg(x) < \overline{Z}(h)$$

and for n sufficiently large

$$\overline{Z}(h_t) + tg(x) < \lambda_n(-s_n - g(x)) \leqslant -\lambda_n g(x).$$

Since $\overline{Z} \in \partial_S g(x)$

$$g(x + rh_t) - g(x) < -r(\lambda_n + t)\, g(x) \quad , \quad r > 0.$$

For $r < r_o$ the right hand scalar will be larger than $-\frac{1}{2}$ and

$$g(x + rh_t) < \tfrac{1}{2}g(x) \leqslant 0.$$

Thus $x + rh_t \in g^{-1}(-S)$ or $h_t \in r^{-1}(g^{-1}(-S)-x)$ which is contained in $P(g^{-1}(-S),x)$ since $g^{-1}(-S)$ is assumed convex. Letting $t \rightarrow o$ one has $h \in P(g^{-1}(-S),x)$.▮

The condition that $g^{-1}(-S)$ be convex is satisfied if g is convex with respect to S. In the terminology of [2;36] [15] says $K(\overline{Z})$ is contained in $P(g^{-1}(-S),x)$, except that $\overline{Z}$ need not belong to $\mathcal{E}$.

The last four paragraphs have partially illustrated the relationships between tangent cones and subgradients. They will be used in combination with various other results to derive necessary conditions for constrained minima. They are rather unsatisfactory, though, since both [13] and [15] require supergradients for minima rather than subgradients. This will be discussed further in the next chapter in the sections on one sided derivatives and stationary point theorems for convex functions.

Local Supportability

[16] Definition: A set $A \subset X$ is locally supportable at x if there is a neighbourhood N of x and a continuous linear function u^+ such that $u^+(x) \leq u^+(y) \quad \forall y \in A \cap N$.

By a theorem of Valentine (1964) if A is locally supportable at each boundary point of A and A is connected with interior then A is convex.

In the proofs of [13] to [16] it is apparent that it would suffice for the subgradient inequality to hold locally. This suggests the following generalization which has been investigated for real valued by Bazaraa et al.(1970).

[17] Definition: $f:X \to Y$ is locally supportable from below at x with respect to S if there is a $\bar{Z} \in B[X,Y]$ and a neighbourhood N of x such that

$$f(y) - f(x) \geq_S \bar{Z}(y - x) \quad \forall y \in N.$$

[18] Theorem: (Bazaraa(1971)). For a continuous function $f:R^n \to R$, f being locally supportable from below at x by $\bar{Z}$ is equivalent to any of the following

(1) $(x, f(x)) \in \mathrm{bd}[\mathrm{Epi}f \cap N]$ for some neighbourhood N of $(x,f(x))$.

(2) $(-\bar{Z}, 1) \in P^+([\mathrm{Epi}f \cap N], (x,f(x))$ for some neighbourhood N of $(x,f(x))$.

(3) Epi f is locally supportable at $(x,f(x))$. ▮

With the appropriate interior conditions on $[\mathrm{Epi}f \cap N]$ the theorem can be extended to any convex domain space X. When f is no longer required to be real valued the same problem that occurs for subgradients occurs here, in that the type of equivalence given by the last theorem seems unobtainable. The equivalences do suggest the next two propositions, though.

[19] Proposition: If X is a normed space and $f: X \to R$ satisfies

(1) $\limsup_{y \to x} \frac{f(y) - f(x)}{\| y - x \|} > -\infty$,

(2) x maximizes f over A,

(3) $(-\bar{Z}, 1) \in wP^{+}(\mathrm{Epi}f, (x, f(x)))$,

then

$-\bar{Z} \in wP^{+}(A, x)$.

Proof: Let $h \neq 0 \in wT(A, x)$. As usual $h_n \to h$ where $h_n = \lambda_n(x_n - x)$, $x_n \in A$, $x_n \to x$ and $\lambda_n \geq 0$. Using (1) there is a subsequence of $\{x_n\}$ (which will not be distinguished notationally) with

$$0 \geq \frac{\lambda_n \; f(x_n) - f(x)}{\lambda_n \; \| x_n - x \|} > -M \qquad n \geq n_o.$$

Since $h_n \to h$, $\lambda_n \| x_n - x \|$ is bounded and $k_n = \lambda_n(f(x_n) - f(x))$ must also be bounded and has a convergent subsequence (which again will not be relabeled) with $k_n \to k_o$.

Then

$$(h, k) = \mathrm{wlim}\, \lambda_n(x_n - x, f(x_n) - f(x)).$$

Thus $(h,k) \in wP(Epi f, (x,f(x))$. Using (3)

$\overline{Z}(h) \leq k.$

Since $f(x_n) \leq f(x)$, $k \leq 0$ and $\overline{Z}(h) \leq 0$. Hence

$-\overline{Z} \in wT^+(A,x) = wP^+(A,x)$. ▮.

Condition (1) is guaranteed by local supportability while condition (3) is very close to [18] (2).

[20] Proposition: Let X,Y be reflexive normed spaces. Suppose $g: X \to Y$ is continuous and that B is any set in Y containing o. Suppose the following conditions hold.

(1) $\limsup\limits_{y \to x} \frac{\| g(y) - g(x) \|}{\| y - x \|} < \infty$

(2) $\exists\ \overline{Z} \in B[X,Y]$ with

$(\overline{Z}, -I)\left[wP(Epi_B g, (x,g(x)))\right] \subset wP(B,g(x)).$

Then $\overline{Z} \in w\mathcal{E}$ (in the notation of [2.31]).

Proof: Let $h \in wT(g^{-1}(B),x)$. Then $h_n = \lambda_n(x_n - x) \to h$ with $x_n \to x$, $g(x_n) \in B$ and $\lambda_n \geq 0$. As in [19](1) can be used to derive the existence of a subsequence (not relabeled) of $\lambda_n(g(x_n) - g(x))$ tending weakly to k. (This uses the reflexivity of Y and the weak compactness of the unit ball in a reflexive space.)

$(h,k) \in wP(Epi_B g, (x,g(x)))$ since $(x_n, g(x_n)) \in Epi_B g$.

Since g is continuous

$k = w\lim \lambda_n(g(x_n) - g(x)) \in wP(B,g(x)).$

Using (2) $\overline{Z}(h) \in k + wP(B,g(x)) \subset wP(B,g(x))$. Thus

$\overline{Z}(wT(g^{-1}(B),x)) \subset wP(B,g(x))$

and since $\overline{Z}$ is supposed linear and continuous

$\overline{Z}(wP(g^{-1}(B),x)) \subset wP(B,g(x))$

as desired. ▮

This last proposition illustrates the problem associated with (local) supportability of epigraphs when f is not real valued. It seems necessary to impose rather strong conditions like (2) which will be easy to verify on the line when I = 1 but are not so convenient in general.

[21] Examples: (1) Let $C \subset R^n$ be a convex set and let $f: R^n \to R$ be defined by $f(x) = 0 \quad x \in C$; $f(x) = \infty \quad x \not\in C$.

$$\partial f(x) = \{ \bar{z} \mid \bar{z}(z - x) \leq 0 \quad \forall z \in C \} = -(C - x)^+.$$

(2) $f(x) = x^3$ is a real valued differentiable function which is not locally supportable at 0.

(3) The duality map J of a Banach pace X into its dual X' defined by

$$J(x) = \{ x' \in X' \mid x'(x) = \|x\| \| x'\| , \ \|x\| = \|x'\| \}$$

is the subgradient of the convex mapping $g(x) = \frac{1}{2}\|x\|^2$. In a Hilbert space J is just the identity.

[22] The last result of this chapter gives a simple necessary condition for a set E to be pseudoconvex at point x. ([2.5]).

Proposition: $E \subset X$ is weakly pseudoconvex at $x \in E$ only if every hyperplane which supports E locally at x supports E globally at x.

Proof: Suppose $x' \in X'$ is such that for some neighbourhood N of x $x'(y) \geq x'(x)$ for all $y \in N \cap E$. Then if

$h \in wT(E,x)$

$$h_n = \lambda_n(x_n - x) \rightarrow h \text{ where } x_n \rightarrow x, \{x_n\} \subset E, \lambda_n \geqslant 0.$$

For $n \geqslant n_o$, $x_n \in N$ and $x'(x_n - x) \geqslant 0$. Since x' is continuous and linear

$$x'(h) = \text{wlim } x'(h_n) = \text{wlim } \lambda_n x'(x_n - x) \geqslant 0.$$

Thus $x' \in wT^+(E,x) = wP^+(E,x)$.

Suppose E is weakly pseudoconvex at x. Then if $y \in E$

$$y - x \in wP(E,x).$$

This together with $x' \in wP^+(E,x)$ gives

$$x'(y) \geqslant x'(x) \quad \forall y \in E. \blacksquare$$

Chapter Five

FIRST ORDER CONDITIONS

This chapter is concerned with first order necessary and sufficient conditions for the optimization problem

$$(P) = \min f(x) \quad \text{subject to } g(x) \in B,\ x \in C$$

where $f: X \to Y$, $g: X \to Z$, $B \subset Y$ is any set as is $C \subset X$. Minimization is taken with respect to a closed convex cone $S \subset Y$ and it will be indicated whether the minimum is weak or strong in each case. X, Y, Z are Hausdorff locally convex spaces unless otherwise noted.

[1] Notation: A, the set of feasible solutions of (P), is defined by

$$A = \{x \mid g(x) \in B,\ x \in C\}.$$

Δ will denote the inverse image of B under g

$$\Delta = \{x \mid g(x) \in B\}.$$

g will be called the constraint function.

f will be called the objective function.

Section one: Fritz John Conditions

The first section of this chapter is restricted to some generalizations of first order necessary conditions for (P) in which g is subject to no 'constraint' qualification. The basic theorems generalize results of Fritz John (1948), Mangasarianand Fromovitz (1967), Nagahisa and Sakawa (1969) and Zlobec and Massam (1973). The sets S, B will sometimes be supposed to closed convex cones with nonempty interior. Thenthe problem will be denoted by (Q) and written as

$$(Q) \quad \min f(x) \quad g(x) \in -B \ (\text{or } g(x) \leq_B 0) \quad x \in C.$$

Such theorems are generally called <u>Fritz John</u> type conditions.

[2] <u>Theorem</u>. Suppose x is a weak (local) minimum with respect to S for (Q). Suppose f, g are compactly differentiable at x. Then there exist $p^+ \in S^+$, $q^+ \in B^+$, not both zero, such that

$$p^+f'(x) + q^+g'(x) \in P^+(C,x) \qquad q^+(g'(x)) = 0.$$

<u>Proof</u>: Let $M = \{(y,z) \mid \exists h \in T(C,x) \text{ with } (f'(x)(h) \leq_S y, \; (g'(x))(h) \leq_B z\}$

$$N = \{(y,z) \mid y \in -S, \; z \in -B\}.$$

M and N are closed convex sets and $N^o \neq \emptyset$.

Suppose that $M \cap N^o \neq \emptyset$. There then exists $h \in T(C,x)$ with $(f'(x))(h) \in -S^o$ and $(g'(x))(h) + g(x) \in -B^o$.

Let $h_n = \lambda_n(x_n - x) \to h$ where $x_n \in C$, $x_n \to x$, $\lambda_n \geq 0$. By the definition of the compact derivative

(1) $$\lambda_n[f(x_n) - f(x)] = \frac{f(x + \lambda_n^{-1}h_n) - f(x)}{\lambda_n^{-1}} \to (f'(x))(h) \in -S^o.$$

Thus, for $n \geq n_o$, $\lambda_n[f(x_n) - f(x)] \in -S^o$. Since S^o is a cone

(2) $f(x_n) <_S f(x)$.

Moreoever, similarly

(3) $$\lambda_n[g(x_n) - g(x)] + g(x) \to (g'(x))(h) + g(x) \in -B^o.$$

For $n \geq n_1$

$$\lambda_n g(x_n) \in -B^o + (\lambda_n - 1)g(x).$$

Since $g(x) \in -B$, $(\lambda_n - 1)\, g(x) \in -B$ for $n \geq n_2$ (Since $h \neq 0$, λ_n can be assumed convergent to ∞) and

(4) $g(x_n) \in \lambda_n^{-1}(-B^o - B) \subset -B^o$.

Examining (2) and (4) one sees that for $n \geq n_3$

$$f(x_n) < f(x), \; g(x_n) \in -B, \; x_n \in C$$

which contradicts the weak minimality of x for (Q).

This means that $M \cap N^o = \emptyset$ so that the Hahn Banach theorem can be invoked to assert the existence of $(p^+,q^+) \neq o \in (Y', Z')$ with

$$p^+(y) + q^+(z) \geq p^+(s) + q^+(b) \quad (y,z) \in M, s \in -S^o, b \in -B^o.$$

In particular $p^+ \in S^+$, $q^+ \in B^+$ and

(5) $p^+((f'(x))(h)) + q^+((g'(x))(h)) + q^+(g(x)) \geq o \quad \forall h \in T(C,x)$

Letting $h = o$ one sees that $q^+(g(x)) \geq o$ which in conjunction with $g(x) \in -B$, $q^+ \in B^+$ gives $q^+(g(x)) = o$. Finally, since $q^+(g(x)) = o$ and $T^+(C,x) = P^+(C,x)$, (5) gives the desired conclusion. ∎

The proof of the following theorem corresponds closely to that of [2]. Note that X is implicitly a sequential space ([2.43]). ∎

[3] Theorem: Suppose in addition to the hypotheses in [2], f and g are supposed boundedly differentiable at x with $f'(x)$ and $g'(x)$ completely continuous. Then there exist $p^+ \in S^+$, $q^+ \in B^+$ not both zero with

$$p^+(f'(x)) + q^+(g'(x)) \in wP^+(C,x) \qquad q^+(g(x)) = o.$$

Proof: The relevant lines (1), (3) follow from the fact that $\{h_n\}$ is now a weakly convergent net and the argument in [2.26] and [2.41]. ∎

[4] Remarks: (1) Theorem [2] was proved in Banach spaces for real valued f by Nagahisa and Sakawa, and was sketched, with an incorrect statement, by Zlobec and Massam (1973) for convex spaces and real valued f.

(2) In return for strengthened differentiability assumptions [3] gives a stronger necessary condition since $wP^+(C,x) \subset P^+(C,x)$ and they need not be equal.

The next results are concerned with the specific case in which C is the null set of a mapping h. That is $h: X \to W$ and

$$C = \{ x \mid h(x) = o \} = N(h).$$

[5] <u>Theorem</u>: Let x be a weak local minimum (Q) and suppose that $h: X \to W$ is compactly differentiable at x with X fully complete and W barrelled. Suppose further that h satisfies the following <u>regularity condition</u>

(1) $R(h'(x))$ is closed in W.

(2) If $R(h'(x)) = W$, then $N(h'(x)) = P(N(h),x)$.

Then there are $p^+ \in S^+, q^+ \in B^+, w^+ \in W'$ not all zero with

$$p^+(f'(x)) + q^+(g'(x)) + w^+(h'(x)) = o \qquad q^+(g(x) = o.$$

Moreoever, in case (2) both w^+ and (p^+,q^+) can be assumed non zero.

<u>Proof</u>: By assumption $R(h'(x))$ is closed. If the range is not W then there is (by Hahn-Banach) some $w^+ \in W' / \{o\}$ with $w^+((h'(x))(y) = o \ \forall y \in X$. In this case $p^+ = q^+ = o$ and the given w^+ suffice in the conclusion. Otherwise let p^+, q^+ be as in [2] and define

$$K(y) = p^+((f'(x))(y)) + q^+((g'(x))(y))$$

Using [2] and (2)

$$y \in P(C,x) = N(h'(x)) \Rightarrow K(y) \geq o.$$

All the hypotheses of the Farkas Lemma [3.6] are satisfied and one can deduce that

$$K(y) + w^+((h'(x))(y)) = o \qquad \forall y \in X, \ w^+ \neq o$$

which gives the desired conclusion. ▌

[6] <u>Theorem</u>: Suppose in [5] f,g are boundedly differentiable and $f'(x), g'(x)$ are completely continuous; the regularity condition (2) can be replaced by

(2)′ If $R(h'(x)) \neq W$, then $N(h'(x)) = wP(N(h),x)$.

Proof: This is as in [5] using [3] rather than [2]. ▌

The theorem of [5], [6] generalizes results of Craven (1970), (1972) and Craven and Mond (1973). The demonstration of this relies on the next definition and proposition.

[7] Definition: A continuous linear map $B: X \to Z$ between Banach spaces is said to be adequate (Craven (1970)) if

(1) $R(B)$ is closed in Z,

(2) If $R(B) = Z$ then there is a continuous projection q of X onto $N(B)$.

[8] Proposition: If $h: X \to W$ is continuously Fréchet differentiable at x and $h'(x)$ is adequate then h satisfies the regularity condition of Theorem [5].

Proof: (1) of [7] is the same as (1) of [5]. Suppose that (2) of [7] holds. The image then satisfies the hypotheses of Halkin's "correction" theorem [2.37] with $T = (1 - q)X$. Using [2.39]

$$N(h'(x)) = P(N(h), x). ▌$$

Note that if W is finite dimensional $h'(x)$ need only be assumed to exist since one may then apply [2.38]. ▌

[9] Proposition: If $h: X \to W$ is continuous and affine then

$$P(N(h), x) = N(h'(x))$$

Proof: This follows from the fact that $h'(x)$ exists and

$$h'(x)(u) = h(u) - h(o) \qquad x, u \in X.$$

This means that when $(h'(x))(u) = o$, $h(u) = h(o)$ and

$$h(x + tu) = h(x) + t(h(u) - h(o)) = h(x)$$

Thus when $x \in N(h)$ and $(h'(x))(u) = o$, $u \in P(N(h), x)$ since

$u = \lim n\,(x + \frac{1}{n}u - x)$. The opposite containment is standard.▮

[10] Theorem: (Craven and Mond (1973)). This is [5] with Y = R, X and W Banach spaces, $f'(x)$ and $g'(x)$ Frechet and h continuous, affine and open or continuously differentiable with $h'(x)$ adequate.

Proof: This follows from [5], [8] and [9].▮

[11] The next results lists some other properties of adequate mappings which can be of use.

Proposition: (1) In [7] if R(B) is finite dimensional or X is a Hilbert space then a continuous projection $q: X \to N(B)$ exists.

(2) If h satisfies the conditions of [8] then when $h'(x)$ is surjective there is a bijection T of a neighbourhood N of o in $h'(x)^{-1}(o)$ onto a neighbourhood U of $x \subset N(h)$ with T and T^{-1} Fréchet differentiable at o and x respectively.

Proof: (1) follows from the fact that projections in Hilbert space are continuous. (2) is proved in Craven (1973, submitted).▮

[12] A generalization of [5] which utilizes a condition similar to the Guignard constraint condition ([16]) is given by the next result.

Theorem: Suppose everything is as in [5] with the provision that (2) is replaced by (2')

$$(2')\ R(h'(x)) = W \text{ implies } N(h'(x)) \cap G = P(N(h),x)$$

for some closed convex cone $G (\subset X)$ with interior.

Then there are p^+, q^+, w^+ as in [5] such that

$$p^+(f'(x)) + q^+(g'(x)) + w^+(h'(x)) \in G^+,\ q^+(g(x)) = o.$$

Proof: In the terminology of [5] one has (when $R(h'(x)) = W$)

$$y \in G,\ (h'(x))(y) = o \Rightarrow K(y) > o.$$

Thus there is no solution to

$$(h'(x))(y) = o,\ K(y) < o\ ,\ y \in G.$$

The transposition theorem [3.10] applies since $h'(x)$ is open and

$$\left[r\ p^+(f'(x)) + r\ q^+(g'(x)) + w^+(h'(x))\right](y) \geq o\ \forall y \in G$$

for $w^+ \in W'$, $r \geq o$. This is the desired result.▌

When $B^o = \phi$ in (Q) Craven (1973, in preparation) has proved, using the concept of adequacy, an extension to [11]. Since this condition on g is a constraint condition generalizations will be given in the section on Kuhn-Tucker type results.

Converse Duality

The following theorem generalizes results of Craven and Mond (1971) and Mangasarian (1969) using the result of [5]. The proof method is taken from Craven and Mond.

[13] Theorem: Suppose X and Z' are reflexive fully complete spaces and X', Z are barrelled spaces. Suppose $f: X \to R$ and $g: X \to Z$ are twice compactly differentiable and that $B \subset Z$ is a closed convex cone with $(B^+)^o \neq \phi$. Let (D_1) (D_2) denote the programs

(D_1) min $f(x)$ subject to $g(x) \in -B$

(D_2) max $f(x) + q^+(g(x))$ subject to $q^+ \in B^+, f'(x) + q^+(g'(x)) = o$

Suppose (x_o, q_o^+) is optimal for (D_2) and that

$M^* = \left[f''(x_o) + q_o^+ g''(x_o)\right]^*$ has zero kernel.

Then if (i) f is pseudoconvex and (ii) $q_o^+ g$ is quasiconvex or $\Delta = g^{-1}(-B)$ is pseudoconvex at x_o, x_o is optimal for D, and the optimal values agree.

Proof: Suppose (x_o, q_o^+) is optimal for (D_2). Applying [5] to (D_2) one has $u \geq o$, $v \in X'' = X$ $q \in (B^+)^+ = B$, not all zero, and

(1) $u\left[f'(x_o) + q_o^+ g'(x_o)\right] + M^*(v) = o$

(2) $u\, g(x_o) + (g'(x_o))(v) + q = o$

(3) $q_o^+(q) = o$.

These are just the necessary conditions of [5] rewritten. Note that [5] can be applied to (D_2) since B^+ has interior and $f'(x_o) + q_o^+ q'(x_o)$ satisfies the regularity condition of [5]. (Craven (submitted) claims that the condition on M is sufficient to prove adequacy of $f'(x_o) + q_o^+ g'(x_o)$;if this is not so then the condition needs to be added.)

Since (x_o, q_o^+) is feasible for (D_2) $f'(x_o) + q_o^+(g'(x_o)) = o$. (1) says then that $M^*(v) = o$ which by hypothesis means $v = o$.

If $u = o$ (2), in turn, forces $q = o$. Since all 3 cannot be zero $u \neq o$. (2) again implies that $g(x_o) = {}^1/_u q \in B$. (3) gives $q_o^+(g(x_o)) = o$.

(4) $(f'(x_o))(x - x_o) = -(q_o^+ g'(x_o))(x - x_o)$.

By hypothesis either $q_o^+ g$ is quasiconvex, which, with $q_o^+(g(x_o)) = o \geqslant q_o^+ g(x)$ when $g(x) \in -B$, gives

(5) $(f'(x_o))(x - x_o) \geqslant o \;\; \forall x \in \Delta$;

or Δ is pseudoconvex at x_o. This means that $(g'(x_o))(\Delta - x_o) \subset P(-B, g(x_o))$ which with $q_o^+ \in B^+$ and $q_o^+(g(x_o)) = o$ gives (5) again. Thus since f is pseudoconvex and (5) holds one has

$$f(x) \geqslant f(x_o) \text{ if } x \in \Delta \quad (g(x) \in -B).$$

Because x_o is feasible for (D_1) x_o is optimal. Moreoever, the extreme values agree because

$$f(x_o) + q_o^+(g(x_o)) = f(x_o) .\blacksquare$$

[14] Remarks: (1) $q_o^+ g$ can be pseudoconvex without Δ being pseudoconvex at x_o and conversely. Convexity guarantees both conditions.

(2) Craven and Mond's result (1971) is given with the hypotheses that f and g are convex.

(3) It is clearly possible to use Kuhn-Tucker type results to prove [13].

Section Two: Real valued objective functions

Kuhn and Tucker (1951) were the first to use the notion of a constraint qualification on g to guarantee the existence of a multiplier for the problem (M).

(M) $\min f(x) \quad g_1(x) \geq 0, \ldots, g_n(x) \geq 0, h_{n+1}(x) = 0, \ldots,$
$h_p(x) = 0$

where all the functions are real valued and are assumed to have continuous first partials at any optimal point.

[15] Definition: Kuhn-Tucker constraint condition. Let $x \in R^n$ satisfy the constraints of (M). The constraint qualification is said to hold at x, if for any $y \neq 0$ with

$$(g_i'(x))(y) \geq 0 \quad \forall i \in \{i \mid g_i(x) = 0\}$$
$$(h_j'(x))(y) = 0 \quad j = n+1, \ldots, p$$

y is tangent to an arc $\alpha(\theta)$ differentiable at x and contained in the constraint region. That is $\alpha(0) = x$, $\alpha'(0) = y$ and $\alpha(\theta)$ satisfies the constraints for $\theta < \theta_1$.

Kuhn and Tucker proved that if x satisfied this constraint condition

$$f'(x) - \sum_{i=1}^{n} u_i g_i'(x) + \sum_{i=n+1}^{p} z_i h'(x) = 0$$

for some $u_i \geq 0 \qquad i = 1, \ldots, n$ with $u_i g_i(x) = 0$.

In other words, in the terminology of [2] the multiplier associated with f could be assumed positive. Varaiya (1967), Guignard (1969), Zlobec (1970) and others have considered constraint qualifications for the problem (P) which are given in tangent cone terms and which generalize Kuhn-Tucker's condition. When f is no longer real valued certain problems arise in the attempt to produce an analogue of the real results. The results in this section assume that f is real valued and give analogous results to those of Guignard and Zlobec for weak cones.

[16] Theorem: Suppose x is a minimum for (P) with f,g boundedly differentiable and f real valued. Let

$$wH = \{ h \in X' \mid h = u^+ g'(x), u^+ \in wP^+(B,g(x)) \}$$

$$wK = \{ y \mid (g'(x))(y) \in wP(B,g(x)) \}.$$

Suppose wH is closed and G is a closed convex cone such that $wK \cap G \subset wP(A,x)$ and $wK^+ + G^+$ is closed, then there is some $u^+ \in wP^+(B,g(x))$ with

$$f'(x) - u^+(g'(x)) \in G^+.$$

The condition is also sufficient if

(1) G is a closed convex cone with $A - x \subset G$.

(2) A is weakly pseudoconvex with respect to Δ at x.

(3) f is pseudoconvex over A at x.

Proof: Guignard proved the necessary condition for strong tangent cones and Fréchet derivatives. Propositions [2.26] and [2.33] provide the only alterations necessary in the proof. The method will be shown later when a similar result is proved for more general objective functions. The sufficiency condition will also be proved more generally later. (2) is in fact more general than Guignard's condition which was A or Δ is pseudoconvex at x. ([2.5]) ∎

If the closure conditions are not assumed in [16] the following asymptotic version holds.

[17] Theorem: In the terminology of chapter 2 suppose $x \in A$ and

(1) $G \subset X$ is a closed convex cone with $E \in w\mathcal{E}$ and

$wK(E) \cap G \subset wP(A,x)$

(2) $D \in wP^+(A,x)$.

Then $\quad D \in \overline{wP^+(B,g(x)E + G^+}$.

Proof: $x \in wK(E) \cap G \Rightarrow D(x) \geqslant 0$ by (1) (2). Thus

$$E(x) \in wP(B,g(x)), x \in G \Rightarrow D(x) \geqslant 0 .$$

This is the same as

$$E(x) \in wP^+(B,g(x))^+, x \in (G^+)^+ \Rightarrow D(x) \geqslant 0.$$

By a proposition of Zlobec & Massam (1973) this last line is equivalent to the conclusion. In fact it relies on a separation argument which is independent of the nature of the closed convex cones. ∎

[18] Corollary: If x is a minimum for (P) where f,g are boundedly differentiable with f real valued, then if $wK(g'(x)) \cap G \subset wP(A,x)$

$$f'(x) \in \overline{wP^+(B,g(x))g'(x) + G^+}.$$

Proof: By [2.26] $f'(x) \in wP^+(A,x)$ while by [2.33] $g'(x) \in w\mathcal{E}$. ∎

[19] Proposition: Suppose that the sufficient conditions (1) (2) (3) of [16] hold and that for some closed convex cone G

$$f'(x) \in \overline{wP^+(B,g(x))\, g'(x) + G^+}.$$

Then x is a minimum for (P).

Proof: This is similar to an argument in Zlobec (1971). ∎

Again it will be shown more generally later. It seems worth emphasising that conditions for $wK^+ + G^+$ to be closed are given in [3.15] and [3.17].

Section Three: Conditions for weak minima

If x is a weak minimum with respect to S it is too much to hope that a direct analogue of [16] or [17] should hold. The next generalization is of some help, however.

[20] Theorem: Suppose x is a weak minimum for (P) and suppose f,g are boundedly differentiable at x with $f'(x)$ completely continuous. Then, if $wK(g'(x)) \cap G \subset wP(A,x)$ for some closed convex cone G, there is $p^+ \in S^+ / \{o\}$ with

$$p^+(f'(x)) \in \overline{wP^+(B,g(x))g'(x) + G^+}.$$

Proof: By proposition [2.26] there is a suitable p^+ with $p^+(f'(x)) \in wP^+(A,x)$. The proof proceeds as in [18].∎

[21] Equally, if the closure conditions are met as in [16] one can actually assert that $p^+(f'(x)) - u^+(g'(x)) \in G^+, p^+ \in S^+ / \{o\}$,

$$u^+ \in wP^+(B,g(x_o)).$$

[22] It was noted in [2.30] that functions f with weak minima can exist for which there is no equivalent real problem. The sufficiency condition of [19] can be rephrased to exclude this possibility.

Theorem: Suppose $x \in A$ and that for some $p^+ \in S^+ / \{o\}$ and some closed convex cone G one has

(1) $p^+(f'(x)) \in \overline{wP^+(B,g(x))g'(x) + G^+}$

(2) $A - x \in G$

(3) A is pseudoconvex with respect to Δ at x.

(4) u^+f is pseudoconvex at x.

Then

$$(P') \quad \min\ (p^+f)(x) \text{ subject to } g(x) \in B, x \in C$$

has a minimum at x.

Proof: This is just [19] for p^+f. ∎

[23] Remarks: (1) [21], [22] in conjunction give that if (P) has a weak minimum at x that (P′) has a minimum at x when the appropriate conditions are met.

(2) [20], [21], [22] could equally well have been phrased for strong tangent cones and compact derivatives without $f'(x)$ in [20] being completely continuous.

[24] The result of [21] can be related to the Fritz John condition [2], [12] as follows.

Theorem: Suppose in [21] that (1) $K(g'(x)) = P(\Delta, x)$ and (2) $P(\Delta, x) \cap P(C,x) = P(A, x)$ then $\exists\ p^+ \in S^+ / \{0\}$, $u^+ \in P^+(B, g(x))$ with

$$p^+(f'(x)) - u^+(g'(x)) \in P^+(C,x)$$

Proof: This is just a special case of [21].

Conditions for (1), (2) to hold were given [2.18] and [2.39].

[25] Remarks: When −B is a closed convex cone then [24](1) is just a regularity condition on g similar to the one imposed on h in [5]. The proof method of [5] could be applied to [24] to obtain a generalization of Craven's (1973, in preparation) announced result which was used in Craven (1973, submitted) for cones B with $B^o = \phi$.

Section Four: Conditions for strong minima

In this section analogous results to the Kuhn-Tucker necessary condition are derived for (P) when x is assumed to be a strong minimum for (P) with respect to S. In this case the multipliers are no longer linear functionals but are continuous linear operators.

[26] Definition: $T \in B[X,Y]$ is said to be positive with respect to two convex cones $K \subset X$, $S \subset Y$ if $T(K) \subset S$.

[27] Definition: The set of all positive mappings of K into S will be called the maximum cone and denoted by K^S. When $S = R^+$ K^S is denoted K^+. When $S = R^{n+}$, K^S is denoted K^{n+}.

In the notation above $g'(x)(P(\Delta, x)) \subset P(B, g(x))$ which was denoted $g'(x) \in \mathcal{E}$ can be rewritten as $g'(x) \in P(\Delta, x)^{P(B,g(x))}$.

[28] Positive mappings and maximum cones have been studied by Ritter (1969, a, b). The next result of Ritter's is central to the generalized Guignard condition.

Theorem: (Ritter (1969,a)) Let X be a normed space and Y a reflexive Banach space. Suppose $S \neq o$ is a closed convex normal cone in Y and that $K_1, K_2, K_3 = K_1 \cap K_2$ are closed convex cones in X. Suppose that either

(1) $K_3^o \neq \phi$ or (2)(i) K_3 has interior relative to $K_3 - K_3$

(ii) K_3 has a point interior to K_2 or K_1

(iii) $\forall y' \in Y'$ $y'(S) = o$ implies $y' = o$ (S is called full).

Then

$$K_1^S + K_2^S = (K_1 \cap K_2)^S. \blacksquare$$

[29] <u>Definition</u>: If $H \subset B[X,Y]$, $S \subset Y$ are convex cones, define $(H)^s$ by

$$(H)^s = \{x \in X \mid T(x) \in S \quad \forall\, T \in H\}$$

[30] <u>Proposition</u>: If $S \neq \{o\}$ is a closed pointed convex cone then $\bar{K} = (K^s)^s$.

<u>Proof</u>: $\Rightarrow$ Let $x \in \bar{K}$, then there is a sequence (net) $\{x_n\} \subset K$ with $x_n \to x$. Then $T(x_n) \subset S$ and since T is continuous and S is closed $T(x) \subset S$ for any T in K^s. Thus $\bar{K} \subset (K^s)^s$. $\Leftarrow$ Suppose $\bar{x} \notin \bar{K}$. Let u^+ be a continuous linear functional which is non negative on $\bar{K}$ but has $u^+(\bar{x}) = -1$. Let $T(y) = su^+(y)$ where $s \in S/\{o\}$. Then

$$T(k) = su^+(k) \in S \qquad \forall k \in \bar{K}$$

and $T \in (\bar{K})^s$. However, $T(\bar{x}) = -s \in -S/\{o\}$ and since S is pointed $T(\bar{x}) \notin S$. Thus $\bar{x} \notin (K^s)^s$ and $(K^s)^s \subset \bar{K}$. ∎

This generalizes the standard result for closed convex cones that $(C^+)^+ = C$.

[31] <u>Proposition</u>: Suppose $H = \{Tg'(x) \in B[X,Y] \mid T \in P(B,g(x))^s\}$; then $H^s \subset K(g'(x))$.

<u>Proof</u>: Let $\bar{y} \in H^s$. Then $(Tg'(x))(\bar{y}) \in S \quad \forall\, T \in P(B,g(x))^s$. Suppose $(g'(x))(\bar{y}) \notin P(B,g(x))$. Then there is some $u^+ \in P^+(B,g(x))$ with $u^+((g'(x))(\bar{y})) = -1$.

Let $s \in S/\{o\}$ and let $T(y) = su^+(y)$. T belongs to $P(B,g(x))^s$ and $T((g'(x))(\bar{y}) \notin S$ which is a contradiction. Thus $(g'(x))(\bar{y}) \in P(B,g(x))$ for any $\bar{y} \in H^s$ which is

$$H^s \subset K(g'(x)). \;∎$$

This result can be improved if one imposes extra conditions on g and $P(B,g(x_o))$.

<u>Theorem</u>: Suppose X,Y,Z are Banach spaces with Y reflexive. Suppose that $S \neq 0 \subset Y$ is a closed convex normal cone and that $g'(x_o)$ has closed range. Suppose also that

(1) $R(g'(x_o)) = Z$ or (2) $R(g'(x_o)) \neq Z$ and

(i) $R(g'(x_o)) \cap P^o(B,g(x_o)) \neq \emptyset$

(ii) S is a full cone.

Then $K^S \subset H$.

<u>Proof</u>: Suppose $T \in K^S$. Then

$$g'(x_o)(h) \in P(B,g(x_o)) \Rightarrow T(h) \in S.$$

The Farkas Lemma [3.6] can be applied since $R(g'(x_o))$ is

closed and

$$T = T_o g'(x_o); \quad T_o \in (P(B,g(x_o)) \cap R(g'(x_o)))^S.$$

The hypotheses guarantee that

$$T_o \in P(B,g(x_o))^S + R(g'(x_o))^S,$$

using [28], so that

$$T_o = T_1 + T_2 \; ; \; T_1 \in P(B,g(x_o))^S$$
$$T_2 \in R(g'(x_o))^{\perp}$$

so that

$$T_o = T_1 g'(x_o) \in H. \blacksquare$$

In particular $(H^S)^S \subset K^S \subset H$ and $H = \overline{H}$. This theorem gives conditions which exclude Zlobec's, (1970) example in which $R(g'(x_o))$ is closed but H is not.

[32] The following is a generalization of Guignard's theorem in the necessary direction.

Theorem: Let X,Z be normed spaces with Y a reflexive Banach space. Let $S \subset Y$ be a closed, convex normal cone. Suppose that f,g are Fréchet differentiable at x and that there is a closed convex cone $G \subset X$ with $G \cap K \subset P(A,x)$ and suppose that

(1) $(G \cap K)^o \neq \phi$ or (2)(i) $G \cap K$ has interior relative to $(G \cap K)-(G \cap K)$

(ii) Some point in $G \cap K$ is interior to G or to K.

(iii) S is full.

Then a necessary condition for x to be a strong minimum with respect to S is

$$f'(x) \in (H^s)^s + G^s.$$

Proof: By [2.24] $f'(x) \in P(A,x)^s$. By the theorem in [28] and (1) or (2)

$$P(A,x)^s \subset (G \cap K)^s = G^s + K^s.$$

By [31] $H^s \subset K$ and it is clear from definition [28] that $K^s \subset (H^s)^s$. Collecting results one has

$$f'(x) \in G^s + (H^s)^s. \blacksquare$$

[33] Corollary: (1) If $(H^s)^s = \bar{H}$, which is true at least for $(Y,S) = (R^n, R^{n+})$, or if $T = g'(x)$ is invertible, then $f'(x) \in \bar{H} + G^s$.

(2) If $(H^s)^s = \bar{H} = H$ then for some $T \in P(B,g(x))^s$

$$f'(x) - T(g'(x)) \in G^s. \blacksquare$$

[34] The result in [33](2) can be proved directly from the Farkas theorem of [3.9] if the appropriate interior conditions are satisfied. These are, perhaps not surprisingly, stronger than

those in [32].

Theorem: (Generalized Guignard) Suppose X,Y,Z are Banach spaces with Y reflexive. Suppose $S \neq o \subset Y$ is a closed convex normal cone and that f,g are Fréchet differentiable with $R(g'(x)) = Z$. Suppose that for some closed convex cone $G \subset X$

$$G \cap K \subset P(A,x)$$

and suppose that there is some $\bar{x} \in G^o$ with $(g'(x))(\bar{x}) \in P(B,g(x))^o$.

Then a necessary condition for x to be a strong local minimum for (P) is

$$f'(x) - T(g'(x)) \in G^s \quad , \quad T \in P(B,g(x))^s.$$

Proof: The conditions of the theorem are sufficient to apply the Farkas theorem of [3.9] to obtain from

$$h \in G, \ (g'(x))(h) \in P(B,g(x)) \ \Rightarrow \ (f'(x))(h) \in S$$

that

$$f'(x) = T(g'(x)) + T_2 \ , T \in P(B,g(x))^s, \ \ T_2 \in G^s. \ \blacksquare$$

[35] Corollary: If G can be chosen to be X, $P(B,g(x))$ need not have interior and S need not be normal.

Proof: In this case one can apply the Farkas Lemma [3.6] to

$$(g'(x))(h) \in P(B,g(x)) \ \Rightarrow \ (f'(x))(h) \in S$$

and proceed as before. $\blacksquare$

The conditions of [34] are clearly stronger than those of [32] since when $\bar{x} \cap G^o$ and $(g'(x))(\bar{x})$ belongs to $P^o(B,g(x))$ $\bar{x}$ must belong to $K^o \cap G^o \subset (G \cap K)^o$ and [32] (1) is satisfied.

[36] Remarks: (1) [32] could have been phrased for weak tangent cones with the appropriate strengthening of the differentiability hypotheses.

(2) In [16] $H^+ + G^+$ closed was required this was guaranteed for $H^s + G^s$ in [32] by the interior conditions of [28] applied to $P(A,x)$.

(3) If S is closed but not normal one can derive

$$f'(x) \in \overline{G^s + (H^s)^s}$$

as can be seen by inspecting Ritter's proof of [28].

[37] Theorem: (Generalized Sufficiency Condition)

Suppose that f,g are boundedly differentiable in convex spaces and that for some point x and some closed convex cone G

(1) $A - x \subset G$

(2) A is weakly pseudoconvex with respect to Δ at x

(3) f is pseudoconvex at x over A with respect to a closed convex cone $S \subset Y$.

Then if

(4) $f'(x) \in \overline{wP(B,g(x))^s g'(x) + G^s}$

x is a strong minimum for (P) with respect to S.

Proof: Using (4) there are nets $\{G_n\} \subset G^s$ and $\{T_n\} \subset wP(B,g(x))^s$ with

(5) $(f'(x))(y - x) = \lim \left[T_n \ (g'(x))(y - x) \ + G_n(y - x) \right]$

Let $y \in A$, then by (1) $y - x \in G$ and thus (6) $G_n(y - x) \in S$.

Now, by (2), $A - x \subset wP(\Delta, x)$. Since g is boundedly differentiable $g'(x) \in w\mathcal{E}$ ([2.33]) and $(g'(x))(y - x) \in wP(B,g(x))$ so that

(7) $T_n \ (g'(x))(y - x) \in S$. Substituting (6) and (7) in (5)

(8) $(f'(x))(y - x) = \lim S_n \in \overline{S} = S$.

Since f is assumed pseudoconvex with respect to S at x over A (8) yields

$$f(y) - f(x) \in S \quad \forall y \in A$$

and x is a strong minimum for (P). ∎

This theorem clearly remains true if all weak pseudotangent cone relations are replaced by strong ones. It is then only necessary for g to be compactly differentiable as one uses [2.32] instead of [2.33].

[38] When $Y = R^n$ a necessary condition can be derived for (P) directly from the Guignard type result proved for real valued maps in [16]. This is phrased for strong tangent cones.

Theorem: Suppose in (P) that $Y = R^n$ and that x is a strong minimum with respect to a pointed cone $S \subset R^n$. Suppose further that g satisfies the Guignard constraint condition for some convex cone G. Then a necessary condition for x to be a strong minimum is

$$f'(x) - Mg'(x) \quad \in \quad K^{-1}(G^{n+})$$

where $M = K^{-1}T$ with K an invertible nxn matrix with rows in S^+ and $T = (u_1^+, \ldots, u_n^+)^T$, $u_i^+ \in P^+(B, g(x))$.

Proof: Since x is a minimum for f over A

$$f'(x) \in P(A,x)^s$$

Since S is pointed in R^n, S^+ has interior. Choose $y_1^+, \ldots, y_n^+$ linearly independent in S^+; then

$$y_i^+ f'(x) \in P^+(A,x) \qquad i = 1, \ldots, n.$$

Applying the necessary condition of [16] to each $y_i^+ f'(x)$ one derives $u_i^+ \in P^+(B, g(x))$ with

(1) $y_i^+(f'(x)) - u_i^+(g'(x)) \in G^+ \qquad i = 1, \ldots, n$

which can be written as

$$K(f'(x)) - T(g'(x)) \in G^{n+}$$

or, since K^{-1} exists, as

$$f'(x) - M(g'(x)) \in K^{-1}(G^{n+}). \blacksquare$$

[39] Corollary: If $S = R^{n+}$, K can be chosen as I and one derives

$$f'(x) - Tg'(x) \in G^{n+} \ , \quad T \in P(B,g(x))^{n+}_{\cdot} \ \|$$

In particular one obtains for (M_1)

$$(M_1) \ \min f(x) \text{ subject to } g_1(x) \geqslant o.. g_m(x) \geqslant o, \ x \in R^{k+}$$

where $f: R^k \to R^m$ and $g_i: R^k \to R$ are supposed continuously differentiable. Minimization is with respect to $S = R^{n+}$. Denote $g = (g_1, \ldots, g_m)^T$.

[40] Theorem: A necessary condition for x to be a strong minimum with respect to R^{n+} is that when x satisfies the Kuhn-Tucker constraint condition of [15] one has

$$f'(x) - T(g'(x)) \in P(R^{k+},x)^{n+}; \ T \in P(R^{m+},g(x))^{n+}.$$

When f is pseudoconvex with respect to R^{n+} at x this is also sufficient if $A - x \subset P(\triangle, x)$.

Proof: The Kuhn-Tucker constraint qualification gives $G = P(R^{k+},x)$. The closure conditions of [16] are met since all the cones involved are polyhedral. Note that since the spaces are finite dimensional weak and strong tangent cones coincide. Thus this necessary condition is just a special case of [39]. Sufficiency follows from [37] since conditions (1) and (2) hold because of the choice of G and because R^{k+} is convex. ∎

The last result of this section can be used to recast Fritz John type results for weak minima in operator form.

[41] Theorem: Suppose G is a closed convex cone in Y and that $p^+ \in S^+, u^+ \in P^+(B,g(x))$ exist with

$$p^+(f'(x)) - u^+(g'(x)) \in G^+.$$

Then there exist $T_1 \in S^s$, $T_2 \in P(B,g(x))^s$, $T_3 \in G^s$ with

$$T_1(f'(x)) - T_2(g'(x)) = T_3 \qquad (T_1,T_2) \neq o.$$

Proof: Define $T_1(y) = sp^+(y)$, $T_2(y) = su^+(y)$, $T_3 = T_1(f'(x)) - T_2(g'(x))$ where $s \in S/\{o\}$. |

Section Five: One Sided Derivatives

One drawback to Zlobec's notion of asymptotic consistency (see [17]) is that it is necessary for the functions involved to be linear and continuous. There are many functions which are not differentiable but which have an associated non-linear variation for which optimization results can be framed.

[42] Definition: Let $f\colon X \to Y$ with X,Y vector spaces. The one sided (Gateaux) derivative of f at x denoted $d^+f(x;)$ is defined by

$$d^+f(x;h) = \lim_{t \to o^+} \frac{f(x + th) - f(x)}{t}.$$

[43] Proposition: Suppose $f\colon X \to Y$ is convex with respect to a closed convex cone S and has a one sided derivative at x. Then

$$d^+f(x;h) = \text{“}\inf_{t > o}\text{”} \frac{f(x + th) - f(x)}{t}$$

and $d^+f(x;h)$ is convex and positively homogeneous as a function of h.

Proof: The convex inequality can be written as

$$(1) \quad \frac{f(x + th) - f(x)}{t} \geqslant_s \frac{f(x + rh) - f(x)}{r} \qquad t \geqslant r,$$

which is the force of the first conclusion. Moreoever, denoting $t^{-1}[f(x + th) - f(x)]$ by $g_t(h)$ it is clear that g_t is convex as a function of h. Thus for $o \leqslant r \leqslant 1$ and $t > o$

(2) $rg_t(h) + (1 - r)g_t(k) \geqslant_S g_t(rh + (1 - r)k)$.

Taking limits one has the convexity of $d^+f(x;h)$ since S is assumed closed. ∎

Note that this result includes a type of subgradient inequality.

$$f(x + h) - f(x) \geqslant_S d^+f(x;h).$$

[44] It is also apparent that the following proposition holds.

Proposition: If $f: X \to Y$ is quasiconvex with respect to a closed S and has a one sided derivative then

$$d^+f(x;h) \in -S \quad \text{when} \quad f(x + h) - f(x) \in -S. \ ∎$$

[45] As a partial generalization of [43] one has:

Proposition: Suppose $g: X \to Y$ is (P) strictly quasiconvex or quasiconvex with respect to S. Then when $d^+g(x,h)$ exists it satisfies

$$d^+g(x;rh + (1 - r)k) \leqslant d^+g(x;k)$$

whenever $d^+g(x;h) < d^+g(x;k)$ and $0 \leqslant r \leqslant 1$.

Proof: If $d^+g(x;h) - d^+g(x;k) \in -S^o$ there is some t_o such that for $t < t_o$

$$\frac{g(x + th) - g(x)}{t} < \frac{g(x + tk) - g(x)}{t}$$

which gives $g(x + th) < g(x + tk) \quad t < t_o$.

Using either of the convexity hypotheses

$$g(x + rth + (1-r)tk) \leqslant g(x + tk) \quad 0 \leqslant r \leqslant 1$$

which in turn gives

$$d^+g(x;rh + (1 - r)k) \leqslant d^+g(x;k). \ ∎$$

The simplest example of a function with a convex one sided derivative is given by $f(r) = |r|$ for which $d^+f(o;h) = |h|$.

The next results give some tangent cone relationships for one sided derivatives. For simplicity the spaces are taken to be convex throughout.

[46] Proposition: Let $f: X \to Y$ have a one sided derivative at x which is uppersemicontinuous with respect to S. Suppose x is a strong (local) minimum for f over $A \subset X$. Then when A is convex

$$d^+f(x,h) \in S \quad \forall h \in P(A,x).$$

Proof: Since A is convex and x is a minimum over A

$$f(x + t(y - x)) \_ f(x) \quad \forall y \in A$$

and

$$d^+f(x;y- x) \geqslant o \qquad \forall y \in A.$$

Since $d^+f(x;h)$ is positively homogeneous in h

$$(1) \quad d^+f(x;t(y - x)) \geqslant o \quad t \geqslant o, \forall y \in A.$$

Let $h \in P(A, x)$. Then there is a net $h_n \to h$ with $h_n = t_n(x_n - x)$, $t_n \geqslant o, x_n \in A$. Using (1) and the uppersemicontinuity of $d^+f(x;h)$, $o \leqslant \lim d^+f(x;h_n) \leqslant d^+f(x,h)$ since S is closed. Thus one is done. ▌

[47] Proposition: Suppose that in the hypotheses of [46] that rather than A being convex f is supposed to be quasiconcave with respect to S. Then

$$d^+f(x;h) \in S \quad \forall h \in T(A,x).$$

Proof: Let $h_n = \lambda_n(x_n - x) \to h$, $x_n \to x$, $x_n \in A$, $\lambda_n \geqslant o$. Then $f(x + t(x_n - x)) \geqslant f(x)$ for $o \leqslant t \leqslant 1$ since $f(x_n) \geqslant f(x)$. Proceeding as before one derives

$$d^+f(x\,;\,\lambda_n(x_n - x)) = d^+f(x;h_n) \in S\,.$$

Since f is uppersemicontinuous and S is closed $d^+f(x;h) \in S$. ∎

[48] Corollary: Suppose in [47] that $d^+f(x;h)$ is concave in h with respect to S. Then

$$d^+f(x;h) \in S \quad \forall h \in P(A,x).$$

Proof: Since $d^+f(x;h)$ is concave and [47] holds

$$d^+f(x;\bar{h}) \in S \qquad \forall\, \bar{h} \in [T(A,x)]\,.$$

Since $d^+f(x;h)$ is uppersemicontinuous in h

$$d^+f(x;\bar{h}) \in S \qquad \forall \bar{h} \in \overline{[T(A,x)]} = P(A,x). \;∎$$

[49] Proposition: Suppose $d^+g(x;h)$ exists and is convex in h with respect to a closed, convex cone B with interior and

$$(1) \quad \exists\; \bar{h} \in X \text{ with } d^+g(x;\bar{h}) + g(x) < 0\,.$$

Then

$$\{h \mid d^+g(x;h) \in P(-B,g(x))\} \subset P(\Delta,x) \quad (\Delta = g^{-1}(-B)).$$

Proof: Let $h_t = th + (1-t)\bar{h}$ $\;0 < t < 1$. Suppose $d^+g(x,h) \in P(-B,g(x))$. There is $\{y_n\} \subset -B$ with $y_n \to g(x)$ and $\lambda_n(y_n - g(x)) \to d^+g(x;h)$. Since $d^+g(x;\;)$ is convex

$$d^+g(x;h_t) + (1-t)g(x) < td^+g(x;h) + (1-t)d^+g(x;\bar{h}) +$$
$$(1-t)g(x) < td^+g(x;h)\,.$$

Thus for $n > n_o$

$$d^+g(x;h_t) + (1-t)g(x) < \lambda_n(y_n - g(x)) \leqslant -\lambda_n g(x)\,.$$

Using the definition of the derivative for $r \leqslant r_o$

$$\frac{g(x + rh_t) - g(x)}{r} < (-\lambda_n + t - 1)g(x)\,.$$

Rearranging one has

$$g(x + rh_t) < g(x) + r(t - (1 + \lambda_n))g(x)$$

Fixing n at n_o, if r is small enough $r(t - (1 + \lambda_n)) \geqslant -\frac{1}{2}$ and

$$g(x + rh_t) < \tfrac{1}{2}g(x) \leqslant 0.$$

This means that $x + rh_t \in g^{-1}(-B) = \Delta$ and $h_t \in T(\Delta, x)$.

Letting $t \to 1$ $h_t \to h \in T(\Delta, x) \subset P(\Delta, x)$. ▮

[50] Corollary: Suppose g is actually compactly differentiable in [49] then $K = P(\Delta, x)$.

Proof: In this case $(g'(x))(h) = d^+g(x;h)$ so that [49] shows that

$$g'(x)^{-1}(P(-B,g(x))) \subset P(\Delta, x).$$

The reverse containment was given in [2.32] so that

$$g'(x)^{-1}(P(-B,g(x))) = P(g^{-1}(-B), x).$$ ▮

With these preliminaries one is in a position to prove a necessary condition which generalizes a problem of Luenberger's (1970).

[51] Theorem: Let $f: X \to R$ $g: X \to Y$, $B \subset Y$, $C \subset X$ and let (P_3) be min $f(x)$ s.t. $g(x) \in -B, x \in C$. Suppose one has

(1) $P(-B,g(x))^{\circ} \neq \phi$,

(2) $d^+f(x;h)$ is convex in h,

(3) $d^+g(x;h)$ is convex in h with respect $-P(-B,g(x))$,

(4) $h \in P(\Delta, x)$ when $d^+g(x;h) \in P(-B,g(x))$,

(5) A is convex or f is quasiconcave,

(6) $d^+f(x;h)$ is uppersemicontinuous in h.

Then if G is a closed convex cone with $G \cap P(\Delta, x) \subset P(A, x)$ and there is some $\bar{h} \in G$ with (7) $d^+g(x;\bar{h}) \in P(-B,g(x))^{\circ}$ a necessary condition for x to be a minimum for (P_3) is

$$d^+g(x;h) - u^+(d^+g(x;h)) \geqslant 0 \quad \forall h \in G$$

for some $u^+ \in P^+(-B,g(x))$.

Proof: Let $E = \{(r,z) \mid d^+f(x;h) \leqslant r,\ d^+g(x;h) - z \in P(-B,g(x)),\ h \in G\}$

and $F = \{(r,z) \mid r \leqslant o,\ z \in P(-B,g(x))\}$

Suppose $E \cap F^{o} \neq \phi$. Then there is some $h \in G$ with

$$d^{+}f(x;h) < o,\ d^{+}g(x;h) \in P(-B,g(x))^{o}.$$

By (4) $h \in P(\Delta,x) \cap G \subset P(A,x)$. Then using (5), (6) and [45] or [46]

$$d^{+}f(x,h) \geqslant o$$

which is impossible. Thus $E \cap F^{o} = \phi$. Moreo-ver, E is convex since G is convex and (using (2), (3)) $d^{+}f(x;h)$ and $d^{+}g(x;h)$ are convex. There is, by the separation theorem for convex sets, a linear functional $z^{+} = (\bar{r},u^{+}) = o$ with $\bar{r} \geqslant o$, $-u^{+} \in P^{+}(-B,g(x))$ such that

$$\bar{r}(d^{+}f(x;h)) + u^{+}(d^{+}g(x;h)) \geqslant o \quad \forall h \in G.$$

If $\bar{r} = o$ then $u^{+}(d^{+}g(x;h)) \geqslant o \quad \forall h \in G$ which contradicts (7).
$\bar{r}$ can be taken to be 1 and the theorem is established.|

[52] <u>Corollary</u>: With (P_3) as in [51] suppose B is a closed convex cone with interior, that $C = X$, and that

(1)′ $d^{+}f(x;h)$ is convex and uppersemicontinuous in h,

(2)′ Either f is quasiconcave or Δ is convex,

(3)′ $d^{+}g(x;h)$ is convex with respect to B.

Then if h exists with $d^{+}g(x;\bar{h}) + g(x) \in -B^{o}$ and x is minimal for (P_3) one has

$$u^{+}(g(x)) = o,\ d^{+}f(x;h) + u^{+}d^{+}g(x;h) \geqslant o \quad \forall h \in G \text{ for some } u^{+} \in B^{+}.$$

<u>Proof</u>: Since $B \subset B + g(x) \subset -P(-B,g(x))$, $P(-B,g(x))^{o} \neq \phi$.
(3) is implied by (3)′ since $B \subset -P(B,g(x))$.
(3)′ and [49] imply (4) while $G \cap P(\Delta,x) \subset P(A,x)$ is satisfied by $G = X$ since $C = X$. Finally $d^{+}g(x;\bar{h}) + g(x) < o$ is sufficient to exclude $\bar{r} = o$ in the proof of [51]. [51] thus gives the

desired result because for B a convex cone

$$u^+ \in -\bar{P}^+(-B, g(x)) \text{ implies } u^+ \in B^+ \text{ and } u^+(g(x)) = 0. \blacksquare$$

[53] Remarks: (1) If $X = R$; $B = \{1/n\}_1^\infty \cup \{0\}$ and $g(x) = 0$ is an example in which $-B^0 = \emptyset$ but $P(-B,0)^0 \neq \emptyset$ so that condition (1) of [51] is weaker than $B^0 \neq \emptyset$.

(2) If f is differentiable (2′) of [52] or (5) of [51] can be dropped.

(3) Luenberger (1969) states [52] without (2′) and $d^+f(x;h)$ uppersemicontinuous in h. This can be established in a direct argument but not from [51].

One also has a sufficiency condition for one sided derivatives similar to [37].

[54] Definition: $f: X \to Y$ is pseudoconvex with respect to S at x over A if

$$d^+f(x;y-x) \geq_S 0 \qquad \text{implies } f(y) \geq_S f(x) \quad \forall y \in A.$$

[55] Theorem: Suppose there are closed convex cones $G \subset X$; $B \subset Z$ with:

(1) $A - x \subset G$.

(2) f is pseudoconvex over A at x with respect to S.

(3) There is some $u^+ \in B^+$ with $u^+(g(x)) = 0$ and

$$d^+f(x;h) + u^+(d^+g(x;h)) \geq 0 \qquad \forall h \in G.$$

(4) u^+g is quasiconvex.

Then x is a strong minimum with respect to S for

$$\min f(y) \quad \text{subject to} \quad g(y) \leq_B 0 \quad y \in C.$$

Proof: Let $y \in A$. Using (1) and (3)

(4) $d^+f(x;y-x) + u^+(d^+g(x;y-x)) \geq 0$.

If $g(y) \leq 0$ then $u^+(g(y)) \leq u^+(g(x)) = 0$. (4) and [44] then yield

(5) $u^+(d^+g(x,y - x)) = d^+u^+(g(x;y - x) \leqslant o.$

Combining (4) and (5) produces $d^+f(x;y - x) \geqslant o \quad \forall y \in A$. Since f is pseudoconvex x is a strong minimum. ▮

Section Six: Convex Optimization Theorems

Let (P_4) denote

(P_4) min $f(x)$ subject to $g(x) \leqslant_B o$, $h(x) = o$, $x \in C$ where $f: X \to R$, $g: X \to Z$, $h: X \to W$ and B is a closed convex cone with interior in Z. C is a convex set in X a reflexive Banach space. The functions will all be supposed to satisfy various convexity assumptions. The fundamental utility of subgradients arises from the following theorem which generalizes some results of Rockafellar's.

[56] Theorem: Suppose $\text{domf}^o \neq o$ and f is continuous and convex on $C = \text{domf}$. Suppose g is convex with respect to B, g is continuous on C, h is affine, continuous and open. Suppose that

(1) $\exists\, x_1 \in C$ with $g(x_1) < o \quad h(x_1) = o$,

(2) $\exists\, x_2 \in C^o$ with $h(x_2) = o$.

Then a necessary condition for $\bar{x}$ to be optimal for (P_4) is;

$$o \in \partial f(\bar{x}) + \partial(u^+g)(\bar{x}) + w^+(\partial h(\bar{x})); u^+(g(\bar{x}) = o$$

for some $u^+ \in B^+$, $w^+ \in W'$.

Proof: The hypotheses guarantee an equivalent unconstrained problem of minimizing

$$\varphi(x) = f(x) + u^+(g(x)) + w^+(h(x))$$

for some $w^+ \in W'$, $u^+ \in B^+$ with $u^+(g(\bar{x})) = o$. This is proved in [6.17].

By the definition of $\varphi(x)$ one has

$$o \in \partial\varphi(\bar{x}) = \partial(f(x) + u^+(g(\bar{x})) + w^+(h(\bar{x}))).$$

f, u^+g and w^+h are real valued convex mappings which are (lower semi) continuous on C^o. By a result of Rockafellar's (1966) or from the general theory of maximal monotone mappings

(3) $o \in \partial f(\bar{x}) + \partial u^+(g(\bar{x})) + \partial w^+(h(\bar{x})) = \partial\varphi(\bar{x})$.

Moreoever, h is affine and continuous and thus differentiable. So

(4) $\partial(w^+h(\bar{x})) = (w^+h)'(\bar{x}) = w^+(h'(\bar{x})) = w^+(\partial h(\bar{x}))$.

Combining (3) and (4) gives the result.∎

[57] Corollary: (1) When X, Z, W are finite dimensional spaces the continuity results are met automatically if $C^o \subset (\text{dom}f)^o \cap (\text{dom}g)^o$

(2) When $B = R^{n+}$ and $Z = R^n$ $\partial u^+(g(\bar{x}))$ can be replaced by $u^+(\partial_B g(\bar{x}))$.

Proof: (1) is just [1.53] while (2) follows from the definition of $\partial_B g$.∎

[58] Corollary: If B is pointed and $g'(\bar{x})$ exists $u^+(\partial_B g)$ can be used to replace $\partial(u^+g)$.

Proof: In this case (4) of [56] holds for u^+ and g.∎

Corollary [57] (2) is the basic theorem of Rockafellar (1970). There appears to be a general problem in the extension of [56] since it seems that $\partial(u^+g)$ and $u^+(\partial g)$ need not in general be equal.

The tangent cone relationships proved in chapter two and in [4.13], [4.15] can be used to derive some more general results about (P_4).

[59] Theorem: Suppose f is either differentiable or uppersemicontinuous and concave Suppose that g is concave with respect to B and there is some $\bar{Z} \in \bar{\partial} g(\bar{x})$ with

(1) $\bar{Z}(h) \in P(-B, g(\bar{x})) \Rightarrow h \in P(\Delta, \bar{x})$.

(2) For some closed convex cone G with $G \cap P(\Delta, \bar{x}) \subset P(A, \bar{x})$

$\bar{Z}(\bar{h}) \in P^{o}(-B, g(\bar{x}))$ for some $\bar{h} \in G$.

A necessary condition for $\bar{x}$ to be a minimum for (P_4) is:

$u^{+}\bar{Z} \in \bar{\partial} f(\bar{x})$ for some $u^{+} \in P^{+}(-B, g(x))$.

Proof: By [2.24] $f'(\bar{x}) \in P^{+}(A, \bar{x})$. By [4.13] $\bar{\partial} f(\bar{x}) \in P^{+}(A, \bar{x})$. Let $\bar{y} \in \bar{\partial} f(\bar{x})$ or let $\bar{y} = f'(\bar{x})$; one can then apply the argument of [51] to $\bar{y}$, $\bar{Z}$ in place of $d^{+}f$, $d^{+}g$ to obtain

$\bar{y} + u^{+}\bar{Z} = 0 \qquad u^{+} \in -P^{+}(-B, g(\bar{x}))$. ∎

[60] Corollary: Suppose $B^{o} \neq \phi$, C = X and for some $\bar{Z} \in \bar{\partial} g(\bar{x})$

$\bar{Z}(\bar{h}) + g(\bar{x}) < 0$.

Then the necessary condition in [59] holds.

Proof: This follows from [59] and [4.15] much as [52] follows from [51] and [50]. ∎

[59] and [60] give results which are essentially backwards as they give necessary conditions for concave functions' minima and not maxima. The next theorem rectifies this for the constraint function.

[61] Theorem: Suppose in (P_4) f is uppersemicontinuous and concave, $h \equiv 0$, g is convex with respect to B and has a subgradient $\bar{Z} \in \partial g(\bar{x})$. Then a necessary condition for $\bar{x}$ to be a minimum in (P_4) is

$$\bar{\partial} f(\bar{x}) \subset \overline{P^{+}(-B,g(x))\bar{Z} + G^{+}}$$

whre G is any closed convex cone with $K \cap G \subset P(A,\bar{x})$.

(K denotes $\{h \mid \bar{Z}(h) \in P(-B,g(\bar{x}))\}$).

Proof: Let $\bar{y} \in \bar{\partial} f(\bar{x})$ then by [4.13] $\bar{y} \in P^{+}(A,\bar{x})$.

Since $\bar{Z} \in \partial g(\bar{x})$ [4.14] guarantees that

$$\bar{Z}(P(g^{-1}(-B),\bar{x})) \subset P(-B,g(\bar{x}))$$

One can apply Zlobec's result (which is [17] with strong cones) to deduce that

$$\bar{y} \in \overline{P^{+}(-B,g(\bar{x}))\bar{Z} + G^{+}}$$

Note that the assumptions of f guarantee that $\bar{\partial} f(\bar{x}) \neq \phi$. ∎

Chapter Six

MULTIPLIER THEOREMS AND MINIMAX THEOREMS

Multiplier Theorems and Minimax Theorems

This chapter is concerned with multiplier theorems for convex and quasiconvex programmes in locally convex spaces. The first section is concerned with various generalizations of a theorem proved by Luenberger (1968) concerning the existence of simpler, but constrained equivalent problems for

(P_5) weakmin$_S$ $f(x)$ subject to $g(x) \in -B$, $x \in C$.

where $f:X \to Y$ is strongly quasiconvex w.r.t.S, a closed convex with interior, $g:X \to Z$ is convex w.r.t.B, a closed convex cone with interior, and $C \subset X$ is a convex set .

[1] Theorem: Suppose (1) f is upper semicontinuous on lines w.r.t.S

(2) $\exists\ x_1 \in C$ with $g(x_1) \in -B^0$.

If x_0 is a weak minimum for (P_5) there is some $u^+ \in B^+/\{0\}$ with (P_5) equivalent to (P_5')

(P_5') weakmin$_S$ $f(x)$ subject to $u^+(g(x)) \leqslant 0, x \in C$.

Proof: Let $E = \{y \mid f(x) < u_0 \ ; \ g(x)-y \in -B,\ x \in C\}$

where $u_0 = f(x_0)$.

(i) E is convex.

Let $y_1, y_2 \in E$. Then there are $x_1, x_2 \in C$ with

$$f(x_1) < u_0;\ f(x_2) < u_0 \qquad g(x_1) \leqslant y_1;\ g(x_2) \leqslant y_2.$$

Since C and g are convex $x_\lambda = \lambda x_1 + (1-\lambda)\, x_2 \in C$ and

$$g(x_\lambda) \leqslant \lambda g(x_1) + (1-\lambda)\, g(x_2) \leqslant \lambda\, y_1 + (1-\lambda)\, y_2.$$

Since f is strongly quasiconvex w.r.t. S, proposition [18] of chapter one gives $f(x_\lambda) < u_0$ since $f(x_1) < u_0, f(x_2) < u_0$.

Thus $\lambda y_1 + (1-\lambda)\, y_2 \in E$ and E is convex.

(ii) By construction $E \cap -B = \emptyset$. Since $B^0 \neq \emptyset$ the Hahn-Banach Theorem is applicable and there is some non zero $u^+ \in B^+$ with

$u^+(y) \geqslant 0 \quad \forall\, y \in E.$

(iii) Suppose that $y_2 \in E$ with $u^+(y_2) = 0$. By the second hypothesis $x_1 \in C$ with $g(x_1) \in -B^o$. Since $y_2 \in E$ there is $x_2 \in C$ with $g(x_2) \leqslant y_2$. Thus for $0 < \lambda < 1$

$$u^+(g(\lambda x_1 + (1-\lambda)x_2)) \leqslant \lambda u^+(g(x_1)) + (1-\lambda)\, u^+(g(x_2)) < 0.$$

Since $\lambda x_1 + (1-\lambda)x_2 \in C$, $g(\lambda x_1 + (1-\lambda)x_2) \notin E$ and

$f(\lambda x_1 + (1-\lambda)x_2) \notin -S^o + u_o$. Suppose that

$f(x_2) \in -S^o + u_o$. By the first hypothesis

$$\lim_{\lambda \to 0} f(\lambda x_1 + (1-\lambda)x_2) - f(x_2) \in -S.$$

Thus one would have

$$\lim_{\lambda \to 0} f(\lambda x_1 + (1-\lambda)\, x_2) \in -S+(-S^o) + u_o \subset -S^o + u_o.$$

Since $-S^o + u_o$ is an open set this would imply that

$$f(\lambda x_1 + (1-\lambda)x_2) \in -S^o + u_o \quad \text{for} \quad \lambda < \lambda_o$$

which is impossible. Thus $f(x_2) \notin - S^o + u_o$ and since x_2 was arbitrary $y_2 \notin E$. Hence, one has $y \notin E$ when $u^+(y) \leqslant 0$. From this one derives that when $x \in C$ and $u^+(g(x)) \leqslant 0$ $f(x) \notin -S^o + f(x_o)$.

By the definition of weak minimization and the feasibility of x_o one has that x_o is a weak minimum for (P_5').∎

[2] As Luenberger notes one need not have complementary slackness, $u^+(g(x_o)) = 0$, in (P_5'). This is shown by

$$f(r_1,r_2) = \begin{cases} 0 & r_1+r_2 > 0 \\ +1 & r_1+r_2 \leqslant 0 \end{cases} \quad \text{and} \quad g(r_1,r_2) = (r_1,r_2)$$

for which (P_5) has a minimum of 1 at $(-1;-1)$ among other points. Setting $u^+ = (1;1) \in R^{2+}$ one sees that u^+ satisfies (P_5') but $u^+(g(x_o)) = -2$.

There are various results about the slackness of $u^+(g(x_o))$. The next proposition generalises one of Luenberger's.

[3] Proposition: With everything as in [1] either

(1) $\exists$ $x_2 \in C$ with $g(x_2) \in -B^o$, $f(x_2) - f(x_o) \in -\text{bdS}$

or

(2) $u^+(g(x_o)) = 0$ for some nonzero u^+ satisfying (P_5').

Proof: Let $A = \{y \mid \exists x \in C, y - g(x) \in B, \; u_o - f(x) \in S\}$

(i) Clearly A is convex.

(ii) If $A \cap -B^o \neq \emptyset$ there is some $\overline{x} \in C$ with $f(\overline{x}) \leqslant u_o$, $g(\overline{x}) < 0$ and if (1) does not hold $f(\overline{x}) - f(x_o)$ cannot belong to $-S / -S^o = \text{bdS}$. This means that $f(\overline{x}) \in -S^o + u_o$ and $g(\overline{x}) \leqslant 0$, $\overline{x} \in C$ which contradicts the definition of x_o. Thus $A \cap B^o = \emptyset$. There is, therefore, a separating hyperplane with $u^+ \in B^+/\{0\}$ and $u^+(A) \geqslant 0$. It is simple to prove that this implies that $u^+(E) \geqslant 0$ and to proceed as in theorem [1] to show u^+ satisfies (P_5').

Moreoever, $u^+ \in B^+$ so that $u^+(g(x_o)) \leqslant 0$. Since $g(x_o) \in A$ $u^+(g(x_o)) \geqslant 0$ and must be zero.❙

[4] Proposition: Suppose that f satisfies the following condition:

(1) $x \neq y$ and $f(x) \leqslant_S f(y)$ implies $f(\lambda x + (1-\lambda)y) <_S f(y)$ $\quad 0 < \lambda < 1$:

then if the conditions of theorem [1] hold either $g(x_o) \in -B^o$ or $u^+(g(x_o)) = 0$.

Proof: By [3] , if $u^+(g(x_o)) \neq 0$ one has some $x_1 \in C$ with $g(x_1) < 0$ and $f(x_1) - f(x_o) \in -\text{bdS}$. This means $f(x_1) \leqslant f(x_o)$. If $x_1 \neq x_o$ one has by property (1) that $f(\lambda x_1 + (1-\lambda)x_o) < f(x_o)$ $\quad 0 < \lambda < 1$. Since $\lambda x_1 + (1-\lambda)x_o \in C$ and $g(\lambda x_1 + (1-\lambda)x_o) \leqslant 0$ this contradicts the

the minimality of x_o. ∎

Note that if f is strictly strongly quasiconvex w.r.t. S, f is strongly quasiconvex and satisfies (1).

Continuity conditions can be imposed on g to insure that for some u^+ (possibly 0) one does have complementary slackness.

[5] Proposition: Suppose in [4] that g is fully uppersemicontinuous with respect to B and that $x_o \in C^o$ then either x_o is a global minimum and $u^+(g(x_o)) = 0$ for $u^+ = 0$ or $u^+(g(x_o)) = 0$ and u^+ can be taken nonzero.

Proof: By [4] $u^+(g(x_o)) \neq 0 \Rightarrow g(x_o) \in -B^o$.
Since x_o is assumed to lie in C^o and since g is fully upper semi-continuous there is a neighbourhood N of x_o in C with $g(N) \subset -S$. Thus

$$f(x) - f(x_o) \not\in -S^o \qquad \forall x \in N.$$

Since property (1) of [4] is stronger that (P) strict quasiconvexity, proposition [80] of chapter one implies that x_o is a weak global minimum. In this case (P_5) is equivalent to (P_5') with $u^+ = 0$. ∎

These results exclude Luenberger's example which satisfies all the conditions except the strict quasiconvexity of f.

Equality Constraints

[6] Definition: $h:X \longrightarrow W$ is subaffine if $A(a) = \{x \mid h(x)=a\}$ is convex $\forall\, a \in W$.

[7] Proposition: h is subaffine if h is any of the following:

(1) affine.

(2) maximal monotone from X to X'.

(3) quasi-convex and-concave with respect to a pointed cone S.

<u>Proof</u>: (1) is immediate. (2) is a standard result of monotone operator theory. (3) Suppose h is quasi-convex and-concave. If $a \notin R(h)$ $A(a) = \emptyset$. If $a \in R(h)$ then $a = h(b)$ and $\{x \mid h(x) = h(b)\} = \{x \mid h(x) \leq h(b)\} \cap \{x \mid h(x) \geq h(b)\}$ because S is pointed.

Since the last two sets are convex $A(a)$ is convex. ▮

[8] <u>Proposition</u>: Suppose that $h\colon X \to W$ is subaffine and that in [1] (P_5) is replaced by

$$(P_6) \text{ weakmin}_S\ f(x) \text{ s.t. } g(x) \leq 0,\ h(x) = 0, x \in C.$$

Then there is some $u^+ \in B^+/\{0\}$ with (P_6) equivalent to

$$(P_6{}'') \text{ weakmin}_S\ f(x) \text{ s.t. } u^+(g(x)) \leq 0\ ,\ h(x) = 0,\ x \quad C.$$

<u>Proof</u>: Let $C'' = C \cap \{x \mid h(x) = 0\}$ which is convex and apply [1] to (P_5) with C replaced by C''. ▮

The next result extends [1] to include a finite dimensional affine constraint.

[9] <u>Theorem</u>: Suppose in [8] that $h\colon X \to R^n$ is actually affine and that the following hold.

(1) $\exists\, x_1 \in C$ s.t. $h(x_1) = 0$, $g(x_1) < 0$.

(2) $\exists\, x_2 \in C^o$ with $f(x_2) < u_o$ $h(x_2) = 0$ and with h open at x_2.

(3) f is actually fully upper semicontinuous with respect to S.

Then there is some $u^+ \in B^+/\{0\}$ and $z^+ \in R^n$ with (P_6) equivalent to $(P_6{}')$.

$$(P_6{}') \text{ weakmin}_S\ f(x) \text{ s.t. } u^+(g(x)) + z^+(h(x)) \leq 0\ ,\ x \in C.$$

<u>Proof</u>: By [8] (P_6) is equivalent to (P_6'').

Let $E_1 = \{(r,z) | f(x) < u_o, \ u^+(g(x)) \leqslant r, \ h(x) = z, \ x \in C\}$ where $u^+ \in B^+/\{o\}$ is as in [8].

As in [1], since h is affine, E_1 is convex.

Set $F_1 = \{(r,o) | \ r \leqslant 0, \ \ o \in R^n\}$. By [8] $E_1 \cap F_1 = \emptyset$ and one can apply Hahn-Banach since the sets are finite dimensional. This produces $\bar{r} \geqslant 0$, $z^+ \in R^n$, $(r,z^+) \neq 0$

$\bar{r}\, u^+(g(x)) + z^+(h(x)) \geqslant 0$ if $x \in C$, $f(x) < u$.

Suppose $\bar{r} = 0$ then

$z^+(h(x)) \geqslant 0$ when $x \in C$, $f(x) < u_o$.

By (2) there is an $x_2 \in C^o$ with $f(x_2) < u_o$ and $h(x_2) = 0$ and with $h(N)$ a neighbourhood in R^n for some N with $x_2 + N \subset C^o$. By (3) there is a neighbourhood N_1 of x_2 with $f(N_1) < u_o$. Combining these facts one sees that $z^+(U) \geqslant o$ for $U = h(N)$ a neighbourhood of 0. This implies that $z^+ = 0$ which is impossible.

Without loss of generality suppose r = 1. One sees that $u^+(g(x)) + z^+(h(x)) \leqslant 0$ when $x \in C$, $f(x) < u_o$. One can now proceed as in [1] to derive the promised equivalence. This uses $u^+ \neq 0$ and (1) in place of $g(x) < 0$. ∎

The argument of [8] and [9] does not extend to infinite dimensional affine constraints because separation can not be guaranteed. The next result gives an infinite dimensional variation on [9].

[10] <u>Theorem</u>: Suppose in (P_6) that the following conditions hold:

(1) h is open; g is continuous and f is fully upper semicontinuous. (on C).

(2) $C^o \neq \emptyset$ and if $C \neq X$, h is continuous.

(3) $\exists\, x_3 \in C^o$ with $f(x_3) < u_o$, $h(x_3) = 0$.

(4) $\exists\, x_1 \in C^o$ with $g(x_1) < 0$, $h(x_1) = 0$.

Then $\exists u^+ \in B^+/\{o\}$, $z^+ \in W'$ with (P_6) equivalent to (P_6').

Proof: Let $A = \{(y,w) \mid f(x) < u_o,\ x \in C^o,\ g(x) < y, h(x) = w\}$

$= \{(y,w) \mid g(x) < y,\ h(x) = w,\ x \in C^o \cap f^{-1}(u_o - S^o)\}$.

The full upper semicontinuity of f guarantees that $\Omega = C^o \cap f^{-1}(u_o - S^o)$ is an open convex set. By the construction of A there is no x in Ω with $h(x) = 0$ and $g(x) \in -B^o$.

This means that the transposition theorem of [3.10] is applicable and $u^+ \in B^+$, $z^+ \in W'$ exist (not both zero) with

(5) $u^+(g(x)) + z^+(h(x)) \geqslant 0$ if $x \in C^o$, $f(x) < u_o$

Suppose $C \neq X$ and $\overline{x} \in \overline{C}$. Since h and g are continuous and f is upper semicontinuous,

$u^+(g(\overline{x})) + z^+(h(\overline{x}))$ is still non negative.

As in [9] since h is open $u^+ \neq 0$. Now suppose $x_2 \in C$ and $u^+(g(x_2)) + z^+(h(x_2)) = 0$. Let $x_\lambda = \lambda x_1 + (1-\lambda) x_2$, then $u^+(g(x_\lambda)) + z^+(h(x_\lambda)) < 0 \quad 0 < \lambda < 1$ and thus $f(x_\lambda) \nless u_o$ by (5). Since f is upper semicontinuous $f(x_2) \nless u_o$ and one has $x \in C$, $u^+(g(x)) + z^+(h(x)) \leqslant 0$ implies $f(x) \nless u_o$ which is the desired statement. ∎

[11] It is clear that x_o is in turn minimal for (P_6''')

$\min_S f(x)$ s.t. $u^+(g(x)) \leqslant 0$, $z^+(h(x)) = 0$, $x \in C$

since if $f(\overline{x}) < f(x_o)$ and $\overline{x}$ is feasible for (P_6''') it is feasible for (P_6') which is a contradiction.

[12] The results of the previous section could have been phrased for multivalued convex constraints; in particular the following generalization of [10] holds.

Theorem: Suppose that g in [10] is replaced by a lower semicontinuous multivalued convex mapping G and that (4) is replaced by

(4') $\exists x_1 \in C^0$ with $h(x_1) = 0, G(x_1) \cap -B^0 \neq \emptyset$.

Then (Q_6) weakmin$_s$ $f(x)$ s.t. $G(x) \cap -B \neq \emptyset$, $h(x) = 0$ is equivalent to

(Q'_6) weakmin$_s$ $f(x)$ s.t. $u^+ G(x) + z^+ h(x) \cap R^- \neq \emptyset$

for some $u^+ \in B^+/\{0\}$, $z^+ \in W'$.

Proof: The transposition theorem [3.10] is still applicable to

$A_1 = \{(y,w) \mid x \in X,\ f(x) < u,\ h(x) = w,\ [y-G(x)] \cap B^0 \neq \emptyset\}$

The argument is derived in exactly the same fashion as in [10] using the multivalued convexity rather than convexity to show that $u^+(y) + z^+(w) = 0$ implies $(y,w) \notin A$. ∎

It was noted in proposition [94] of chapter one that a real valued multivalued convex function with compact images has a single valued restriction. When G is upper semicontinuous $G(x)$ is compact (Chapter 3) and thus $u^+(G(x))$ is compact and there is a single valued $k(x) \subset u^+(G(x))$. This means that (Q_6) is equivalent to (Q''_6) weakmin$_s$ $f(x)$ s.t. $k(x) + z^+(h(x)) \leqslant 0$.

Equivalent convex and quasiconvex constraints.

In connection with (P_5) Luenberger noted that if g was only assumed quasiconvex (with respect to the orthant in R^n) that there would generally be an equivalent convex constraint. The discussion below makes this more precise.

[13] Proposition: Let Y be any real sequence space with index I. Suppose $g:X \to Y$ is strongly quasiconvex w.r.t. the coordinate

ordering then each coordinate mapping $g_i(x)$ is quasiconvex.

Proof: $g(x) \leq z \Longleftrightarrow g_i(x) \leq z_i \quad \forall\, i \in I$.

Suppose $g_1(x) < \bar{z}_1$ and $g_1(y) < \bar{z}_1 \qquad \bar{z}_1 \in R$.

Then $g_i(x) \leq \max(g_i(x), g_i(y)) = \bar{z}_i \quad i \neq 1$,

similarily $g_i(y) \leq \bar{z}_i$.

Let $\bar{z} = \{\bar{z}_i \mid i \in I\}$. By construction one has $g(x) \leq \bar{z}$ and $g(y) \leq \bar{z}$. Since g is strongly quasiconvex $g(\lambda x + (1-\lambda)y) \leq \bar{z}$, $0 \leq \lambda \leq 1$. From this it is immediate that $g_1(\lambda x + (1-\lambda)y) \leq \bar{z}_1$, $0 \leq \lambda \leq 1$, and each coordinate is quasiconvex. This argument required that all the g_i are finite together or infinite together. ▮

[14] Proposition: Let $g: X \to R$ be quasiconvex with $S(o) = \{x \mid g(x) \leq 0\}$ a closed bounded set with nonempty interior. There is an equivalent convex constraint f with $f(x_o) < 0$ for some x_o and with f finite everywhere.

Proof: Without loss of generality let $x_o = 0 \in S^o(o)$.

Let $S_1 = (S(o), o) \subset (X,R)$ and let $s_1 = (0,-1) \in (X,R)$. Let C be the cone of lines through points in S_1 with vertex s_1.

(1) C is a closed convex set: $C = \{(\lambda s, \lambda - 1) \mid s \in S(o), \lambda \geq 0\}$.

From the expression C is obviously convex. C is closed because $S(o)$ is closed and bounded. Suppose $c_n = (\lambda_n s_n, \lambda_n - 1) \to c_o$. One has $\lambda_n - 1 \to \lambda_o - 1$, $\lambda_n s_n \to b_o$. If $\lambda_o \neq 0$ then actually $s_n \to b_o/\lambda_o \in S(o)$ while if $\lambda_o = 0$, $\lambda_n s_n \to 0$ since $S(o)$ is bounded. In either case $c_o \in C$.

(2) Let $f(x) = \min\{r \mid (x,r) \in C\}$.

f is convex. Since if $f(x_1) = r_1 \qquad f(x_2) = r_2$

$(x_1,r_1)\in C_1, (x_2,r_2)\in C$ so that $(\lambda x_1 + (1-\lambda)x_2, \lambda r_1+(1-\lambda)r_2)\in C$

and $$f(\lambda x_1 +(1-\lambda)x_2)\leq \lambda f(x_1) + (1-\lambda)f(x_2).$$

(3) $f(o) = -1$.

(4) If $f(x)\leq 0$, $\exists\, r \leq 0$ with $(x,r)\in C$ since C is closed.

Thus $(x,r) = (ts,t-1)$, $o \leq t \leq 1$ and $s \in S(o)$.

Since $o \in S(o)$ $ts = ts +(1-t)o\in S(o)$ and $g(x) = g(ts) \leq 0$.

(5) If $g(x)\leq 0$ $(x,o) \in C$ and $f(x) \leq 0$.

(6) <u>f is finite everywhere</u>. By construction $0\in S^o(o)$ and

$N(o) \subset S^o(0)$;

then $x\in N(o) \Rightarrow (tx,t-1)\in C$ $\forall t > 0$.

Let $y\in X$. For $o<d<d_o$ $dy \in N(o)$ and $(tdy,t-1)\in C$.

In particular $(y, d^{-1}-1)\in C$ and $f(y) \leq d^{-1} -1$. ∎

The construction could be performed without S bounded. It would appear, however, that (4) need not hold in that case.

The next result guarantees the existence of equivalent convex constraints for a class of quasiconvex constraints.

[15] <u>Theorem</u>: Let Y by a sequence space indexed by I. Let Y have the coordinate ordering θ, and let $G:X \to Y$ be strongly quasi-convex w.r.t.θ. Suppose for each i that

(1) $\{x \mid g_i(x) \leq 0\}$ is closed and bounded with a common interior point a_o.

(2) $|g_i(x)| < \infty \iff |g_j(x)| < \infty$.

Then there is a convex mapping $F:X \to Y$ with

$$\{x \mid F(x) \leq 0\} = \{x \mid G(x) \leq 0\}$$

with F everywhere finite and with $F(a_o) < 0$.

<u>Proof</u>: $G(x)\leq 0 = x\in\bigcap_{i\in I}\{x \mid g_i(x) \leq 0\}$. By (2) and the strong quasiconvexity of G one has, using [13], that each g_i is quasiconvex. [14] and (1) now imply the existence of f_i convex

everywhere defined with

$$\{x \mid f_i(x) \leq 0\} = \{x \mid g_i(x) \leq 0\} \text{ and with } f_i(a_o) < 0.$$

Setting $F(x) = \{f_i(x)\}$ one has the desired constraint. ∎

[16] Remarks: (1) Condition (2) implies no loss of generality since if $A = \bigcap_i \{x \mid g_i(x) < \infty\}$ one can redefine $\bar{g}_i(x)$ by by $\bar{g}_i(x) = \begin{cases} g_i(x) & x \in A \\ \infty & x \notin A \end{cases}$ and $\bar{G}(x) = \{\bar{g}_i(x)\}$

$\bar{G}(x)$ then satisfies (2) and $\{x \mid G(x) \leq z\} = \{x \mid \bar{G}(x) \leq z\} \quad \forall z$.

(2) $\{x \mid G(x) \leq 0\}$ can be closed without $\{x \mid g_i(x) \leq 0\}$ closed as is seen by $G(x) = (g_1(x), g_2(x))$ with $g_2(x) = |x - \frac{1}{2}|$ and

$g_1(x) = \begin{cases} \infty & |x| \geq 1 \\ 0 & |x| < 1 \end{cases}$. The closedness of each component level set would be guaranteed by the semicontinuity of G.

At least in the cases covered by [15] equivalent convex constraints exist. This does not mean necessarily that u^+F and u^+G are equivalent.

Convex Multiplier Theorems

By requiring in (P_5) that f be convex w.r.t. S one returns to the standard problem of convex programming. Moreover, it is clear that any convex f is upper semicontinuous on lines w.r.t.S so that [1] still applies to any convex function.

The first result of this section gives a multiplier theorem for infinite dimensional affine constraints.

[17] Theorem: Let X, W, Z be convex spaces and $B \subset Y$ a closed convex cone with interior. Suppose the following hold:

(1) $C \subset X$ is convex with interior,

(2) $f:C \subset X \to R$ is convex and continuous on C,

(3) $g:C \subset X \to Z$ is convex w.r.t. B and continuous on C,

(4) $h:X \to W$ is open and affine and is continuous if $C \neq X$,

(5) $\exists x_1 \in C$ with $g(x_1) \in -B^o$, $h(x_1) = 0$,

(6) $\exists x_2 \in C^o$ with $h(x_2) = 0$.

Then

$$(P_7) \inf f(x) \text{ s.t. } g(x) \in -B \ , \ h(x) = 0 \ , \ x \in C$$

is equivalent to:

$$\inf f(x) + u^+(g(x)) + z^+(h(x)) \ , \ x \in C; \text{ for some } u^+ \in B^+, z^+ \in W'.$$

Moreoever, if the infimum u_o is achieved in (P_7) at x_o then

$$u^+(g(x_o)) = 0.$$

Proof: Let $\overline{f}(x) = (f(x)) - u_o, g(x))$ and $S = R^+ \times B$.

The Transposition theorem [3.10] can be applied to $\overline{f}$, h and S.

There is no solution to $h(x) = 0$, $\overline{f}(x) \in -S^o$ and $x \in C^o$ since $\overline{f}(x) \in -S^o \Longleftrightarrow f(x) < u_o$ and $g(x) \in -B^o$.

Thus there is $\overline{u} = (\overline{r}, u^+) \in (R^+, B^+)$, $z^+ \in W'$ with $(u^+, z^+) \neq 0$ and

$$(7) \ \overline{r}(f(x)) + u^+(g(x)) + z^+(h(x)) \geqslant \overline{r}(u_o) \quad \forall x \in C^o.$$

If $\overline{r}$ and u^+ are 0 then $z^+(h(x)) \geqslant 0 \ \forall x \in C^o$. Since h is open and by (6) $z^+(N_2) \geqslant 0$ where $N_2 = h(N(x_2))$ is a neighbourhood of 0 and $N(x_1) \subset C$. This implies $z^+ = 0$ which is impossible.

If $\overline{r} = 0$ and $u^+ \neq 0$ (5) gives some $x_1 \in C$ with $u^+(g(x_1)) + z^+(h(x_1)) = u^+(g(x_1)) < 0$ in contradiction with (7). Thus $\overline{r} \neq 0$ and w.l.o.g. $\overline{r} = 1$.

Suppose $\{x_r\}$ is a feasible net with $f(x_r) \to u_o$. (7) gives

$$f(x_r) + u^+(g(x_r)) + z^+(h(x_r)) \geqslant u_o .$$

Moreover, since x_r is feasible $g(x_r) \in -B$, $h(x_r) = 0$ and

$$(8)\ f(x_r) \geqslant f(x_r) + u^+(g(x_r)) + z^+(h(x_r)) \geqslant u_o$$

and $\inf_{x \in C} f(x) + u^+(g(x)) + z^+(h(x)) = u_o$.

If the infimum is achieved at x_o (8) shows that $u^+(g(x_o)) = 0$.∎

[18] Remark: If in (P_7) the affine constraint is finite dimensional the theorem can be proved without most of the continuity conditions analogously to the quasiconvex theorem [9] but using the next theorem rather than [1].

[19] Theorem: Let $f: C \subset X \to R$ be convex and let $G: C \subset X \to Z$ be convex as a multivalued map with respect to a closed convex cone B with interior and let $C \subset X$ be convex. Let (P_8) denote:

$$(P_8)\ u_o = \inf f(x) \quad \text{s.t. } G(x) \cap -B \neq \emptyset, \quad x \in C.$$

If there is some x_1 such that $G(x_1) \cap -B^o \neq \emptyset$ (P_8) is equivalent to:

$$\inf \quad f(x) + u^+(G(x)) \qquad x \in C.$$

If the infimum is attained there is some $y_o \in G(x_o)$ with $u^+(y_o) = 0$.

Proof: Let $A = \{(r,z) \mid f(x) \leqslant r,\ [z - G(x)] \cap B \neq \emptyset,\ x \in C\}$

(1) A is convex. Let (r_1,z_1), $(r_2,z_2) \in A$. There are $x_1, x_2 \in C$ with (i) $f(\lambda x_1 + (1-\lambda)x_2) \leqslant \lambda f(x_1) + (1-\lambda) f(x_2) \leqslant \lambda r_1 + (1-\lambda) r_2$ (ii) $\lambda x = \lambda x_1 + (1-\lambda)x_2 \in C$ (iii) $[z_i - G(x_i)] \cap B \neq \emptyset$ so there are y_1, y_2 with $y_i \in G(x_i)$ and $z_i \geqslant y_i$. Since G is m.v. convex there is some y_λ in $G(x_\lambda)$ with $y_\lambda \leqslant \lambda y_1 + (1-\lambda)y_2$ and $[\lambda z_1 + (1-\lambda) z_2 - G(\lambda x_1 + (1-\lambda) x_2)] \cap B \neq \emptyset$. This establishes that A is convex.

(2) By construction there is no (r,z) in A with $r < u_o$ and

$z \in -B^o$. Again the Hahn Banch theorem guarantees an $\bar{r} \geqslant o$ and $u^+ \in B^+$ with

$$\bar{r}(r) + u^+(z) \geqslant \bar{r}(u_o) \quad \text{if } (r,z) \in A.$$

Since in particular for any $z \in G(x)$ $(f(x),z) \in A$ when $x \in C$ one has has $\bar{r}f(x) + u^+(z) \geqslant \bar{r}u_o \qquad \forall x \in C.$

If $\bar{r} = 0$ then $u^+(G(x)) \geqslant 0 \quad \forall\, x \in C$ which contradicts $x_1 \in C$ and $G(x_1) \cap -B^o \neq \emptyset$. Thus $\bar{r} > o$ and can be taken to be 1.

Then

$$\inf_C f(x) + u^+(G(x)) \geqslant u_o$$

and they are in fact equal since there is some net $\{x_r\} \subset C$ with $f(x_r) \to u_o$ and $G(x_r) \cap -B \neq \emptyset$. Let $y_r \in -B \cap G(x_r)$ then

$$f(x_r) \geqslant f(x_r) + u^+(y_r) \geqslant u_o$$

and the infima are the same. If the infimum is attained at x_o this means $u^+(G(x_o)) \geqslant 0$. For some $y_o \in G(x_o)$, however, $y_o \in -B$ and thus $u^+(y_o) = 0$. ∎

It is apparent that [17] could also have been phrased with a multivalued constraint. The proof of multiplier theorems with equality constraints could be managed by applying [19] to

$$\inf f(x) \text{ s.t. } u^+(g(x)) + z^+(h(x)) \leqslant 0\ ,\ x \in C$$

under the conditions of [9] or [10]. To see this it is only necessary to note that since $u^+ \neq 0$ the point x_1 of [7] , [10] satisfies $u^+ g(x_1) + z^+ h(x_1) \leqslant 0$ and to remark that u^+g+z^+h is convex. One recovers complementary slackness in [19]. Finally, the conditions (2) of [9] and (3) of [10] are only necessary to guarantee the constraint g is active. If g is not active the reuult is still true but is not deducable from [9] or [10] .

Luenberger (1969) claims in a problem that the theorem of remark [18] (and so analogously [9]) holds without using in condition (2) that $x \in C^o$. This seems unlikely.

[20] Remarks: The continuity conditions on f and g can be weakened in [17] by applying [19] (which is in the single valued case just the standard result) to (P_7) deducing the existence of $u^+ \in B^+$ with $u^+(g(x_o)) = 0$ and $u_o = \inf f(x) + u^+(g(x))$ s.t. $h(x) = 0,\ x \in C$. The transposition theorem in [17] is then applied to $\bar{f}(x) = (f(x) + u^+(g(x) - u_o)$ and $S = R$. The statement of [17] could then be weakened by replacing (2), (3) by the condition that $f + u^+g$ is continuous on C. This has the drawback of using in the hypotheses the multiplier whose existence is desired.

A simple minimization theorem for problems with non-affine equality constraints in proved below. It seems worth including because it involves several previously discussed concepts.

[21] Theorem: Suppose $f:X \rightarrow Y$ is fully upper semicontinuous and (P) strictly quasiconvex with respect to S. Suppose that $h:X \rightarrow Z$ is a Fréchet differentiable map between Banach spaces with $R(h'(\bar{x})) = Z$ and with h continuously differentiable in a neighbourhood of $\bar{x}$. Suppose $\bar{x}$ is weakly minimal for

(P_8) min $f(x)$ subject to $h(x) = 0$

Then $\bar{x}$ is also minimal for

(P_8') min $f(x)$ subject to $(h'(\bar{x}))\ (x - \bar{x}) = 0$.

Proof: Since f is (P) strictly quasiconvex and fully upper semicontinuous [2.23] shows that $\bar{x}$ is minimal for f over

$$\bar{x} + T\ (N(h),\ \bar{x}) = \bar{x} + N(h'(\bar{x}))$$

where the equivalence follows from the discussion of regularity in [2.40]. Since $y \in \bar{x} + N(h'(\bar{x}))$ exactly when $(h'(\bar{x}))(x - \bar{x}) = 0$ the theorem is established.▮

[22] Corollary: If f is actually real valued continuous and convex then $\bar{x}$ is a minimum of $f(x) + z^{+}(h'(\bar{x})(x - \bar{x})$ for some $z \in Z'$.

Proof: The transposition theorem [3.10] guarantees, since $h'(\bar{x})$ is surjective and hence open, that for some $z^{+} \in Z'$, $\bar{r} \geqslant 0$

$$\bar{r}(f(x) - f(\bar{x})) + z^{+}(h'(\bar{x}))(x - \bar{x}) \geqslant 0 \quad \forall x \in X.$$

$\bar{r}$ can be taken non zero by [3.12]. Since $\bar{x}$ is feasible and

$$f(x) + z^{+}(h'(\bar{x}))\ (x - \bar{x}) \geqslant f(\bar{x})$$

the result is proved.▮

Convex Programs with $f:X \rightarrow Y$

The preceeding theorems on convex multipliers can be adapted to prove results when $f:X \rightarrow Y$ is convex w.r.t. S and x_o is a weak minimum for (P) with respect to S.

[23] Proposition: Let $f:X \rightarrow Y$ be convex w.r.t. S, $S^o \neq \emptyset$ and A be a convex set. If x_o is a weak minimum for f over A there is some $s^{+} \in S^{+}/\{0\}$ with x minimal for $s^{+}(f(x))$ over A. $s^{+}f$ is clearly still convex.

Proof: Let $E = \{z \mid f(x) \leqslant z + f(x_o), x \in A\}$. E is a convex set and is separated from $- S^o$ because x_o is a weak minimum. The Hahn-Banach Theorem then guarantees some non zero s^{+} with

$$s^{+}(z) \geqslant 0 \geqslant s^{+}(-s) \quad \forall z \in E,\ s \in S$$

which means that $s^{+} \in S^{+}/\{0\}$ and $s^{+}(f(x)) \geqslant s^{+}(f(x_o))\ \forall x \in A$.▮

[24] Theorem: Suppose x_o is a weak minimum with respect to S for

(P_9) min $f(x)$ s.t. $g(x) \in -B$, $x \in C$

with $f:X \to Y$ convex w.r.t.S, $g:X \to Z$ convex w.r.t.B, $B^o \neq \emptyset$, and $C \subset X$ a convex set. Suppose there is some $x_1 \in C$ with $g(x_1) \in -B^o$. Then there is some $T \in B^S$ with $Tg(x_o) = 0$ such that x_o is a weak minimum for

$$f(x) + Tg(x) \quad \text{s.t.} \quad x \in C.$$

Proof: [23] and [19] give in conjunction the existence of $u^+ \in B^+$, $s^+ \in S^+/\{0\}$ with $u^+(g(x_o)) = 0$ and

(1) $s^+(f(x_o)) \leq s^+(f(x)) + u^+(g(x))$ if $x \in C$.

Let $s \in S^o$ then $s^+(s) > 0$ since $s^+ \neq 0$.

Define T by $T(x) = s^+(s)^{-1} s\, u^+(x)$. Clearly $T(B) \subset S$ and T is continuous. Hence $T \in B^S$. Equally clearly $T(g(x_o)) = 0$.

Suppose now that there is some $x \in C$ with

(2) $f(x) + T(g(x)) < f(x_o) + T(g(x_o))$.

Then

(3) $s^+(f(x) + T(g(x)) < s^+(f(x_o)) + s^+(T(g(x_o)) = s^+(f(x_o))$

but $s^+T = u^+$ and (3) contradicts (1). Thus T satisfies the claims of the statement. ∎

It is worth remarking that the same trick could be performed in chapter four to rewrite the Fritz John Theorems on weak minimization in operator form.

Strong minima:

Turning now to strong minima, one faces much more of a problem in obtaining multiplier theorems. The only easy avenue appears to lie in application of the generalized Kuhn-Tucker results of the last chapter. One case in which direct theorems can be proved does exist and this is dealt with first. The problem is still (P_9) of [24].

[25] Proposition: $f: X \to Y$ has a strong minimum (with respect to S a closed convex cone) over A at x_o if and only if $s^+ f$ has a minimum at x_o for all $s^+ \in S^+$.

Proof: $f(x) \geqslant_S f(x_o) \quad \forall x \in A \Rightarrow s^+(f(x)) \geqslant s^+(f(x_o)) \quad \forall x \in A$.

Conversely if $s^+(f(x)) \geqslant s^+(f(x_o)) \quad \forall x \in A, \; s^+ \in S^+$ then

$$f(x) - f(x_o) \in (S^+)^+ \qquad \forall x \in A$$

which since S is closed gives $f(x) \geqslant_S f(x_o) \quad \forall x \in A$. ∎

[26] Theorem: Let $f: X \to R^n$ be convex w.r.t. a closed convex pointed cone S. Let $g: X \to Z$ be convex w.r.t. B, a closed cone with interior. Let $C \subset X$ be convex and suppose some $x_1 \in C$ has $g(x_1) < 0$. Let $S_1 \supset S$ be defined by

$$x \in S_1 \Leftrightarrow Kx \in R^{n+}$$

where K is an invertible matrix with rows in S^+. Suppose x_o is a strong minimum with respect to S for (P_9). Then x_o is a strong minimum with respect to S_1 for

$$f(x) + T_o(g(x)), \quad x \in C$$

where $T_o \in B^{S_1}$ and $T_o(g(x_o)) = 0$.

Proof: Since $S \subset R^n$ S is pointed $\Leftrightarrow (S^+)^o \neq \emptyset$. Thus there is at least one such matrix K and cone S_1. By [25] x_o is minimal for

(1) $\min s_i^+(f(x))$ subject to $g(x) \in -B, \; x \in C \quad s_i^+ \in S^+, i=1,\ldots,n.$

[19] can be applied to (1) since $s_i^+ f$ is convex. One derives that x_o is minimal for

(2) $s_i^+(f(x)) + u_i^+(g(x))$ s.t. $x \in C \;\; ; \;\; u_i^+ \in B^+, \; u_i^+(g(x_o)) = 0$

for $i = 1,\dots,n$. Letting $s_1^+,\dots,s_n^+$ be the rows of K and setting $T = (u_1^+,\dots,u_n^+)^T$ one has

(3) $K(f(x) - f(x_o)) + Tg(x) \in R^{n+} \quad \forall x \in C.$

Since K^{-1} exists this can be written as

(4) $K\left[f(x) - K^{-1}T(g(x)) - f(x_o)\right] \in R^{n+} \quad \forall x \in C.$

But this is just the definition of

$$f(x) - K^{-1}\,T(g(x)) \geqslant_{S_1} f(x_o) \quad \forall x \in C.$$

Setting $T_o = K^{-1}\,T$; $\quad T_o(g(x_o)) = 0$ and $T_o \in B^{S_1}$. |

In the case that S is completely determined by n linearly independent constraints the cones S_1 and S agree and the unconstrained problem is equivalent to the initial problem. This is certainly true if $S = R^{n+}$ and in that situation K can be chosen to be the identity.

If in [26] the minimum was only required to be weak one could require the existence of $(s_1^+,\dots,s_n^+)$ linearly independent in S^+ with x_o minimal for $s_i^+f(x)$ s.t. $g(x) \in -B$, $x \in C$. It is apparent, from the last theorem that this is equivalent to x_o being a strong minimum over S_1.

The preceeding discussion is therefore incomplete unless $S = S_1$. More general results can be derived from the differential conditions of the last chapter:

[27] Theorem: Let X, Y, Z be Banach Spaces with Y reflexive and $S \subset Y$ a closed convex normal cone. Suppose $B \subset Y$ is a closed convex cone with interior and that $f:X \to Y$ is convex on C w.r.t.S and $g:X \to Z$ is convex on C w.r.t.B. Suppose f and g are differentiable at x_o with $g'(x_o)$ surjective. Suppose there is some $x_1 \in C^o$ with $x_1 \in g^{-1}(-B)^o$.

Then if s_o is a strong minimum for f over $g^{-1}(-B) \cap C$ w.r.t.S there is some $T_o \in B[Z,Y]$ with $T_o(B) \subset S$ and $T_o(g(x_o)) = 0$ such that x_o is a strong minimum w.r.t.S for

$$f(x) + T_o g(x) \qquad \forall x \in C.$$

Proof: The conditions guarantee that

(1) $P(C,x_o) \cap P(\Delta, x_o) = P(A,x_o)$ [2.18].

(2) $g'(x_o)^{-1}(P(B,g(x_o)) = P(\Delta, x_o)$ [5.50].

Thus the generalized Guignard condition of [34] of the last chapter gives T_o with the requisite properties such that

(3) $f'(x_o) + T_o(g'(x_o)) \in P(C, x_o)^S$.

$T_o(P(-B,g(x_o)) \subset -S$ implies $T_o(B) \subset S$ and $T_o(g(x_o)) = 0$ since B is a convex set and S is pointed (normal).

Let $x \in C$, since C is convex $x - x_o \in P(C,x_o)$.

(3) gives

$$(f'(x_o))(x - x_o) + T_o(g'(x_o))(x - x_o)) \in S$$

Because f and g are convex one has

(4) $(f'(x_o))(x - x_o) \leqslant_S f(x) - f(x_o)$

(5) $(g'(x_o))(x - x_o) \leqslant_B g(x) - g(x_o)$.

Substituting these in the previous expression one derives

$$f(x) - f(x_o) + T_o(g(x) - g(x_o)) \geqslant_S 0$$

and since $T_o(g(x_o)) = 0$ this is the desired result.

[28] Definition: In keeping with the terminology for the real valued case $f(x) + T(g(x))$ is called a Lagrangian and is denoted $L(T,x)$. It is a mapping from $B(Z,Y) \times X$ to Y.

It is immediate that whenever a multiplier T_o exists, with $T_o \in B^S$, $T_o g(x_o) = 0$, (T_o, x_o) is a saddle point of the Lagrangian over $B^S \times C$.

[29] Proposition: The existence of $T_o \in B^S$ with $T_o(g(x_o)) = 0$ such that $f(x_o) = \min_C f(x) + T_o g(x)$ is equivalent to (T_o, x_o) being a saddle point of $L(T,x)$ over $B^S \times C$.

S is supposed to be a pointed cone.

Proof: Suppose $f(x_o) \leqslant_S f(x) + T_o(g(x))$.

Since $T_o(g(x_o)) = 0$ this gives

$$L(T_o,x_o) \leqslant L(T_o,x)$$

Since $T_o g(x_o) = 0$ and $g(x_o) \in -B$ one has, for any $T \in B^S$, $T(g(x_o)) \in -S \leqslant T_o(g(x_o))$ and so

$$L(T,x_o) \leqslant L(T_o,x_o).$$

Conversely if (T_o,x_o) is a saddle point over $B^S \times C$

$$f(x_o) + Tg(x_o) \leqslant f(x_o) + T_o g(x_o) \leqslant f(x) + T_o g(x)$$

from which it is clear that $T_o(g(x_o)) \in S \cap -S = \{0\}$ and x_o is thus minimal over C.▮

[30] Sensitivity:

Theorem: Consider the problems

strong $\min_S f(x)$ s.t. $g(x) \leqslant z_i, x \in C \quad i = 1,2$.

Suppose that multipliers T_i exist for each problem then

$$T_2(z_1 - z_2) \leqslant_S f(x_1) - f(x_2) \leqslant_S T_1(z_2 - z_1).$$

Proof: The multiplier T_1 gives

$$f(x_1) = f(x_1) + T_1(g(x_1) - z_1) \leqslant f(x) + T_1(g(x) - z_1) \quad \forall x \in C.$$

Setting $x = x_2$ one has

$$f(x_1) - f(x_2) \leqslant T_1(g(x_2) - z_1) \leqslant T_1(z_2 - z_1)$$

since $g(x_2) \leqslant_B z_2$ and $T(B) \subset S$.

The same argument can be applied to T_2, x_2 and produces the other inequality.▮

The primal function

[31] Suppose that for z in some neighbourhood N of o the problem

$$\text{strong}\ \min_S f(x) \quad \text{s.t.}\ g(x) \leqslant z \ , \quad x \in C$$

has solution. Let $\Pi(z)$ denote this solution. $\Pi(z)$ is called the primal function associated with the given problem. S is assumed pointed throughout.

[32] Proposition: Suppose $g: X \to Z$ is convex w.r.t. B and C is convex.

(1) If f is convex w.r.t. S Π is convex w.r.t.S on N.

(2) If f is strongly quasiconvex w.r.t.S Π is also strongly quasiconvex w.r.t.S on N.

Proof: Suppose f is convex. Let $z_1, z_2 \in N$. Let $f(x_1) = \Pi(z_1)$; $f(x_2) = \Pi(z_2)$. Then $tx_1 + (1-t)x_2 \in C \quad 0 \leqslant t \leqslant 1$

Since f is convex $f(tx_1 + (1-t)x_2) \leqslant_S t\Pi(z_1) + (1-t)\Pi(z_2)$.

Since $g(x_1) \leqslant z_1$, $g(x_2) \leqslant z_2$ and g is convex w.r.t.B $g(tx_1 + (1-t)x_2) \leqslant tz_1 + (1-t)z_2$. Thus $tx_1 + (1-t)x_2$ is a potential solution for $\Pi(tz_1 + (1-t)z_2)$. Since it is supposed that the strong minimum exists $(tz_1 + (1-t)z_2 \in N)$

$$\Pi(tz_1 + (1-t)z_2) \leqslant_S f(tx_1 + (1-t)x_2) \leqslant_S t\Pi(z_1) + (1-t)\Pi(z_2).$$

(2) Suppose f is strongly quasiconvex. With x_1, x_2 as before one has, if $\Pi(z_1) \leqslant y$, $\Pi(z_2) \leqslant y$, that $f(tx_1 + (1-t)x_2) \leqslant y$. As in (1) $tx_1 + (1-t)x_2$ is feasible for the problem associated with $\Pi(tz_1 + (1-t)z_2)$ and this means $\Pi(tz_1 + (1-t)z_2) \leqslant y$. ∎

It is clear that if $z_1 - z_2 \in B$ then $\Pi(z_2) - \Pi(z_1) \in S$ (for $z_1, z_2 \in N$).

[33] <u>Definition</u>: The <u>dual function</u> φ of Π defined on B^S is given by $\varphi(T) = \text{strongmin}_{s\; x \in C} (f(x) + Tg(x))$ where as was the case with Π the minimum is supposed to exist on some non trivial set $K = \text{dom}\,\varphi$. Let ∇ denote dom Π.

[34] <u>Proposition</u>: If f is convex then φ is concave and

$$\varphi(T) = \min_{z \in \nabla} \left[\Pi(z) + T(z) \right] \text{ if it exists.}$$

The domain of definition may well not be a convex set. It is assumed throughout that definitions of convexity (concavity) are modified to take this into account.

<u>Proof</u>: (1) It is simple to verify that φ is concave.

Then if $T \in B^S$ and $z \in \nabla$

$$\varphi(T) = \min_{x \in C} \left[f(x) + T(g(x)) \right].$$

If $g(x) \leq z$ then $T(g(x)) \leq T(z)$ and

$$\varphi(T) \leq \Pi(z) + T(z).$$

Also, if $z_1 = g(x_1)$

$$f(x_1) + T(g(x_1)) \geq \Pi(z_1) + T(z_1).$$

This with the definition gives

$$\varphi(T) \geq_s \min_{z \in \nabla} \Pi(z) + T(z) \geq_s \varphi(T). \blacksquare$$

[35] <u>Theorem</u>: (Generalized Lagrange duality)

Suppose the conditions of [27] are met and that $f(x_o) = \Pi(0)$.

Then if $\varphi(T)$ exists on K

$$\Pi(0) = \max_{s\; T \in B^S \cap K} \varphi(T) = \varphi(T_o)$$

and $T_o(g(x_o)) = 0$.

Proof: $\varphi(T) = \min_{C\,s}\left[f(x) + T(g(x))\right]$ exists on K. Let

$T \in B^s$, then if $g(x) \in -B$ $T(g(x)) \in S$

and $\varphi(T) \leq_s f(x)$ $\forall\, x \in C$ with $g(x) \in -B$.

Thus $\varphi(T) \leq_s f(x_o)$.

The result of [24] says that $\varphi(T_o) = f(x_o)$ and that

$T_o(g(x_o)) = 0$. This concludes the theorem. ∎

The major limitations on this extension are that the domains of definition of φ, Π need not be convex even when the conditions of [27] are satisfied. Thus one may well be maximizing over non convex sets.

It is also difficult to verify in general when $\varphi(T)$ exists even if one knows $\varphi(T_o)$ exists. The result, however, does indicate that the duality can be extended.

The result of [35] can be reworded as

$$\min_{s\; z \in -B \cap \Gamma} \Pi(z) = \max_{s\; T \in B^s \cap K} \varphi(T) .$$

Minimax Theorems

It seems natural while examining the various extensions of programming theory from real valued to more general objective functions to investigate minimax theorems. It turns out that the Sion minimax theorem has an entirely adequate extension with the essential and limiting proviso, unlike the real valued case, that both the minimax and maximin have to be assumed to exist. The proof is derived from Browder's (1968) investigation of fixed points of multivalued mappings.

[36] Theorem: Let $K_1,\dots,K_n$ be compact convex sets in topological vector spaces $E_1,\dots,E_n$. Let $f_j : \bigtimes_{i=1}^{n} K_i = K \to Y$ a convex space with a closed convex cone B with nonempty interior. Suppose $\hat{K}_j = \bigtimes_{j\neq i} K_i$ and

(1) For each $j=i,\dots,n$ $f_j(x_j,\hat{x}_j)$ is fully lower semicontinuous with respect to B on $\hat{K}_j$ for fixed x_j in K_j.

(2) $f_j(x_j,\hat{x}_j)$ is strongly quasiconcave with respect to B on K_j for fixed $\hat{x}_j$ in $\hat{K}_j$.

(3) Let $\{a_1,\dots,a_n\} \subset Y$. Suppose for each j and $\hat{x}_j \in \hat{K}_j$ there is $y_j \in K_j$ with $f_j(y_j,\hat{x}_j) - a_j \in B^o$.

Then there is $u \in K$ with $f_j(u) - a_j \in B^o$ $j = 1,\dots,n$.

Proof: Let $S_j = \{u \mid u \in K,\ f_j(u) - a_j \in B^o\}$ then $S_j(x_j) = \{\hat{x}_j \mid \hat{x}_j \in \hat{K}_j, f_j(x_j,\hat{x}_j) - a_j \in B^o\}$ is open by (1) and $S_j(\hat{x}_j) = \{x_j \mid x_j \in K_j\ ,\ f_j(x_j,\hat{x}_j) - a_j \in B^o\}$ is convex by (2) and proposition [18] of chapter one.

By (3) $S_j(\hat{x}_j) \neq \emptyset$ if $\hat{x}_j \in \hat{K}_j$.

The sets $S_j, S_j(x_j), S_j(\hat{x}_j)$ then satisfy all the conditions of Theorem 11 of Browder (1968) (which was constructed to prove the case Y = R) and the theorem allows one to deduce that some $u \in K$ exists with

$$u \in \bigcap_{j=1}^{n} S_j$$

This u satisfies the claim of the theorem. ∎

[37] Theorem: Suppose f maps $K_1 \times K_2$ into Y where K_1 and K_2 are non empty compact convex sets in separated topological vector spaces E_1 and E_2. Let Y be a convex space and $B \subset Y$ a pointed

closed convex cone with interior. Suppose that for fixed y_2 in K_2 $f(x,y_2)$ is fully lower semicontinuous w.r.t. B and strongly quasiconvex on K_1. Suppose, similarly that for fixed x_1 in K $f(x_1,y)$ is fully upper semicontinuous and strongly quasiconcave on K_2. If both minimization and maximization are supposed to be strong with respect to B and

$$\min_{K_1}\max_{K_2} f(x,y) = A_1$$

$$\max_{K_2}\min_{K_1} f(x,y) = A_2$$

then $A_1 = A_2$.

(It is implicit in the hypotheses that the minimizations and maximizations are well defined).

<u>Proof</u>: $f(x,y) \leqslant \max_{y\in K_2} f(x,y)$ and thus

$$\min_{x\in K_1} f(x,y) \leqslant \min_{x\in K_1}\max_{y\in K_2} f(x,y) = A_1$$

This clearly implies that if the strong max. of the left hand side exists that $A_2 \leqslant A_1$.

Set $f_1(x,y) = -f(x,y)$; $f_2(x,y) = f(x,y)$.

Then f_1, f_2 satisfy conditions (1) (2) of [36].

Let $a \in B^o$ then

$$\min_{x\in K_1}\max_{y\in K_2} f(x,y,) = A_1 = \min_{x\in K_1} f(x,y(x)).$$

For each $x \in K_1 = \hat{K}_2$ $f(x,y(x)) \geqslant A_1$

and $f_2(x,y(x)) > A_1 - a$.

Similarly for each $y \in K_2$ $f(x(y),y) < A_2 + a$

and $f_1(x(y),y) > -A_2 - a$.

These points $x(y)$, $y(x)$ satisfy condition (3) of [36].

Since all the conditions are satisfied one has some point

$u = (x,y)$ with $f_1(x,y) > -A_2 - a$

$f_2(x,y) > A_1 - a,$

or $f(x,y) < A_2 + a \qquad A_1 - a < f(x,y).$

Since $a \in B^o$ is arbitrary this gives

$$A_1 \leqslant f(x,y) \leqslant A_2$$

which with $A_2 \leqslant A_1$ and $B \cap -B = 0$ implies that $A_1 = A_2$.∎

It is apparent from the theorem that it is too much to hope that the existence of A_1 is sufficient for $A_1 = A_2$ or vice versa. This is clarified by the following example.

[38] Example: Let A be the matrix $\begin{bmatrix} (0,0) & (1,0) \\ (0,0) & (0,1) \end{bmatrix}$ with entries a_{ij} in R^2. Let $f(x,y) = \sum_{i,j=1}^{2} x_i a_{ij} y_j = x^T A y$ and let

$K_1 = \{x \mid x = (x_1, x_2)\ ,\ x_1 + x_2 = 1\ ,\ 0 \leqslant x_i \leqslant 1\}$

$K_2 \quad \{y \mid y = (y_1, y_2)\ ,\ y_1 + y_2 = 1\ ,\ 0 \leqslant y_i \leqslant 1\}$

(a) $\max_{y \in K_2} \min_{x \in K_1} f(x,y) = \max_{y \in K_2} (o,o) = (o,o)$

(b) $\min_{x \in K_2} \max_{y \in K_2} f(x,y) = \min_{x \in K_1} x_2 \max_{y \in K_2} (y_1, y_2)$

but this interior maximum does not exist strongly since all the points $(y_1, 1 - y_1)$ are incomparable. It is clear that all the other conditions of theorem [37] are satisfied since f is continuous and linear in each variable. In a sense (o,o) is still a minimax since $x_2 = 0$ in (b) is minimal for all $(x_2 y_1, x_2 y_2)$.

Despite this type of drawback it seems worth phrasing the following generalisation of the Von Neumann minimax theorem.

[39] Theorem: Suppose A is a mxn matrix $[a_{ij}]$ with entries in a topological vector space Y with a pointed closed convex cone with interior $B \subset Y$.

Let $K_1 = \{x \mid x = (x_1, x_2, \ldots, x_m)\ \sum x_i = 1\ ,\ x_i \geq 0\}$

and let $K_2 = \{y \mid y = (y_1, \ldots, y_n), \sum y_j = 1, \quad y_j \geq 0\}$

Suppose that (1) $\max_i a_{ij}$ exists strongly w.r.t. B for each j and (2) $\min_j a_{ij}$ exists strongly for each i.

Then

$$\min_{x \in K_1} \max_{y \in K_2} xAy = \max_{y \in K_2} \min_{x \in K_1} xAy$$

if all the strong optimizations are well defined.

Proof: All the conditions of [37] are met since xAy is linear and continuous.▮

The next proposition shows that condition (1) and (2) of [39] are essential.

[40] Proposition: Let $a_1, \ldots, a_n$ be elements of Y. A necessary and sufficient condition for

(1) strong $\max_B\limits_{C_n} \sum \lambda_i a_i$ to exist ,

where $$C_n = \left\{\lambda \,\middle|\, \sum_{i=1}^{n} \lambda_i = 1,\ \lambda_i \geq 0\right\},$$

is for

(2) strong $\max_B\limits_{1 \leq i \leq n} a_i$ to exist.

Proof: The proof proceeds by induction. Suppose n is the smallest natural number such that a set $\{a_i, \ldots, a_n\}$ exists satisfying (1) but not (2). Let $\bar{\lambda}$ be optimal in (1). Then

$$\bar{\lambda}_1 a_1 + \cdot \ \cdot + \bar{\lambda}_{n-1} \ a_{n-1} + \bar{\lambda}_n a_n \geqslant \sum_{i=1}^{n} \lambda_i a_i \quad \forall \lambda \in C_n.$$

This can be rewritten as

$$(3) \qquad \sum_{i=1}^{n-1} \bar{\lambda}_i (a_i - a_n) \geqslant \sum_{i=1}^{n-1} \lambda_i (a_i - a_n) \quad \lambda \in C_n.$$

Since $\bar{\lambda}_n = 1$ is impossible, as it would imply a_n is maximal for (2), one has $\sum_{i=1}^{n-1} \bar{\lambda}_i = L > 0$.

Now in particular (3) holds for any $\lambda \in C_n$ with $\sum_{i=1}^{n-1} \lambda_i = L \leqslant 1$

or equivalently

$$\sum_{i=1}^{n-1} L^{-1} \ \bar{\lambda}_i (a_i - a_n) \geqslant \sum_{i=1}^{n} \lambda_i (a_i - a_n) \quad \lambda \in C_{n-1}.$$

This means that $(L^{-1}\bar{\lambda}_1, \ldots, L^{-1}\bar{\lambda}_{n-1})$ is optimal for (1) with the set $\{a_1 - a_n, \cdot \ \cdot \ , a_{n-1} - a_n\}$ over C_{n-1}. The induction hypothesis then implies that this set has a largest member with respect to B; that is

$$(a_j - a_n) - (a_i - a_n) \in B \qquad i = 1, \ldots, n-1$$

and equivalently $a_j \geqslant a_i \qquad i=1, \ldots, n-1$. There is no loss in assuming that $j = 1$. One then has that the maximum in (1) must be $\bar{\alpha} a_1 + (1-\bar{\alpha}) a_n$ for some $0 \leqslant \bar{\alpha} \leqslant 1$. The values 0 and 1 can be excluded since a_1, a_n are assumed incomparable. Setting $\alpha = \bar{\alpha} + \epsilon$ where $\epsilon > 0$ is chosen such that $0 < \alpha = \bar{\alpha} + \epsilon < 1$ one can deduce from (3) that

$$\bar{\alpha} a_1 + (1-\bar{\alpha}) a_n \geqslant (\bar{\alpha} + \epsilon) a_1 + (1 - (\bar{\alpha} + \epsilon)) a_n$$

which gives $\epsilon a_n \geqslant \epsilon a_1$ a contradiction. Thus no such set exists and (1) implies (2). The converse is immediate. ∎

This condition (which is always met in a total ordering) is thus neccessary in [39] to even begin looking for a minimax. It is clear that if there is a saddle point

$a_{i_o j_o}$ with $a_{ij_o} \geqslant a_{i_o j_o} \geqslant a_{i_o j}$ $A_{j_o} \leqslant A_j$ $_{i_o}A \leqslant {}_iA$

then it is certainly a minimax. (Here $A_j, {}_iA$ denote rows and columns respectively)

Relating Saddle points and Minimax points one has for pointed cones.

[41] Proposition: If the $\min_C \max_D f(x,y)$ and $\max_D \min_C f(x,y)$ are defined any minimax point is a saddle point and vice versa.

Proof: If (x_o,y_o) is a saddle point for f over $C \times D$

$$f(x_o,y) \leqslant_B f(x_o,y_o) \leqslant_B f(x,y_o) \quad x \in C, y \in D \text{ so}$$

$$\min_C \max_D f(x,y) \leqslant \max_{D\,B} f(x_o,y) = f(x_o,y_o) = \min_C f(x,y_o)$$

$$\leqslant \max_D \min_C f(x,y)$$

and since the other inequality always hold (x_o,y_o) is a minimax.

The converse is clear. ∎

Chapter Seven

SECOND ORDER CONDITIONS

Second Order Necessary and Sufficient Conditions

When the generalized Kuhn-Tucker conditions are not also sufficient it is possible to have the Lagrangian stationary at a point which is not optimal for the associated problem. In this case further information concerning the nature of optimal points can be extracted by examining the second derivative of the Lagrangian - assuming that it exists.

Second order necessary conditions have been derived by McCormick (1967) and others for the problem (P_1) min $f(x)$ subject to

$$g_i(x) \leqslant 0 \qquad i = 1,\ldots,n$$
$$h_j(x) = 0 \qquad j = n+1,\ldots,p$$

where all the functions concerned are real valued and defined on R^m. McCormick introduces a second order constraint condition which he uses in conjunction with the Kuhn-Tucker constraint qualification to derive his necessary condition. The first theorem of this section generalizes this result to a qualification which can be used in conjunction with the Guignard constraint qualification. McCormick's condition is then derived as a special case.

[1] Theorem: Suppose $f: X \to R$, $g: X \to Z$ are twice compactly differentiable at a point x_o which is optimal in a sequential, barrelled space for

(P) min $f(x)$ s.t. $g(x) \in B$, $x \in C$.

Suppose further that the following conditions are met:

(1) The Guignard Necessary condition holds at x_o. That is for some closed convex cone G such that $G \cap K \subset P(A,x_o)$ and some $u^+ \in P^+(B,g(x_o))$

$$f'(x_o) - u^+g'(x_o) \in G^+.$$

(2) If $h \in (G \cap K) \cap -(G \cap K)$ then for some nets $\{x_n\} \subset A$, $\{\lambda_n\} \geqslant o$ with $x_n \to x_o$ and $h_n = \lambda_n(x_n - x_o) \to h$ there exists a net $\{k_n\}$ with $k_n \in -P(A,x_n)$ and such that $k_n \to h$ and $\lim_n \lambda_n(k_n - h) = z \in -G$.

(3) For this net $\{x_n\}$, $P(B,g(x_n)) \subset P(B,g(x_o))$ if $n \geqslant n_o$.

(4) $\overline{\lim_n} (f'(x_n))(\lambda_n k_n) \geqslant o$.

Then when $h \in G \cap -G$ and $(g'(x_o))(h) \in P(B,g(x_o)) \cap -P(B,g(x_o))$ one has

$$(f''(x_o))((h,h)) - u^+(g''(x_o))((h,h)) \geqslant o.$$

<u>Proof</u>: $(g''(x_o))((h,h)) = \lim_n \dfrac{(g'(x_n) - g'(x_o))(h)}{\lambda_n^{-1}}$

where $x_n = \lambda_n^{-1}h_n + x_o$, $h_n \to h$ as in (2).

$$\lim_n \frac{(g'(x_n) - g'(x_o))(h)}{\lambda_n^{-1}} = \lim_n \lambda_n \left[(g'(x_o + \lambda_n^{-1}h_n)(k_n) - (g'(x_o))(h)\right] - \lim_n (g'(x_n))(\lambda_n(k_n - h))$$

with k_n as guaranteed by (2).

Thus $(g''(x_o))((h,h)) = \lim_n \lambda_n \left[(g'(x_n))(k_n) - (g'(x_o))(h)\right] - (g'(x_o))(z)$

where the final limit can be taken since $\lambda_n(k_n - h) \to z$ and since $g'(x_n) \to g'(x_o)$.

Now $(g'(x_o))(h) \in P(B,g(x_o))$ by hypothesis and by (2) $k_n \in -P(A,x_n) \subset -P(\Delta,x_n)$. Using proposition [2.32] one has $(g'(x_n))(-k_n) \in P(B,g(x_n))$. For $n \geqslant n_o$ one derives using (3) that

$$-(g'(x_n))(k_n) \in P(B,g(x_o)).$$

Since $P(B,g(x_o))$ is a closed convex cone and $(g'(x_o))(h)$ is assumed in $P(B,g(x_o)) \cap - P(B,g(x_o))$

$$\lambda_n[(g'(x_n))(k_n) - (g'(x_o))(h)] \in -P(B,g(x_o)) \qquad n \geqslant n_o.$$

This in turn gives

(5) $(g''(x_o))((h,h)) + (g'(x_o))(z) \in - P(B,g(x_o))$.

In the same way one derives

$$(f''(x_o))((h,h)) + (f'(x_o))(z) = \lim_n \lambda_n[(f'(x_n))(k_n) - f'(x_o))(h)].$$

Since $h \in G \cap - G$ and $f'(x_o) - u^+g(x_o) \in G^+$, $(f'(x_o))(h) - u^+(g(x_o))(h) = o$. Moreoever, since $u^+ \in P^+(B,g(x_o))$ and $h \in K \cap - K$, $u^+(g'(x_o))(h) = 0$ which means that $(f'(x_o))(h) = 0$. Condition (4) then gives

(6) $(f''(x_o))((h,h)) + (f'(x_o))(z) \geqslant o$.

Collecting (5) and (6) one has since $u^+ \in P^+(B,g(x_o))$

(7) $(f''(x_o))((h,h)) - u^+(g'(x_o))((h,h))$

$$\geqslant - [(f'(x_o))(z) - u^+(g'(x_o))(z)].$$

Since $f'(x_o) - u^+g'(x_o) \in G^+$ and $z \in - G$ (7) becomes $(f''(x_o))((h,h)) - u^+(g''(x_o))(h,h)) \geqslant o$ if $h \in G \cap - G$ and $(g'(x_o))(h) \in P(B,g(x_o)) \cap - P(B,g(x_o))$. |

An equality constraint can be incorporated in [1] if G is specified more closely. This mirrors the situation for regularity conditions in the first order Fritz John conditions.

[2] Theorem: Suppose $h: X \to W$ is twice differentiable with $h'(x_o)$ surjective. Suppose X is fully complete and W is barrelled. Suppose that in Theorem [1] $C = N(h)$ and that $G = P(N(h), x_o) = h'(x_o)^{-1}\{o\}$. Then if $(h'(x_o))(h) = o$ and $(g'(x_o))(h) \in P(B,g(x_o) \cap - P(B,g(x_o))$ one has

$(f''(x_o))((h,h)) - u^+(g''(x_o))((h,h)) + z^+(h''(x_o))((h,h)) = o$

with $z^+ \in W'$ and u^+ as above.

Proof: The Guignard condition (1) [1] now becomes $(h'(x_o))(h) = 0$ implies $(f'(x_o))(h) - u^+(g'(x_o))(h) = 0$. Since $R(h'(x_o)) = W$ the Farkas Lemma [3.6] can be used to derive $z^+ \in W'$ with

(1) $(f'(x_o))(x) - u^+(g'(x_o))(x) + z^+(h'(x_o))(x) = 0, \forall x \in X.$

As in [1] one can also show

(2) $\lim \lambda_n[(h'(x_n))(k_n) - (h'(x_o),h)] = (h''(x_o))(h,h)) + (h'(x_o))(z).$

By hypothesis $(h'(x_o))(h) = 0$ and since $k_n \in - P(A,x_n)$ $k_n \in - P(N(h),x_n)$, $-k_n = \lim_m \lambda_{nm}(x_{nm} - x_n)$ with

$h(x_{nm}) = 0 = h(x_n)$ and $\lim_m \lambda_{nm}[h(x_{nm}) - h(x_n)] = 0.$

This last limit is $-(h'(x_n))(k_n)$ so that (2) gives

(3) $(h''(x_o))((h,h)) + (h'(x_o))(z) = 0.$

Multiplying (2) by z^+ and adding it to the equation (7) of [1] one derives that

$(f''(x_o)(h,h)) - u^+(g''(x_o))((h,h)) + z^+(h''(x_o))((h,h)) \geq$

$(f'(x_o))(z) - u^+(g'(x_o))(z) + z^+(h'(x_o))(z).$

The right hand side is 0 using (1) which gives the conclusion.|

Note that $z \in - G$ is <u>not</u> necessary in this formulation.

It will now be demonstrated that McCormick's condition is subsumed by [1].

[3] Definition: Let x_o be a point satisfying the constraints of (P_1) and assume $g_1,\ldots,g_n,h_{n+1},\ldots h_p$ are twice differentiable continuously at x_o. The <u>second order qualification</u> (McCormick) holds at x_o if the following is true. Let y be any vector such that $(g_i'(x_o))(y) = 0$ for all $i \in B_o = \{i | g_i(x_o) = 0\}$ and such that

$(h_j'(x_o))(y) = 0 \quad j = n+1,\ldots,p$. Then $y = \alpha'(o)$ where $\alpha(\theta)$ is a twice continuously differentiable arc $(\theta \geq o)$ along which $g_i(\alpha(\theta)) = 0$ if $i \in B_o$ and $h_j(\alpha(\theta)) = 0$ if $\theta < \theta_o$ and with $\alpha(o) = x_o$.

[4] Theorem: (McCormick) if f, g_i and h_j $i = 1,\ldots,n$ and $j = n+1,\ldots,p$ are twice differentiable at x_o and the Kuhn-Tucker and Second order constraint conditions hold at x_o then a necessary condition for x_o to be a minimum for (P_1) is that there exist $u^+ = (u_1^+,\ldots,u_n^+)$, $z^+ = (z_{n+1}^+,\ldots,z_p^+)$ such that $u_i^+ \geq o$ and $u_i^+ g_i(x_o) = 0$ with

$$f'(x_o) + \sum_{i=1}^{n} u_i^+ \, g_i'(x_o) + \sum_{j=n+1}^{p} z_j^+ h_j'(x_o) = o$$

and such that for any y with $(g_i'(x_o))(y) = 0 \ \forall i$ such that $g_i(x_o) = 0$ and with $(h_j'(x_o))(y) = 0$, it follows that

$$\left[f''(x_o) + \sum_{i=1}^{n} u_i^+ (g_i''(x_o)) + \sum_{j=n+1}^{p} z_j^+ (h_j''(x_o)) \right](y,y) \geq o \ .$$

Proof: Set $C = R^n$, $B = \{x \mid x = (x_1,\ldots,x_n,o_{n+1},\ldots,o_p) x_i \leq o \ i = 1,\ldots,n\}$ and set $g = (g_1,\ldots,g_n \ h_{n+1},\ldots,h_p)$. It will now be shown that the conditions of [1] hold.

(1) The Kuhn-Tucker constraint condition being satisfied by γ can be rewritten as: $(g'(x_o))(y) \in P(B, g(x_o))$ implies $y = \lim_n n(\gamma(\frac{1}{n}) - \gamma(0))$ which implies (since $\gamma(\frac{1}{n}) \in A$) that $y \in P(A, x_o) = K$ and that G can be taken as R^n in [1](2).

(2) The second order condition can be written as:

if $(g'(x_o))(y) \in P(B, g(x_o)) \cap -P(B, g(x_o))$ then $y = \alpha'(o)$ with $\alpha(\theta)$ satisfying the second order qualification. Set

$x_n = \alpha(\frac{1}{n})$ and $\lambda_n = n$ then $h = \alpha'(o) = \lim_n n(x_n - x_o)$ and

$\alpha''(0) = z = \lim_n \; n\left[\alpha'(\frac{1}{n}) - \alpha'(o)\right]$.

Set $k_n = \alpha'(\frac{1}{n})$. Then $k_n = \lim_m \; -m\left[\alpha(\frac{1}{n} - \frac{1}{m}) - \alpha(\frac{1}{n})\right]$.

Since $\alpha(\theta)$ is contained in the constraint region $\alpha(\frac{1}{n})$ and $\alpha(\frac{1}{n} - \frac{1}{m}) \in A$ and $k_n \in -P(A,x_n)$.

(3) Moreoever, from the second order constraint condition one has

$$P(B,g(x_n)) \subset P(B,g(x_o)) \qquad n \geq n_o$$

as can be seen by examining components.

Since continuously differentiable mappings from R^m to R are Fréchet differentiable it only remains to verify (4). This is proved as a separate proposition. ▮

[5] Proposition: Let α be a twice continuously differentiable arc satisfying the second order qualification with $\alpha'(o) = h$ $\alpha''(o) = z$. Setting $\alpha(\frac{1}{n}) = x_n$, $\alpha'(\frac{1}{n}) = k_n$ and $\lambda_n = n$ [1](4) is satisfied.

Proof: $(f''(x_o))((h,h)) + (f'(x_o))(z) = \lim_n \lambda_n(f'(x_n) - f'(x_o))(h)$. Since h satisfies the constraint condition the above discussion shows that $(g'(x_o))(h) \in P(B,g(x_o) \cap - P(B,g(x_o))$. Because the Kuhn-Tucker condition holds

$$(f'(x_o))(h) = u^+(g'(x_o))(h) = o.$$

Let $\hat{g}(\theta) = f(\alpha(\theta))$. Then $\hat{g}: R \to R$ and has a local minimum at 0 since $\alpha(\theta) \in A$ for $\theta < \theta_o$. (See [6] (1).)

Thus

$$\hat{g}'(\theta) = (f'(\alpha(\theta)))(\alpha'(\theta)) \text{ and } \hat{g}'(o) = (f'(x_o))(h) = o$$

while

$$\hat{g}''(\theta)\Big|_o = (f'(\alpha(\theta)))(\alpha''(\theta))\Big|_o + (f''(\alpha(\theta)))(\alpha'(o), \alpha'(\theta))\Big|_o$$
$$= (f'(x_o))(z) + (f''(x_o))((h,h))$$

which must be non negative since 0 is a minimum and $\hat{g}'(o) = 0$. ▮

[6] Remarks: (1) It is not clear from the statement of the second order constraint condition that the promised arc lies inside the constraint region. An inspection of the definition shows that this follows from the continuity of the finite number of constraints $g_1, \ldots, g_n$. This observation is essential in the proof of the proposition in [5].

(2) McCormick shows that if the set E

$$E = \{g_i \mid i \in \{i \mid g_i(x_o) = 0\} \cup \{h_j \mid j = n+1, \ldots, p\}$$

is independent the second order condition holds. He gives examples to show that the two conditions are not strictly comparable.

[7] Suppose now that in (P) the objective function f is assumed to map X into Y and that x_o is a strong minimum with respect to a pointed S for (P). The following extension of [1] holds.

Theorem: Suppose that in the statement of [1] the following alterations are made.

(1)′ For the given cone G there is some $T \in P(B, g(x_o))^{-s}$ with

$$(f'(x_o) + Tg'(x_o))(h) \in S \quad \forall h \in G.$$

(4)′ $\overline{\lim}\,(f'(x_n))(\lambda_n k_n) \in S.$

Then a necessary condition for x_o to be a strong minimum for (P) is

$$(f''(x_o) + Tg''(x_o))((h,h)) \in S$$

$$\forall h \in G \cap -G \text{ with } (g'(x_o))(h) \in P(B, g(x_o)) \cap -P(B, g(x_o)).$$

Proof: With these changes the proof in [1] can be mirrored exactly. ∎

[8] Remarks: It would also be possible to phrase a version of [1] dealing with weak minima. This could be done by considering $s^+f'(x_o)$ rather than $f'(x_o)$ and by using the same trick as in [24]

of the last chapter to write the condition in operator form.

[9] Corollary: (to [1], [7] or [8])

If in condition (2) of [1] or [7] it is required that $P(A,x) \subset P(A,x_n)$ $n \geq n_o$ then (5), (5)' can be replaced by

$\lim \lambda_n f'((x_n))(h) \in S.$

Proof: In this case one can choose $k_n = h$, $z = 0$. ∎

Second Order Sufficiency Conditions in Reflexive Normed Spaces

Second order conditions which are sufficient for the existence of a minimum when the first order sufficiency conditions of chapter 5 are not met have been studied by McCormick (1967), Fiacco (1968), Guignard (1969) and Zlobec (1971). The results were all phrased in R^n or in finite dimensional normed spaces for real valued objective functions. The finite dimensionality was required to employ a compactness argument. By considering weak pseudotangent cones in reflexive normed spaces an infinite dimensional extension can be made. Two definitions are needed first.

[10] Definition: A point x_o is called an isolated intermediate (local) minimum for f over A with respect to a cone S if there is no sequence $\{x_n\} \subset A$ with $x_n \neq x_o$, $x_n \to x_o$ and such that $f(x_n) - f(x_o) \in -S$. Clearly such a minimum is a weak minimum if $S^o \neq \emptyset$.

[11] Definition: A point x_o will be said to have property (F) (with respect to S, f, A) if whenever there is a sequence $\{x_n\}$ in $A/\{x_o\}$ with $x_n \to x_o$, $f(x_n) \leq f(x_o)$ then there is a sequence x_n' such that $x_n' \to x_o$, $x_n' \in A/\{x_o\}$, $f(x_n') \leq f(x_o)$ and such that $\|x_n' - x_o\|^{-1}(x_n' - x_o)$ has a weakly convergent subsequence with limit $y_o \neq 0$.

Remark: (1) The substance of property (F) lies in the assertion that $y_o \neq 0$. This is because any bounded sequence in a reflexive normed space has a weakly convergent subsequence. This subsequence may, of course, have limit zero. This possibility is excluded for $\|x_n' - x_o\|^{-1}(x_n' - x_o)$ by property (F).

(2) If X is finite dimensional every point has property (F) since in

this case $\|x_n - x_o\|^{-1}(x_n - x_o)$ has a convergent subsequence k_n with $\|k_n\| = 1$. This means that the limit can not be 0.

(3) Clearly any isolated intermediate local minimum has property (F) since no such sequence $\{x_n\}$ can be found initially.

One is now ready to formulate the following extension of Guignard's sufficiency condition.

[12] Theorem: Suppose $f : X \to Y$, $g : X \to Z$ are twice continuously Fréchet differentiable at x_o and that x_o has property (F) with respect to a closed convex pointed cone S. Suppose the following conditions hold (where $A = g^{-1}(B) \cap C$):

(1) G is a closed convex cone such that if $x \in A$ and $\| x - x_o \| < \epsilon$ for some $\epsilon > 0$ then $x - x_o \in G$.

(2) There is some $s^+ \in S^+/\{o\}$ and $u^+ \in wP^+(B, g(x_o))$ with $s^+ f'(x_o) - u^+ g'(x_o) \in G^+$.

(3) $wP(B, g(x_o))$ is such that if $\|g(x) - g(x_o))\| < \epsilon$ for some given $\epsilon > 0$ then $g(x) - g(x_o) \in wP(B, g(x_o))$.

Then a sufficient condition for x_o to be an isolated local intermediate minimum for f over A w.r.t.S is the following:

For any nonzero element $h \in G$ such that either

(4) (i) $s^+(f'(x_o))(h) = 0$ and

$$(g'(x_o))(h) \in wP(B, g(x_o)) \cap -wP(B, g(x_o))$$

or

(4) (ii) $u^+(g'(x_o))(h) = 0$ and

$$(g'(x_o))(h) \in wP(B, g(x_o)) / - wP(B, g(x_o))$$

one has

$$s^+ f''(x_o))((h,h)) - u^+(g''(x_o))((h,h)) > 0.$$

Proof: Suppose by way of contradiction that there is a sequence $\{x_n\} \subset A / \{x_o\}$ with $x_n \to x_o$ and $f(x_n) \le f(x_o)$.

Let $k_n = \|x_n - x_o\|^{-1}(x_n - x_o)$. Since x_o is assumed to have property (F) there is another sequence with $k_n' = \|x_n' - x_o\|^{-1}(x_n' - x_o)$ such that $k_n' \rightharpoonup k_o \neq 0$. There is no loss of generality in assuming $x_n = x_n'$.

Since G is a closed convex cone satisfying (1) there is some n_o such that for $n \geq n_o$ $x_n - x_o \in G$ which in turn implies that $k_n \in G$. Since G is closed and convex G is weakly closed and $k_o \in G$. Clearly $k_o \in wP(A, x_o) \subset wP(\Delta, x_o)$.

Suppose that $s^+(f'(x_o))(k_o)$ is nonzero. By proposition [32] of chapter two $(g'(x_o))(k_o) \in wP(B, g(x_o))$ since $k_o \in wP(\Delta, x_o)$. Since $k_o \in G$, (2) produces $s^+(f'(x_o))(k_o) > 0$. By proposition [26] of chapter 2

$$(s^+f'(x_o))(k_o) = \lim \frac{s^+f(x_o + \|x_n - x_o\| k_n) - s^+f(x_o)}{\|x_n - x_o\|}.$$

For $n \geq n_1$ one has $s^+(f(x_n) - f(x_o)) > 0$ which since $s^+ \in S^+$ contradicts $f(x_n) \leqslant f(x_o)$. Hence $s^+(f'(x_o))(k_o) = 0$.

(i) Suppose now that $(g'(x_o))(k_o) \in wP(B, g(x_o)) \cap -wP(B, g(x_o))$ then since $s^+(f'(x_o))(k_o) = 0$ (4)(i) is satisfied.

(ii) Suppose next that $(g'(x_o))(k_o) \in wP(B, g(x_o)) / -wP(B, g(x_o))$.

If $u^+(g'(x_o))(k_o) > 0$, (2) produces $s^+(f'(x_o))(k_o) > 0$ which is again a contradiction. Thus $u^+(g'(x_o))(k_o) = 0$.

In either case (4)(i) or (4)(ii) then guarantees that since $k_o \neq 0$

(5) $s^+f''(x_o))(k_o, k_o)) - u^+ g''(x_o)((k_o, k_o)) > 0$.

Let $s^+f(x) - u^+g(x) = L(x)$. Taylor's theorem produces (since f,g are twice continously differentiable)

(6) $L(x_n) - L(x_o) = L'(x_o))(x_n - x_o) + \frac{1}{2}(L''(x_o))(x_n - x_o, x_n - x_o)) + o(\|x_n - x_o\|^2)$.

For $n \geq n_o$ $x_n - x_o \in G$ so that $(L'(x_o))(x_n - x_o) \geq 0$. Also, (5) produces $(L''(x_o))(k_o, k_o)) > 0$. Since $L''(x_o)$ is continuous, and hence weakly continuous, in each variable it follows that $(L''(x_o))(k_n,k_n)) > \epsilon_3$ if $n \geq n_2$, for some $\epsilon_3 > 0$. One can derive from (6), therefore, that

(7) $$\frac{L(x_n) - L(x_o)}{\|x_n - x_o\|^2} > 0 \qquad \text{if} \quad n \geq n_3.$$

This in turn produces

$$s^+(f(x_n) - f(x_o)) > u^+(g(x_n) - g(x_o)) \quad \text{if} \quad n \geq n_3.$$

From hypothesis (3) it is apparent (since g is continuous and $x_n \to x_o$) that $u^+(g(x_n) - g(x_o)) \geq 0$ if $n \geq n_4$.

Thus $s^+(f(x_n) - f(x_o)) > 0$ which again contradicts $s^+ \in S^+$ and $f(x_n) - f(x_o) \leq 0$. ▌

[13] Corollary: (Guignard) Suppose in the statement of [12] that f is real valued and X is finite dimensional and that all weak cones are replaced by corresponding strong cones, then one has the Guignard sufficiency condition for an isolated local minimum.

Proof: Since X is finite dimensional the unit ball is compact and $k_o \in P(A,x_o)$. Property (F) holds as was remarked in [11] and $(g'(x_o))(k_o) \in P(B,g(x_o))$. Since $wP(B,g(x_o))$ only enters the theorem through this relationship it can clearly be replaced by $P(B,g(x_o))$. The theorem is then included in [12] since, with $S=R^+$, s^+ can be taken to be 1 and since an isolated local minimum is an isolated intermediate minimum in the terminology of [10]. ▌

[14] Remarks: (1) It is clear that if $g'(x_o)$ is completely continuous, as is the case if Z is finite dimensional, $wP(B,g(x_o))$ can be replaced by $P(B,g(x_o))$ in [12].

(2) Guignard did not include (4)(ii) in her statement of [13]. The proof given in Guignard (1970) contains the erroneous assertion that if $(g'(x_o))(h) \in P(B,g(x_o))/-P(B,g(x_o))$ that $u^+(g'(x_o))(h)$ is positive which enables her to exclude the possibility of (ii).

Zlobec's asymptotic generalization (1970) did include (4)(ii) but he still gives Guignard's result as a corollary without it.

The next example shows that McCormick's sufficiency condition is included in [12] with $G = X$.

[15] Example: When one applies [12] to (P_1) one notices that if one replaces $L(x)$ by the reduced Langrangian,

$$L_1(x) = f(x) + \sum_{u_i^+>0} u_i^+ g_i(x) + \sum_{j=n+1}^{p} z_j^+ \; h_j(x) \; ,$$

that the proof method can be applied to $L_1(x)$ instead of $L(x)$.

In fact in this case $(g'(x_o))(k_o) \in P(B,g(x_o)) /- \; P(B,g(x_o))$ means that

$$(h'_j(x_o))(k_o) = 0 \quad j = n+1, \ldots, p \text{ and}$$

$$(g'_i(x_o))(k_o) < 0 \text{ for some } i_1 \text{ with } g_{i_1}(x_o) = 0.$$

so that $u^+ g'(x_o) = 0$ implies that $u^+_{i_1} = 0$.

Now

$$0 = f'((x_o)(k_o) = -\sum_{u_i^+>0} u_i^+ \; g'(x_o)(k_o)$$ so that from this point on in the proof it suffices to examine $L_1(x)$ and the sufficiency condition:

(1) If $h \neq 0 \in G$ and $(f'(x_o))(h) = 0$, $(h'_j(x_o))(h) = 0 \quad j=n+1,\ldots,p$ and $(g'_i(x_o))(h) = 0$ if $i \in \{i \mid u_i^+ > 0\}$

then

$$f''(x_o) + \sum_{i=1}^{n} u_i^+ \; (g''_i(x_o))(h,h)) + \sum_{j=n+1}^{p} z_j^+(h''_j(x_o))(h,h) > 0$$

is adequate to prove the result.

Note that in the general case there is no way of separating out inactive multipliers.

[16] <u>Example</u>: McCormick gives the following example to show the second order conditions can isolate behaviour that first order conditions do not.

Let $f(x,y) = (x-1)^2 + y^2$; $g(x,y) = x - {}^1/_k y^2$, $k > 0$.

Then $f'(0,0) = \begin{bmatrix} -2 \\ 0 \end{bmatrix}$, $g'(0,0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and

$$f'(0,0) + 2\, g'(0,0) \equiv \begin{bmatrix} 0 \\ 0 \end{bmatrix} \ \forall\ k > 0.$$

For $k = \frac{1}{4}$ $(0,0)^T$ is not a local minimum of (P_1) while for $k = 3$ it is. Thus the first order condition is inadequate for the purpose of locating the minima. In the second order condition

$$(1)\quad f''(0,0) + 2g''(0,0) = \begin{bmatrix} 2 & 0 \\ 0 & 2 - 4/k \end{bmatrix} = M.$$

It is easy to verify that all the conditions for the second order sufficiency and necessity theorems are met. The only one worth remarking is that since there is only one constraint Remark [6](2) guarantees the constraint qualification.

The sufficiency result then elicits that for those y for which $(g'(x_o))(y) = 0$ one must have $y^T M y > 0$. These y are all of the form $(0,a)^T$ and (1) implies that $2-4/_k > 0$ is sufficient for x_o to be a minimum. Thus $(0,0)^T$ is a minimum for (P_1) if $k > 2$.

The necessity condition shows similarly that if $k < 2$ $(0,0)^T$ is not a local minimum. For k=2 the necessary condition holds but the sufficiency condition does not.

Finally, the condition which is lacking for the application of the first order sufficiency condition is the pseudoconvexity of the constraint set. Since f is convex and G = X the other conditions are met. $A_k = \left\{(x,y) \mid y^2/_k \geq x\right\}$ is clearly not pseudo-convex at (0,0) since $(a_1,a_2) \in P(A_k,(0,0))$ implies that $a_1 \leq 0$ and hence $A_k \not\subset P(A_k,(0,0))$.

Asymptotic second order sufficiency:

[17] The next result extends Zlobec's asymptotic version of [13].

<u>Theorem</u>: Suppose in the statement of [12] that (1) and (3) hold and that (2) is replaced by (2′),

$$(2') \lim_i (s^+ f'(x_o) - u_i^+ g'(x_o))(g) = g^+(g) \cdot \forall g \in G$$

where $\{u_i^+\}$ is a sequence in $wP^+(B, g(x_o)), g^+ \in G^+$.

In addition

$$(5') \lim_i (s^+ f''(x_o) - u_i^+ g''(x_o)(\cdot\,,\,)) = L(\ ) \text{ exists for all } w \in \overline{W}$$

where

$$W = \left\{ w \,|\, w = (z - x_o) / \| z - x_o \| \,,\ z \in A,\ 0 < \| z - x_o \| < \epsilon \right\}$$

and $w_k \rightarrow h \Rightarrow L(w_k) \rightarrow L(h) \quad \forall\, w_k \in \overline{W}$,

and (4) becomes:

If $0 \neq h \in G$ is such that either

$$4'(\text{i}) \lim_i (u_i^+ g'(x_o))(h) = 0 \text{ and } (g'(x_o))(h) \in wP(B, g(x_o)) / -wP(B, g(x_o))$$

or

$$4'(\text{ii}) (f'(x_o))(h) = 0 \text{ and } (g'(x_o))(h) \in wP(B, g(x_o)) \cap -wP(B, g(x_o))$$

it then follows that

$$\lim (s^+ f''(x_o) - u_i^+ g''(x_o))(h,h)) > 0;$$

from which it follows that x_o is an intermediate local minimum.

<u>Proof</u>: As before one may assume (using (F)) that $\{x_n\} \subset A$

$\| x_n - x_o \|^{-1} (x_n - x_o) = k_n \rightarrow k_o \neq 0$ and with $x_n \rightarrow x_o$, $f(x_n) \leq f(x_o)$.

Then by (2′) rather than (2) since $k_o \in G$

$$(6) \quad \lim_i (s^+ f'(x_o) - u_i^+ g'(x_o))(k_o) \geq 0.$$

As before $(s^+ f'(x_o))(k_o) \leq 0$ and $(g'(x_o))(k_o) \in wP(B, g(x_o))$.

Hence if $(g'(x_o))(k_o) \in wP(B, g(x_o)) \cap - wP(B, g(x_o))$

then $u_i^+ (g'(x_o))(k_o) = 0$ and using (6)

$$s^+ (f'(x_o))(k_o) \geq 0$$

which means $s^+ (f'(x_o))(k_o) = 0$.

If $(g'(x_o))(k_o) \in wP(B,g(x_o)) / -wP(B,g(x_o))$ and $\lim_i u_i^+(g'(x_o))(k_o) > 0$ then, using (6) again, $s^+(f'(x_o))(k_o) > 0$ which is impossible. Again the sufficiency condition (4)′(i), (ii) must hold for k_o.

Setting $L_i(x) = s^+f(x) - u_i^+(g(x)$ and using Taylor's Theorem

$$\lim_n \lim_i \frac{L_i(x_k) - L_i(x_o)}{\| x_n - x_o \|^2} \geq \tfrac{1}{2} \lim_n \lim_i (L_i''(x_o)(k_n,k_n))$$

where the limit on the right exists because of (5′).

This limit is

$$\lim_i (L_i''(x_o))(k_o,k_o)) = L(k_o) > 0$$

since $k_n \rightarrow k_o$. As in [12] this sufficiency condition then implies that

$$\lim_i \left[L_i(x_n) - L_i(x_o)\right] > 0 \quad \text{if } n > n_o.$$

But $\lim_i u_i^+\left[g(x_n) - g(x_o)\right] \geq 0$ (since $u_i^+ \in wP^+(B,g(x_o))$) if $n \geq n_1$.

This means that for $n \geq \max(n_1,n_o)$

$s^+ f(x_n) - s^+f(x_o) > 0$ which contradicts the choice of $\{x_n\}$.|

A stronger condition but one which emphasises the operator nature of the conditions is given by the next result. It is stated for [12], but could have been stated for [17].

[18] Theorem: Suppose in [12] that (1), (3) continue to hold and that (2) is replaced by (2″)

(2″) $f'(x_o) - T g'(x_o) \in G^s$, $T \in wP(B,g(x_o))^s$.

A sufficient condition for a point x_o with property (F) to be a local intermediate minimum is:

For any non trivial $h \in G$ such that

(4″) (i) $(f'(x_o))(h) = 0$ and $(g'(x_o))(h) \in wP(B,g(x_o)) \cap - wP(B,g(x_o))$

or

(4″) (ii) $(Tg'(x_o))(h) = 0$ and $(g'(x_o))(h) \in wP(B,g(x_o)) / -wP(B,g(x_o)$

one has

$(f''(x_o))((h,h)) - T(g''(x_o))((h,h)) \in$ weak interior $(S) \neq \phi$.

<u>Proof:</u> With $\{x_n\}$ as in [12] suppose that $(f'(x_o))(k_o) \notin -S$.

Then

$$(5) \qquad \frac{f(x_n) - f(x_o)}{\|x_n - x_o\|} \longrightarrow (f'(x_o))(k_o) \in X/-S$$

since $k_n \longrightarrow k_o$. $X/-S$ is a positively homogenous weakly open set since $-S$ is a (weakly) closed convex cone. Hence (5) would imply that

$$f(x_n) - f(x_o) \in X/-S \quad \text{for } n \geq n_o$$

which contradicts $f(x_n) - f(x_o) \in -S$. So $(f'(x_o))(k_o) \in -S$.

(2″) gives $(f'(x_o))(k_o) \geq T(g'(x_o))(k_o)$

and $T(g'(x_o))(k_o) \geq 0$ since $(g'(x_o))(k_o) \in wP(B, g(x_o))$.

Since S is pointed

$$(f'(x_o))(k_o) \in S \cap -S = 0.$$

Thus either (4″)(i) or (4″)(ii) holds and

$$(f'(x_o))(k_o,k_o)) - T(g''(x_o))(k_o,k_o) \in \text{weakint } (S).$$

Setting $L(x) = f(x) - Tg(x)$ and applying the generalised Taylor Theorem (which can be done because L is twice continuously differentiable) one again sees that:

$$L(x_n) - L(x_o) = (L'(x_o))(x_n - x_o) + \tfrac{1}{2}(L''(x_o))(x_n - x_o, x_n - x_o)) + o(\|x_n - x_o\|^2).$$

(2″) gives $(L'(x_o))(x_n - x_o) \in S$. Since $L''(x_o)$ is continuous and hence weakly continuous

$$L''(x_o)\left(\frac{x_n - x_o}{\|x_n - x_o\|}, \frac{x_n - x_o}{\|x_n - x_o\|}\right) \longrightarrow L''(x_o)((k_o,k_o))$$

thus for $n \geq n_o$ and some weak neighbourhood N of 0

$$N + (L''(x_o))\left(\frac{x_n - x_o}{\|x_n - x_o\|}, \frac{x_n - x_o}{\|x_n - x_o\|}\right) \in \text{weakint } S$$

and

$$N + \frac{L(x_n) - L(x_o)}{\|x_n - x_o\|^2} \div \frac{o(\|x_n - x_o\|)^2}{\|x_n - x_o\|^2} \in \text{weakint } S.$$

Since any point of $o(\|x_n - x_o\|^2) / \|x_n - x_o\|^2$ belongs to any given neighbourhood N of 0 for $n \geq n_1$, one has for $n \geq n_2$ that

$$\frac{L(x_n) - L(x_o)}{\|x_n - x_o\|^2} \in \text{weakint } S.$$

Proceeding as before

$$f(x_n) - f(x_o) - T(g(x_n) - g(x_o)) \in \text{weakint } S.$$

and using (3) again since g is continuous and $x_n \rightarrow x_o$

$$g(x_n) - g(x_o) \in wP(B, g(x_o)) \text{ if } n \geqslant n_3.$$

Hence

$$f(x_n) - f(x_o) \in \text{weakint } S + S \subset \text{weakint } S.$$

Thus once again a contradiction has been reached.∎

[19] Remarks: (1) Again it is apparent that if $g'(x_o)$ is completely continuous that $wP(B, g(x_o))$ can be replaced by $P(B, g(x_o))$.

(2) If in addition $f''(x_o)$ is completely continuous S^o can be used instead of weakint(S). This follows because when $g'(x_o)$ is completely continuous $g''(x_o)$ is and so, therefore, is $L''(x_o)$. One would then have

$$(L''(x_o))(k_n, k_n)) \rightarrow (L''(x_o))(k_o, k_o)) \in S^o$$

which would suffice to deduce that

$$\frac{L(x_n) - L(x_o)}{\|x_n - x_o\|^2} \in S^o \qquad \text{if } n \geqslant n_o .$$

(3) Finally it seems worth mentioning that the two part sufficiency conditions in [12], [17], [18] could easily have been written in simpler form. This was not done to facilitate comparison with Guignard's and Zlobec's results.

Chapter Eight

OPTIMIZATION IN HILBERT SPACE AND VARIATIONAL INEQUALITIES FOR OPTIMIZATION

Section one: Optimization in Hilbert Space

In this section a variety of results specific to Hilbert space are discussed. The central advantages of Hilbert Space lie in the existence of a unique closest point to a closed convex set C and in the fact that dual cones can be more fully described since they lie in the space itself.

The first propositions give simpler equivalent constraints for $A(x) \in B$ when $A: X \to Z$ is a linear map between Hilbert spaces.

[1] Definition: A densely defined map A between two Hilbert spaces H_1 & H_2 is said to have pseudo-inverse $A^\dagger$ if there is a map satisfying $D(A^\dagger) = H_2$ and

(1) $R(A) \subset D(A^\dagger)$ $\qquad R(A^\dagger) \subset D(A)$

(2) $AA^\dagger = P_{\overline{R(A)}}$ $\qquad A^\dagger A = P_{\overline{R(A^\dagger)}}$.

A is often called the Moore-Penrose or generalised inverse. Such inverses have been studied by many mathematicians. The basic prop-erties used here are given in Charnes & Ben-Israel (1963) which includes an extensive bibliography. Extensions to non Hilbert spaces have been considered by Hille & Phillips (1957) among others.

[2] Proposition : (Charnes & Ben-Israel)

(1) Every closed densely defined linear operator has a unique closed psuedo-inverse.

(2) $N(A^*) = R(A)^\perp = N(A^\dagger)$.

(3) $R(A^\dagger) = N(A)$.

(4) $(A^\dagger)^\dagger = A$.

(5) If A^{-1} exists $A^\dagger = A^{-1}$.

Moreover, if A is bounded with R(A) closed, A is bounded and

(6) $A^\dagger AA^\dagger = A^\dagger \qquad AA^\dagger A = A$.

[3] <u>Proposition</u>: Suppose $A \in B[X,Z]$ with $R(A)$ closed then

$Ax \in D$ iff $x \in \hat{D}$

where $\hat{D} = A^{\dagger}(D \cap R(A)) \oplus N(A)$.

<u>Proof</u>: $\Rightarrow$ If $Ax \in D$ then $Ax \in D \cap R(A)$ and $A^{\dagger} Ax \in A^{\dagger}(D \cap R(A))$.

Using (2) of [1]

$$P_{\overline{R(A^{\dagger})}}\, x \in A^{\dagger}(D \cap R(A)).$$

Using (3) of [2]

$$P_{N(A)^{\perp}}\, x \in A^{\dagger}(D \cap R(A))$$

and since $P_{N(A)^{\perp}} = I - P_{\overline{N(A)}}$

$$x \in A^{\dagger}(D \cap R(A)) \oplus \overline{N(A)}.$$

Since A is continuous $\overline{N(A)} = N(A)$ and $x \in \hat{D}$.

Conversely if $x \in \hat{D}$

$$P_{\overline{R(A^{\dagger})}}\, x \in A^{\dagger}(D \cap R(A))$$

so that

$$AA^{\dagger} Ax \in AA^{\dagger}(D \cap R(A))$$

Using (6) of [2] at (2) of [1]

$$Ax \in P_{\overline{R(A)}}(D \cap R(A)) = D \cap R(A).$$ ∎

If $R(A)$ is not closed one still has that $Ax \in D$ implies

$$x \in A^{\dagger}(D \cap R(A)) \oplus N(A)$$

[4] <u>Definition</u>: $\mathbb{P}_A(x)$ will be used to denote the closed point to x in a closed convex set A and will be called the <u>projection</u> of x on A.

The fact that $\mathbb{P}_A$ is well defined is a standard result of Hilbert space convexity theory and is proved in Luenberger (1969). The use of the term projection is suggestive of linear projections on closed subspaces to which the notion in [4] reduces if A is a closed subspace.

[5] Proposition: Suppose $f:X\to Y$ is compactly differentiable at x_o and that $A:X\to Z$ is a continuous linear map with $R(A)$ closed. A necessary condition for x_o to be a strong minimum for f with respect to S subject to $Ax\in D$ is

$$(f'(x_o))\ (h)\in S \quad \forall h\in P(A^{\dagger}\ (D\cap R(A))\oplus N(A),x_o).$$

If f is pseudoconvex with respect to S at x_o, this is also sufficient.

Proof: Since $Ax\in D$ if and only if $x\in\hat{D}$ this is just [2.24] .∎

If f is supposed Fréchet differentiable then the cone P can be replaced by wP when $f'(x_o)$ is completely continuous.

[6] Proposition: Suppose that D is a convex set then if $f:X\to R$ a necessary condition for x_o to be a minimum for f subject to $Ax\in D$ is

$$\mathbb{P}_{T(\hat{D},x_o)}\ (-f'(x_o))=0$$

Again, if f is pseudoconvex at x_o, the condition is sufficient.

Proof: It is well known that the nearest point can be characterised by

(1) $(\mathbb{P}_A a_o-a_o,\ \ \mathbb{P}_A a_o-a)\leq 0 \quad \forall a\in A.$

Since D is supposed convex, $\hat{D}$ is convex and

$$P(\hat{D},x_o)\ = T(\hat{D},x_o).$$

The condition of [5] gives

(2) $(0-(-f'(x_o)),\ 0-t)\leq 0 \quad \forall t\in T(\hat{D},x_o)$

which on inspection of (1) is equivalent to

$$\mathbb{P}_{T(\hat{D},\ x_o)}\ (-f'(x_o))=0.\blacksquare$$

This result says in geometric terms that the closest point to the tangent cone for $-f'(x_o)$ is the vertex. In the last result it is apparent that convexity is only used to replace $P(D,x_o)$ or $wP(D,x_o)$ by $T(D,x_o)$.

This gives rise to the following generalisation.

[7] Proposition: Suppose $f : X \to R$ is (Fréchet) compactly differentiable. A necessary condition for x_o to be a minimum of min $f(x)$ subject to $Ax \in D$ with A pseudo-invertible is

$$\mathbb{P}_{P(D,x_o)}(-f'(x_o))=0$$

$$(\mathbb{P}_{W\ P(D,x_o)}(-f'(x_o))=0)\ .\ \blacksquare$$

Projections have been studied by McCormick & Tapia (1972) to provide gradient descent methods for solving min $f(x)$ subject to $x \in B$, a closed convex set. The most extensive investigation of projections has been made by Zarantonello (1971).

The result of [7] can be rewritten as follows.

[8] Proposition: If y is any point of X

$$\mathbb{P}_{P(\hat{D},x_o)}(y)=\mathbb{P}_{P(A^{\dagger}(D\cap R(A)),A^{\dagger}Ax_o)}(A^{\dagger}Ay)\oplus\mathbb{P}_{N(A)}(y)$$

Proof: $x_o \in A^{\dagger}Ax_o \oplus P_{N(A)}(x_o)$ since $R(A^{\dagger}A)$ is closed by definition.

(1) $P(\hat{D},x_o) = P(A^{\dagger}(D\cap R(A)),\ A^{\dagger}A(x_o))\oplus N(A)$.

It is clear that the right hand side contains theleft. Conversely, since $N(A)$ is a closed subspace,

$T(A^{\dagger}(D\cap R\ (A)),A^{\dagger}Ax_o)\oplus N(A)\subset P(\hat{D}\ ,x_o)$ which easily gives (1).

The definition of $\mathbb{P}$ implies that

$$\mathbb{P}_{P(\hat{D},x_o)}(y) = \mathbb{P}_{P(A^{\dagger}(D\cap R(A),A^{\dagger}Ax_o)}(A^{\dagger}Ay)\oplus P_{N(A)}y.\ \blacksquare$$

[9] Corollary: The condition of [7] is equivalent to

(1) $\mathbb{P}_{P(A^{\dagger}(D\cap R(A)),A^{\dagger}Ax_o)}(-A^{\dagger}Af'(x_o)) = 0$

(2) $\mathbb{P}_{N(A)}(f'(x_o)) = 0$. $\blacksquare$

If A is invertible (1) just becomes $\mathbb{P}_{P(A^{-1}(D),x_o)}(-f'(x_o)) = 0$ while (2) disappears. And if A is the zero mapping (1) collapses while (2) reduces to $f'(x_o) = 0$ which is the standard result for unconstrained minimization. ▌

It is apparent that when (P_1) is the minimization with respect to S given by (P_1) =min f(x) subject to $g(x) \in D$ with D closed and convex and g nonlinear there is not the same possibility of complete solution. It is, however, immediate on setting $G(x) = (I - \mathbb{P}_D)g(x)$ that (P_1) is equivalent to (P_1') $\min_S f(x)$ subject to $G(x) = 0$.

Unfortunately, G need not be differentiable even when g is. The following differential result does hold though.

[10] Proposition: Suppose $g: X \to Z$ is compactly differentiable at x_o, $d'G(x_o;h)$ exists and equals

$$(g'(x_o))(h) - \mathbb{P}_{T(D,g(x_o))}((g'(x))(h)) = \mathbb{P}_{T^-(D,g(x_o))}((g'(x_o))(h)).$$

Proof: $d^+G(x_o;h) = \lim_{t\to o^+} \frac{g(x_o+th)-g(x_o)}{t} - \lim_{t\to o^+} \frac{\mathbb{P}_D g(x_o+th) - \mathbb{P}_D g(x_o)}{t}$.

Now $\lim_{t\to o+} \frac{\mathbb{P}_D g(x_o+th) - \mathbb{P}_D g(x_o)}{t} = \lim_{t\to o+} t^{-1}\mathbb{P}_{(D-g(x))}(g(x_o+th)-g(x_o))$.

This last equality follows from

$$\mathbb{P}_{D-g(x_o)}(y-g(x_o)) = \mathbb{P}_D y - g(x_o) \text{ and } \mathbb{P}_D g(x_o) = g(x_o).$$

From the characterisation of $\mathbb{P}_A x$ one has

$$(\mathbb{P}_A x - \mathbb{P}_A y, x-y) = (\mathbb{P}_A x - \mathbb{P}_A y, x - \mathbb{P}_A x) + (\mathbb{P}_A y - \mathbb{P}_A x, y - \mathbb{P}_A y) +$$

$$\| \mathbb{P}_A x - \mathbb{P}_A y \|^2 \geq \| \mathbb{P}_A x - \mathbb{P}_A y \|^2$$

since the first two terms on the right are non negative by virtue of (1) $(\mathbb{P}_A z - t, z - \mathbb{P}_A z) \geq 0 \ \forall t \in A$. This now yields $\| x-y \| \geq \| \mathbb{P}_A x - \mathbb{P}_A y \|$.

In particular

(2) $$\| \mathbb{P}_{t^{-1}(D-g(x_o))} t^{-1}\left[g(x_o + h)-g(x_o)\right] - \mathbb{P}_{t^{-1}(D-g(x_o))}(g'(x_o))(h)\|$$
$$\leq \| \frac{g(x_o + th) - g(x_o)}{t} - (g'(x_o))(h) \| .$$

It is a consequence of (1) or the definition of $\mathbb{P}_A$ that

$$\mathbb{P}_{t^{-1}(D-g(x_o))} t^{-1}((g(x_o+th)-g(x_o)) = t^{-1}\mathbb{P}_{D-g(x_o)}(g(x_o+ th)-g(x_o)).$$

This was attributed to Rockafellar by McCormick & Tapia.

Rockafellar also noted that for $x \in Z$

(3) $$\mathbb{P}_{t^{-1}(D-g(x_o))}x \rightarrow \mathbb{P}_{T(D,g(x_o))}x \text{ as } t \rightarrow 0^+.$$

Combining (3) with $x = (g'(x_o))(h)$ and (2) one sees that

(4) $$\mathbb{P}_{T(D,g(x_o))}(g'(x_o))(h) = \lim P_{t^{-1}(D-g(x_o))} t^{-1}\left[g(x + th)-g(x_o)\right].$$

Thus

$$d^+G(x_o;h) = (g'(x_o))(h) - P_{T(D,g(x_o)}(g'(x_o)(h).$$

The final equality is a consequence of Zarantonello's result that for convex closed cones $I - \mathbb{P}_c = \mathbb{P}_{c^-}$ which is proved later.▮

[11] <u>Remark</u>: It is apparent from the proof of [10] that if only $d^+g(x_o; \ )$ exists, $d^+G(x_o;h) = \mathbb{P}_{T^-(D,g(x_o))}d^+g(x_o;h)$.

[12] An application of [10] to [6] produces the next proposition.

<u>Proposition</u>: Suppose D is closed and convex in [6] then a necessary condition for x_o to be a minimum for $f(x)$ subject to $Ax \in D$ is

$$\frac{d}{dt}\mathbb{P}_{\hat{D}}\ (x_o-tf'(x_o))\Big|_{t=o^+} = 0.$$

<u>Proof</u>: $\hat{D}$ is clearly convex since D is. Suppose now that $d'_n \in \hat{D}$ and $d'_n \rightarrow d'_o$. Then $d'_n \in A^\dagger(D\cap R(A)) \oplus N(A)$, $Ad'_n \in (D\cap R(A))$.

The continuity of A implies that $Ad'_o \in (D \cap R(A))$ and $A^{\dagger}Ad'_o \in A^{\dagger}(D \cap R(A))$ which in turn implies that $d_o \in A^{+}(D \cap R(A)) \oplus N(A) = \hat{D}$. Thus $\hat{D}$ is closed.

[10] with $g(x) = x$ provides that

$$\lim_{t \to 0^+} \frac{\mathbb{P}_{\hat{D}}(x_o - tf'(x_o)) - \mathbb{P}_D(x_o)}{t} = \mathbb{P}_{T(\hat{D},\, x_o)}(-f'(x_o)).$$

The last quantity is 0 from [6] and this is the desired result.∎

The differential condition of [10], while interesting, is not of much use in optimization because $d^+G(x\ ,\ )$ need not in general be convex, let alone linear. The only case in which d^+G will certainly be convex occurs when D is a closed subspace and then d^+G is in fact $(\mathbb{P}_D g)'$ and the optimization results can be applied directly. A necessary condition can be developed, though, using the notion of a projection (constraint) qualification.

[13] Definition: The projection qualification will be said to be satisfied at x_o if there is a closed convex cone G such that

$$\mathbb{P}_G\, \mathbb{P}_K = \mathbb{P}_{G \cap K} = \mathbb{P}_{P(A, x_o)},$$

where K and A retain their usual meanings. Zarantonello has shown that this implies that $G \cap K = P(A, x_o)$.

The following propositions on projections are necessary to the first order conditions.

[14] Proposition: (Zarantonello) If C is a closed convex cone in a Hilbert space then any point x can be expressed uniquely as

$$x = x_1 + x_2\ ,\ x_1 \in C,$$
$$x_2 \in C^- \quad \text{and} \quad x_1 = \mathbb{P}_C x\ ,\ x_2 = \mathbb{P}_{C^-} x.∎$$

Proof: From the projection inequality

(1) $(x-\mathbb{P}_C x, c-\mathbb{P}_C x) \leq 0 \quad \forall c \in C.$

Since C is closed convex cone, 0 and $2\,\mathbb{P}_C x_o$ belong to C.

Thus

$$(x-\mathbb{P}_C x, \mathbb{P}_C x) = 0.$$

Moreover, $C + \mathbb{P}_C x \subset C$ and hence (1) implies $x-\mathbb{P}_C x \in C^-$.

Now let $x = x_1 + x_2$, $x_1 \in C$, $x_2 \in C^-$, $(x_1, x_2) = 0$.

Then

$$(x-x_1, c-x_1) = (x_2, c-x_1) = (x_2, c) \leq 0 \quad \forall c \in C$$

and $x_1 = \mathbb{P}_C x$, similarly $x_2 = \mathbb{P}_{C^-} x$.

Since $x = \mathbb{P}_C x + (x-\mathbb{P}_C x)$ and $x-\mathbb{P}_C x \in C^-$ and $(\mathbb{P}_C x, x-P_C x) = 0$

this means that $x-\mathbb{P}_C x = \mathbb{P}_{C^-}(x)$. ▮ Note that this includes:

$\mathbb{P}_C x = 0$ if and only if $x \in C^-$.

[15] Proposition: (Zarantello) Let C_1, C_2 be closed convex cones.

Suppose $\mathbb{P}_{C_1}\mathbb{P}_{C_2} = \mathbb{P}_{C_2}\mathbb{P}_{C_1}$ then

(1) $\mathbb{P}_{C_1 \cap C_2} x = \mathbb{P}_{C_1}\mathbb{P}_{C_2} x$ if and only if

(2) $(x-\mathbb{P}_{C_1}\mathbb{P}_{C_2} x, \mathbb{P}_{C_1}\mathbb{P}_{C_2} x) = 0.$ ▮

It is immediate that [15] holds for any x if one of C_1, C_2 is the whole space or if both C_1, C_2 are closed subspaces. Zarantonello has proved in addition that if C_1, C_2 are finte dimensional closed convex cones then $\mathbb{P}_{C_1 \cap C_2} = \mathbb{P}_{C_1}\mathbb{P}_{C_2}$ whenever C_1, C_2 commute.

[16] Proposition: Let $H = \{h \mid h = u^+ g'(x),\ u^+ P^+(B, g(x_o))\}$,

$$K = \{k \mid (g'(x_o))(k) \in P(B, g(x_o))\}.$$

Then $K^+ = \overline{H}$.

Proof: That $K^+ \subset \overline{H}$ is a consequence of [31] of chapter five.

Suppose, conversely, that $h \in H$. Then $h = u^+ g'(x_o)$,

$u^+ \in P^+(B,g(x_o))$.

For any $k \in K$, $(g'(x_o))(k) \in P(B,g(x_o))$ by definition.

Thus $(h,k) = u^+(g'(x_o))(k) \geq 0$ and since h was arbitrary $H \subset K^+$.

Since K is a closed cone and $K^+ \subset \overline{H}$ one has $K^+ = \overline{H}$.|

[17] The cone $G = P(A,x_o)$ will always satisfy the constraint qualification since

$$\mathbb{P}_{P(A,x_o)} \mathbb{P}_K = \mathbb{P}_{P(A,x_o) \cap K} = \mathbb{P}_K \mathbb{P}_{P(A,x_o)} = \mathbb{P}_{P(A,x_o)}$$ as can be easily verified. For the purpose of proving a necessary condition it would suffice to require that

$\mathbb{P}_{P(A,x_o)}(-f'(x_o)) = \mathbb{P}_G \mathbb{P}_K(-f'(x_o))$.

This would be satisfied by the cone $G = (\mathbb{P}_K(-f'(x_o)))^-$ but, since for this cone G one has $\mathbb{P}_G \mathbb{P}_K(-f'(x_o)) = 0$ whether $\mathbb{P}_{P(A,x_o)}(-f'(x_o)) = 0$ or not, it would fail to discriminate and would have no chance of giving any useful necessary condition. For this reason the constraint qualification of [13] is used since in many cases it is also sufficient.

[18] <u>Theorem</u>: Suppose X is a Hilbert Space and $f:X \to R$, $g:X \to Z$ are compactly differentiable at a point x_o. Suppose that G is a closed convex cone satisfying the projection constraint qualification at x_o. A necessary condition for x_o to minimize

$(P) = \min f(x)$ subject to $g(x) \in B$, $x \in C$ is given by

$\lim_i f'(x_o) - u^+_i g'(x_o) \in G^+, u^+_i \in P^+(B,g(x_o))$, where

$\lim_i u^+_i g'(x_o) = \mathbb{P}_{\overline{H}} f'(x_o)$.

<u>Proof</u>: The result of proposition [7] gives $\mathbb{P}_{P(A,x_o)}(-f'(x_o)) = 0$. The projection qualification gives $\mathbb{P}_G \mathbb{P}_K(-f'(x_o)) = 0$ or $\mathbb{P}_K(-f'(x_o)) \in - G^+$.

This is equivalent to $f'(x_o) + (I-\mathbb{P}_K)\,(-f'(x_o)) \in G^+$.

Using [14] and [16] $I - \mathbb{P}_K = \mathbb{P}_{-\overline{H}}$.

Moreover; $\mathbb{P}_{-\overline{H}} -f'(x_o) = -\mathbb{P}_{\overline{H}} f'(x_o)$

so that $f'(x_o) - \mathbb{P}_{\overline{H}} f'(x_o) \in G^+$.

This last equation is equivalent to the claimed result.▮

As a special case of [18] one has

[19] Proposition: Suppose that H is closed and that

$$g'(x_o)^{-1}\,[P(B,g(x_o))] = P(g^{-1}(B),x_o),$$

then a necessary condition for min $f(x)$ subject to $g(x) \in B$ is (1) $f'(x_o)-u^+g'(x_o) = 0$ or (2) $u^+g'(x_o) = \mathbb{P}_H f'(x_o)$ and $f'(x_o) \in H$.

Proof: Since $H = \overline{H}$ $\mathbb{P}_H x = \mathbb{P}_{\overline{H}}\, x$. Since $K = P(A,x_o) = P(A,x_o)$, G can be taken to be X and the result follows.▮

[20] Theorem: (Sufficiency) Let $f:X \to R$ $g:X \to Z$ be differentiable at x_o. Suppose (1) $\mathbb{P}_G\, \mathbb{P}_K(-f'(x_o)) = \mathbb{P}_{P(A,x_o)}\,(-f'(x_o))$ and (2) A is pseudoconvex at x_o and f is pseudoconvex at x_o over A then $f'(x_o) - \mathbb{P}_{\overline{H}} f'(x_o) \in G^+$ is sufficient for x_o to be a minimum for (P).

Proof: Working back through the proof of [18]

$$\mathbb{P}_G\, \mathbb{P}_K(-f'(x_o)) = 0,$$

Using (1) $\mathbb{P}_{P(A,x_o)}(-f'(x_o)) = 0$ Equivalently $f'(x_o) \in P^+(A,x_o)$.

The result now follows from the standard sufficiency argument.▮

[21] Remarks:

(1) [18] and [19] can equally well be framed for bounded

differentiation and weak pseudotangent cones.

(2) By using $s^+f'(x_o)$, $s^+\in S^+/\{0\}$ instead of f' the previous results can be adjusted to deal with (P) when $f:X\to Y$ and x_o is a weak minimum with respect to S.

(3) If $P(A,x_o) = 0$ then G can be chosen to be K^-. This covers the case , for example, when C is a discrete set and provides an example in which G can be considerably larger than $P(A,x_o)$.

An alternative formulation of a first order necessary condition but one which appears to avoid restraint qualification is given by the next theorem.

[22] Proposition: Suppose x_o is a minimum for (P) then a necessary condition is given by:

$$f'(x_o) - \mathbb{P}_{\overline{H}}f'(x_o)\in \mathbb{P}_{-K}(P(A,x_o)).$$

Proof:
$$\begin{aligned} f'(x_o) - \mathbb{P}_{\overline{H}}f'(x_o) &= (I - \mathbb{P}_{\overline{H}})(f'(x_o)) \\ &= \mathbb{P}_{(-H^+)}(f'(x_o)) \\ &= \mathbb{P}_{-K}(f'(x_o)). \end{aligned}$$

Then, since x_o is a minimum, $f'(x_o)\in P^+(A,x_o)$ and $\quad f'(x) - \mathbb{P}_{\overline{H}}f'(x_o)\in \mathbb{P}_{-K}(P^+(A,x_o))$. ∎

[23] Theorem: Let $H = \overline{H}$ and $G = [\mathbb{P}_{-K}(P^+(A,x_o))]^+$. Then a necessary condition is

$$f'(x_o) - u^+g'(x_o)\in G^+;\ u^+g'(x_o) = \mathbb{P}_H f'(x_o).$$

If (1) $\mathbb{P}_{-K}P^+(A,x_o)\subset P^+(A,x_o)$,

(2) A is pseudoconvex at x_o,

(3) f is pseudoconvex over A at x_o; then the condition is sufficient as well.

Proof: Necessity follows from $H = \overline{H}$ and

$$\left[\overline{\mathbb{P}_{-K}P^+(A,x_o)}\right] = (\mathbb{P}_{-K}\,P^+(A,x_o))^{++} = G^+,$$

whence $f'(x_o) - u^+g'(x_o) \in \mathbb{P}_{-K}P^+(A,x_o) \subset G^+$.

Sufficiency is proved by noting that with $G = \overline{[\mathbb{P}_{-K}P^+(A,x)]}^+$

$$f'(x_o) - u^+g'(x_o) \in G^+$$

so that by (1) $f'(x_o) - u^+g'(x_o) \in P^+(A,x_o)$

and hence $\quad (f'(x_o))(h) \geqslant u^+((g'(x_o))(h)) \quad \forall h \in P(A,x_o)$.

Since A is pseudoconvex $(f'(x_o))(x-x_o) \geqslant u^+(g'(x_o))(x-x_o) \geqslant 0$,

$$\forall x \in A.$$

Which, since f is pseudoconvex implies that x_o is a minimum for f over A.▌

[24] Remarks:

(1) Condition (1) $\mathbb{P}_{-K}P^+(A,x_o) \subset P^+(A,x_o)$ is essentially Guignard's sufficiency condition that $A-x_o \subset G$ because A is pseudoconvex by (2).

(2) It is only in cases in which G does satisfy some sort of constraint condition that [22] is at all useful. This constraint condition might actually be $\mathbb{P}_{-K}P^+(A,x_o) \subset P^+(A,x_o)$ which is certainly met if $K = P(A,x_o)$.

(3) McCormick and Tapia give an explicit characterisation of $\mathbb{P}_C$ when C is what they call a positive cone with respect to an orthogonal set in X. That is $C = \{x \in X \mid (x, \delta_\alpha) \geqslant d_\alpha \quad \alpha \in A\}$ and $\{\delta_\alpha\}$ is orthogonal in X. In the case that K is positive this allows an explicit statement of [22] to be made.

The final Hilbert space result uses a series of results by Zarantonello (1971) on spectral mappings with respect to non linear projections.

[25] Definition: $J:X \to X$ is said to be a spectral mapping with respect

to a spectral resolution $\{\mathbb{P}_\lambda\}$ if J is a mapping which can be written as

$$J = \int_{-\infty}^{\infty} \lambda d \; \mathbb{P}_\lambda$$

where (roughly) the meaning of integration is analogous to that in the standard (linear) projection theory. Zarantonello has developed a complete theory of such spectral integrals with respect to projections on convex cones.

[26] Theorem: Suppose $f: X \to R$ is compactly differentiable and $J: X \to X$ is spectral with respect to a spectral resolution $\{\mathbb{P}_\lambda\} = \{\mathbb{P}_{D_\lambda}\}$. Suppose that D is a closed convex cone such that $D^-_\lambda = D$ for some D_λ in the resolution. Then a necessary condition for x_o to minimize f(x) subject to $Jx \in D$ is given by

$$\mathbb{P}_{T(\partial T^*(0),\ x_o))}(-f'(x_o)) = 0$$

where $T(x) = \frac{1}{2}(\mathbb{P}_{D^-}(x), J(x))$ and ∂T^* is the subgradient of the convex conjugate of T.

Proof: $J(x) \in D$ if and only if $(I - \mathbb{P}_D)(J(x)) = 0$

Using [14] $J(x) \in D$ if and only if $\mathbb{P}_{D^-}(J(x)) = 0$.

By Lemma 9.1 of Zarantonello (1971) $\mathbb{P}_{D^-}J(x)$ is a spectral map since J is, and by Theorem 9.8 of Zarantonello

(1) $\mathbb{P}_{D^-}(J(x)) = \frac{d}{dx} \frac{1}{2}(\mathbb{P}_D x, J(x)) = \frac{d}{dx} \; T(x).$

It is reasonably simple to verify that so defined is convex from X to R. In fact any spectral map is the gradient of a convex map.

Now $0 = \mathbb{P}_{D^-}J(x))$ if and only if $0 = \frac{d}{dx} T(x) = \partial T(x)$

and $0 \in \partial T(x) \Leftrightarrow x \in \partial T^*(0)$ by a theorem of Rockafellar (1966).

Hence, using (1),

$$0 \in \mathbb{P}_{D^-}J(x) \text{ if and only if } x \in \partial T^*(0).$$

Since $\partial T^*(0)$ is a convex set and since the problem is equivalent to

$$\min f(x) \text{ subject to } x \in \partial T^*(0),$$

[6] gives the desired necessary condition.∎

[27] Remarks:

(1) Spectral resolutions containing any given P_C can be simply constructed, as Zarantonello indicates.

(2) Again sufficiency is guaranteed by the pseudoconvexity of f at x_o. The constraint set is necessarily pseudoconvex at x_o since $\partial T^*(0)$ is convex.

Section two: Variational Inequalities

Suppose that X is a convex space and $f:X \to R$ is lower semicontinuous and convex. Let $T:X \to X'$ be a given mapping. Let (,) be the associated bilinear form.

[28] Definition: An inequality of the form

$$(1) \quad (Tu, u-v) \leq f(v)-f(u) \quad u,v \in \text{domf}$$

is called a variational inequality and is said to have solution u_o if $\quad (Tu_o, u_o-v) \leqslant f(v)-f(u_o) \quad \forall v \in \text{domf}$.

The study of such abstract variational inequalities is well developed in the work of Browder, Stampacchia, Lions and others. The primary motivation for the study has come from partial differential equation theory, but as Browder has remarked (1966,a) there are very close correspondences with optimization theory. Kosca (1969) has shown that any variational inequality can be considered as one in which f is just the indicator of a closed convex set.

It will be seen from (1) that a solution to the variational inequality is equivalent to $0 \in R(T + \partial f)$. Thus any theory which guarantees solutions to operator equations $T_1(u) = 0$ also can be invoked in the variational context. Such theorems are usually somewhat less constructive than the corresponding direct proof of solutions to (1) but are generally much more immediate. In this section a brief survey of results from monotone operator theory is made and these results are then applied to two optimization problems. For the remainder of this discussion X is a real Banach space with dual X'.

[29] Definition: (Browder & Hess (1971)) A mapping (multivalued) T from X into X' is said to be generalised pseudo-monotone if the following holds:

For any sequences $\{u_j\}$ in X and $\{w_j\}$ in X' with $w_j \in Tu_j$ $u_j \rightharpoonup u_o$ $w_j \rightharpoonup w_o$ such that $\limsup (w_j, u_j - u_o) \leq 0$; $w_o \in Tu_o$ and $(w_j, u_j) \to (w_o, u_o)$.

[30] Definition: $T: X \to X'$ is said to be monotone if

$$(w-v, x-y) \geq 0 \qquad \forall x, y \in D(T) \;\; \forall w \in Tx, \forall v \in Ty.$$

[31] Definition: T_1 is said to be maximal monotone if $G(T_1) \subset (X, X')$ is maximal among the graphs of monotone maps $G(T)$; or equivalently if $(z-w, u-v) \geq 0 \quad \forall\; u \in D(T)$ and $\forall z \in Tu$ implies $v \in D(T)$ and $w \in Tv$.

Browder and Hess have proved that any maximal monotone map is generalised pseudomonotone so in particular any subgradient map is generalised pseudomonotone. Maximal monotonicity of such maps was proved in Rockafellar (1970).

[32] Definition: Let $T:X \to X'$ be a multivalued map. T is said to be quasibounded if $\forall M > 0$ there exist $K(M) > 0$ such that if $w \in Tu$ and $(w,u) \leqslant M \|u\|$, $\|u\| \leqslant M$ then $\|w\| \leqslant K(M)$.

[33] Definition: T is strongly quasibounded if for each $M > 0$ there exists $K(M) > 0$ such that if $w \in Tu$ and $(w,u) \leqslant M$, $\|u\| \leqslant M$ then $\|w\| \leqslant K(M)$.

In addition to bounded maps strongly quasibounded maps include those maximal monotone maps which have $0 \in D(T)^{\circ}$ (Rockafellar (1969)).

[34] Definition: T is coercive if there is a map $c: R \to R^{+}$ with $\lim_{r \to \infty} c(r) = \infty$ such that $(w,u) \geqslant c(\|u\|)\|u\|$ $\forall (u,w) \in G(T)$.
With these definitions one can state the following theorem of Browder and Hess.

[35] Theorem: Let X be a reflexive Banach space and T be a maximal monotone map from X into $2^{X'}$ with $0 \in D(T)$. Let T_o be generalised pseudomonotone and coercive with the property (2) that T_o is regular, that is $R(T_o + T_2) = X'$ for any bounded, everywhere defined, single-valued maximal monotone mapping T_2. Suppose further that T_o is quasibounded or T is strongly quasibounded: then $R(T_o + T) = X'$. ∎

From the previous discussion it is clear that this includes the result $0 \in R(T_o + \overline{w}_o + \partial f)$, $\overline{w}_o \in X'$ when the conditions on T hold with $T = \partial f$. For these maps, therefore, one has a solution $w_o \in Tu_o$ to

(3) $\quad (w_o - \overline{w}_o, u_o - v) \leqslant f(v) - f(u_o) \quad \forall v \in \text{dom} f.$

[36] If in fact one only wishes to solve (3) with $\overline{w}_o = 0$ the coercivity conditions can be weakened in many situations. For example, in the case that T_o is quasibounded and $0 \in T(0)$ the condition: $\exists M_1 \ni (T_o u,u) > 0$ if $\|u\| > M_1$, suffices as can be seen by inspecting the proofs in Browder and Hess.

Many alterations are possible in the type of theorem that can be proved. Browder and Hess (1971) and Mosca (1969) provide more than enough variations to indicate the depth of the subject. Mosca's paper which deals with approximations of inequalities provides conditions under which monotone mappings, which are not necessarily maximal, can be used.

With this brief discussion behind one can turn to the use of variational inequalities in optimization. The most immediate example is provided by the notion of the subgradient itself. As has already been pointed out any variational inequality can be considered as a statement that the subgradient of f contains vectors of certain forms. Within this framework it seems worth noting the following theorem which relies on a result of Rockafellar (1970,c) that:

[37] Proposition: If $K(u,v)$ is a convex-concave semicontinuous saddle function then ∂K is maximal monotone where ∂K denotes $\partial K(u,v) = (\partial_u K, -\partial_v -K)$. This last notation is used to denote the standard subgradients of $K(\ ,v)$ and $-K(u, \ )$ both of which are convex functions.

[38] Theorem: Let $X = (Y,Z)$ where Y and Z are reflexive Banach Spaces. Let $K(y,z)$ be a convex-concave semicontinuous saddle function on X into R. Let T_o be a coercive generalised pseudomonotone map

satisfying (2) of [35] . Suppose that either

(1) T_o is quasibounded and $(0,0) \in D(\partial K)$, or

(2) ∂K is strongly quasibounded (i.e. $(0,0) \in D^o(\partial K)$).

Then given any $(y'_o, z'_o) \in X'$ there exist $(y_o, z_o) \in D(\partial K)$ simultaneously solving

(3a) $K(y,z_o) - K(y_o,z_o) \geq (y'_o - \pi_1 T_o(y_o,z_o), y - y_o)$

(3b) $K(y_o,z_o) - K(y_o,z) \leq (z'_o - \pi_2 T_o(y_o,z_o), z - z_o)$

for all $(y,z) \in (Y,Z)$ where π_1, π_2, are the projections of X' on Y', Z' respectively.

<u>Proof</u>: (3a) can be rewritten as $y'_0 - \pi_1 T_o \in \partial_{y_o} K(y_o,z_o)$ and (3b) as $-(z'_o - \pi_2 T_o) \in \partial_{z_o} -K(y_o,z_o)$, or using [37] $(y'_o, z'_o) - (\pi_1 T_o(x_o,y_o), \pi_2 T_o(x_o,y_o)) \in \partial K(y_o,z_o)$. This in turn is equivalent to requiring that $R(T_o + \partial K) = X'$. By [37] K is maximal monotone. An application of the Theorem of [35] gives the desired results. ∎

Note that $(0,0) \in D(\partial K)$ if and only if K has a saddlepoint at some point (y_o,z_o) since the existence of such a saddle point is equivalent to

$$K(y,z_o) - K(y_o,z_o) \geq (y-y_o,0)$$
$$-K(y_o,z) - (-K(y_o,z_o)) \leq (z-z_o,0)$$

which in turn says

$$0 \in \partial K(y,z_o)\big|\, y_o \qquad 0 \in \partial -K(y_o,z)\big|\, z_o$$

or

$$(0,0) \in \partial K(y_o,z_o).$$

It is simple to verify that if $T_o(y,z) = (T_1 y, T_2 z)$ and T_1, T_2 are both coercive, strongly quasibounded generalised pseudo-monotone then T_o has these properties. It is clear that [38] includes the standard variational inequality as the case Z = 0 in which case (3b) is vacuously satisfied.

Theorem [38] might be said to give solutions to a pair of coupled variational inequalities with the coupling taking place T_o.

The next results return to simple variational inequalities and to a discussion of the complementarity problem. Suppose that C is a closed convex cone in a Banach Space and that $T:X \rightarrow X'$ is a single valued mapping which will subsequently be required to be of various monotone types. The complementarity problem is defined to be: (P) minimize (Tu,u) subject to $Tu \in C^+$, $u \in C$.

This problem finds its origins in linear programming when given an n vector b and an n×n matrix M one wants to find $y, x \in R^n$ such that $y = Mx + b$ and $(y_i, x_i) = 0$, $i = 1,.,n$.

Karamardian (1969) showed that when $C \subset R^n$ and $T = f$ was continuous and satisfied $(f(u)-f(v),u-v) \geq k \| u-v \|^2$ on C that the minimum in (P) is 0.

Bazaraa (et al.) (1972) have showed that 0 is the minimum when T is bounded hemicontinuous and satisfies $C \subset D(T)$ and $(Tu-Tv,u-v) \geq \alpha(\| u-v \|) \| u-v \|$ for some strictly increasing α with $\lim_{r \to \infty} \alpha(r) = \infty$. This property is called <u>α-monotonicity</u>.

Their proof relies on variational inequalities proved in Mosca (1969) concerning perturbations of mappings satisfying the various conditions listed above. In the case that T is everywhere defined, such a T is maximal monotone and their result is included in the following results.

Generally, one has:

[39] <u>Proposition</u>: The solution to (P) is 0 if and only if there is a solution to the variational inequality $(Tv,u-v) \geq 0 \ \forall u \in C$.

<u>Proof</u>: Note first that setting $f_C(u) = \begin{cases} 0 & u \in C \\ \infty & u \notin C \end{cases}$ this has the

form of [28] (1). Suppose u_o is a variational solution, then $(Tu_o,u_o) \leqslant (Tu_o,u) \quad \forall\, u \in C$. Since $0 \in C$, $(Tu_o,u_o) \leqslant 0$. For any $\lambda > 0$ $\lambda u_o \in C$ since C is a cone.

Thus $\forall \lambda > 0$

$$(Tu_o,u_o) \leqslant \lambda(Tu_o,u_o)$$

which is impossible unless $(Tu_o,u_o) = 0$. This in turn implies that $(Tu_o,u) \geqslant 0 \quad \forall\, u \in C$ or equivalently that $Tu_o \in C^+$. Conversely if $(Tu_o,u_o) = 0$, $Tu_o \in C^+$, $u_o \in C$ then $(Tu_o,u-u_o) = (Tu_o,u) \geqslant 0 \quad \forall\, u \in C$ since $Tu_o \in C^+$.

Finally since $u_o \in C$, u_o is actually a solution to the variational inequality. ∎

A solution to $(Tu_o,u_o) = 0$ must be a minimum for (P) because $Tu \in C^+$, $u \in C$ implies $(Tu,u) \geqslant 0$. The solution will be unique if one has $(Tu-Tv,u-v) > 0$, $u \neq v$, since with $u_o, u_1 \in C, Tu_o, Tu_1 \in C^+$

$$(Tu_o,u_o) = (Tu_1,u_1) = 0 \text{ implies}$$
$$(Tu_1-Tu_o,u_1-u_o) = -(Tu_1,u_o) - (Tu_o,u_1) \leqslant 0.$$

Thus one has, using the remark of [36]:

[40] <u>Theorem</u>: (P) has solution u_o with $(Tu_o,u_o) = 0$ whenever T is a quasibounded generalised pseudomonotone mapping which is regular and satisfies $(Tu,u) \leqslant 0$ if $\|u\| \leqslant M$.

Browder and Hess (1972) remark that for a monotone map, regularity is equivalent to maximality when $0 \in D(T)$ so that [40] includes all maximal monotone maps with $0 \in D(T)$.

If one requires a stronger property than generalised pseudomonotonicity then regularity is not needed in [40]. The definition is initially due to Brezis (1968).

[41] Definition: A multivalued mapping $T:X \rightarrow X'$ is said to be pseudomonotone on $C \subset D(T)$ when

(a) Tu is a nonempty closed convex subset of X if $u \in C$.

(b) T is upper semicontinuous as a multivalued mapping from $C \cap F$ into X with the weak topology for any finite dimensional subspace F.

(c) Whenever $\{u_j\} \subset C$ $u_j \rightharpoonup u_o, w_j \in Tu_j$ then $\limsup(w_j, u_j - u) \leqslant 0$ implies that for each $v \in C$ $\exists\, w(v) \in Tu$ with $\liminf(w_j, u_j - v) \geqslant (w(v), u-v)$.

[42] Definition: A single valued function $T:X \rightarrow X'$ is said to be demicontinuous if $x_n \rightarrow x$ implies $Tx_n \rightharpoonup Tx$ while T is said to be hemicontinuous, if

$$T(tx + (1-t)\,x_o) \rightharpoonup Tx \quad \text{when} \quad t \rightarrow 1.$$

Proposition: A single valued hemicontinuous monotone mapping with $C \subset D(T)$ is pseudomonotone on C if (1) $D(T)$ is open or (2) if $D(T)$ is a dense subspace and T is locally bounded.

Proof: (a) is immediate since T is single valued.

(b) Kato (1967)(1964) has shown that hemicontinuity implies demicontinuity for a class of sets $D(T)$ which include open sets and includes dense subspaces when T is locally bounded. Demicontinuity is clearly stronger than the continuity property of [41] (b).

(c) Let $\{u_j\} \subset C$ with $u_j \rightharpoonup u$ and suppose $\overline{\lim}(Tu_j, u_j - u) \leqslant 0$. By monotonicity and $C \subset D(T)$, $(Tu, u_j - u) \leqslant (Tu_j, u_j - u)$ which means that

$$(Tu_j, u_j - u) \rightarrow 0.$$

Let x be any point in C. Then

$$(Tu_j, u_j - x) = (Tu_j, u_j - u) + (Tu_j, u - x)$$

so that

$$\liminf(Tu_j, u_j - x) = \liminf\ (Tu_j, u-x)$$

Also $(Tx, u_j - x) \leqslant (Tu_j, u_j - x)$ so that

$$(Tx, u-x) \leqslant \liminf\ (Tu_j, u-x).$$

Let $v \in C$ and let $x_t = tv + (1-t)u$ which since C is convex is in $C \subset D(T)$. Setting $x_t = x$

$$(Tx_t, u-x_t) \leqslant \liminf\ (Tu_j, u-x_t).$$

Since T is hemicontinuous

$$\lim_{t \to 0} Tx_t = Tu$$

and

$$(Tu, u-v) \leqslant \liminf\ (Tu_j, u-v) = \lim\ (Tu_j, u_j - v)$$

and T is pseudoconvex on C. ▮

For pseudomonotone maps one has the following theorem of Browder and Hess, which they prove directly.

[43] Theorem: Let X be a reflexive Banach space, and C a closed convex subset of X with T pseudomonotone on C. Then if

(1) $w \in Tu$, and $(w,u) \leqslant 0$ implies $\|u\| \leqslant M$ for some $M > 0$,

there is a solution $u_0 \in C$, $w_0 \in Tu$ to

$$(w_0, u-u_0) \geqslant 0 \qquad \forall\, u \in C.$$

Proof: Browder and Hess prove this theorem for T coercive and assert that $(w_0 - \overline{w}, u-u_0) \geqslant 0 \quad \forall u \in C$ has solution for any $\overline{w} \in X'$. Inspection shows that their proof holds for the present theorem. ▮ Note, conversely, that coerciveness of T would provide that $T-\overline{w}$ satisfied the hypothesis (1) for all $\overline{w} \in X'$.

Using [42], [39] one has as a corollary

[44] Corollary: There is a solution to

$$(Tu_0, u_0) = 0, \quad u_0 \in C, \quad Tu_0 \in C^+$$

whenever T is hemicontinuous single valued monotone satisfying

$$(1) \quad (Tu,u) > 0 \quad \text{if} \quad \|u\| > M_1$$

and either (2)(a) $C \subset D(T)$ and $D(T)$ is open or (2)(b) T is locally bounded with D(T) a dense subspace.

[45] Remarks:

(1) Kato in fact shows (1964) that D(T) need only be "quasi dense" in (2)(b) by which he means for each $u \in D(T)$ there is a dense subset M_u in X such that for each $v \in M_u$ $u + tv \in D$ if $0 < t < \epsilon(v)$. In this circumstance hemicontinuity and local boundedness still imply demicontinuity.

(2) Any α-monotone T satisfies the condition in (1) so that for D(T) quasidense [44] includes the result of Bazaraa et al.

[46] The variational inequalities theorems discussed above are all valid for T multivalued so that they all provide solutions to

$$(P')\, 0 \in (Tu_o, u_o) \quad Tu_o \cap C^+ \neq \emptyset \ , \quad u_o \in C$$

which might be considered as the multivalued nonlinear complementary problem.

Bazaraa et al. note that their method, which as was already noted relies on Mosca's approximation theory and which doesn't appear to have direct extension to non monotone mappings, also gives information on approximation, perturbation and continuous dependence of solutions. The perturbation theory for the variational inequalities discussed above is buried in the operator analysis. The penultimate results of this section use a perturbation argument directly to establish the existence of a solution to the complementarity problem for generalised pseudomonotone mappings which are demicontinuous and satisfy another weak monotone type condition.

[47] Definition: (Hess(1972)). A single valued map $T X \rightarrow X'$ is said to be of type (P) if whenever $u_n \rightarrow u_o$ then $\limsup(Tu_n, u_n - u_o) \geqslant 0$.

[48] Definition: $J: X \rightarrow X'$ is called the duality map and is defined by

$$Ju = \{ v \in X' \mid (u,v) = \|u\| \, \|v\| \; , \; \|v\| = \|u\| \}.$$

It is known (Hess(1972)) that every reflexive Banach space X has an equivalent norm in which both X and X' are locally uniformly convex and that with these equivalent norms J is single valued and demicontinuous.

[49] Theorem: Let X be a reflexive Banach space and let $T: X \rightarrow X'$ be a demicontinuous type (P) mapping then

$T_\lambda = T + \lambda J$ is pseudomonotone on any closed set contained in $D(T)$ when X, X' have the locally uniformly convex norms mentioned above.

Proof: J is single valued demicontinuous so that T_λ is also $\forall \lambda > 0$.

J is known to have the property that whenever $u_j \rightharpoonup u_o$ and $\liminf(Ju_j, u_j - u_o) \leqslant 0$ then $u_j \rightarrow u_o$. As Hess remarks it is straightforward to verify that T_λ shares this property. This in turn allows one to verify that T_λ is pseudomonotone on any closed $C \subset D(T)$. ▮

[50] Theorem: Suppose X is a reflexive Banach Space and that

$T:X \to X'$ satisfies the following properties for a closed convex cone C.

(1) T is generalized pseudomonotone on C.

(2) T is type (P) on C.

(3) T is demicontinuous on C.

(4) T is quasibounded on C.

(5) If $(Tu,u) \leq 0$ and $u \in C$ then $\|u\| \leq M$. Then there is a solution u_o to the complementarity problem with $(Tu_o,u_o) = 0$.

Proof: There is no loss in assuming that X and X' are locally uniformly convex since all the hypothese are invariant under equivalent norms as is the conclusion. [49] then implies that T_λ is pseudomonotone on C and coercive. Using [43] there is a solution $u_\lambda \in C$ to

$$(T_\lambda u_\lambda, u_\lambda) = 0 \qquad u_\lambda \in C,\ T_\lambda u_\lambda \in C^+.$$

Now $\quad (T_\lambda u_\lambda, u_\lambda) = (Tu_\lambda, u_\lambda) + \lambda(Ju_\lambda, u_\lambda) = (Tu_\lambda, u_\lambda) + \|u_\lambda\|^2$

so that $\quad \forall \lambda > 0$

$$(T u_\lambda, u_\lambda) \leq 0.$$

Property (5) implies that $\{u_\lambda\}$ is bounded set in C.

Since X is reflexive there is some sequence $\lambda_n \to 0$ with $u_n = u_{\lambda_n} \rightharpoonup u_o$. u_o belongs to C because C is convex and closed and hence weakly closed. Using $(Tu_n,u_n) \leq 0$ and $\|u_n\| \leq M$ one has, by (4), $\|Tu_n\| \leq K(M)$ so that for some subsequence which will not be relabeled $Tu_n \rightharpoonup w_o$.

Since $\quad T_{\lambda_n} u_n = Tu_n + \lambda_n J u_n \in C^+$

and $\quad \| T_{\lambda_n} u_n - Tu_n \| = \|\lambda_n J u_n\| = \lambda_n \|u_n\| \leq \lambda_n M$

$$\| T_{\lambda_n} u_n - Tu_n \| \to 0 \text{ and since } Tu_n \rightharpoonup w_o$$

$$T_{\lambda_n} u_n \rightharpoonup w_o \in C^+ \text{ as again } C^+ \text{ is weakly closed.}$$

Now $\quad (T_{\lambda_n} u_n, u_n - u_o) \leqslant 0$

because $u_o \in C$ and $(T_{\lambda_n} u_n, u_n) = 0$. Thus one has

$$(Tu_n, u_n - u_o) + \lambda_n (Ju_n, u_n - u_o) \leq 0 .$$

This in turn implies that

$$\liminf (Tu_n, u_n - u_o) \leqslant 0.$$

because $\lambda_n \to 0$ and $\|(Ju_n, u_n - u_0)\| \leqslant \|u_n\| \; \|u_n - u_0\| \leqslant M_1$.

Since T is generalised pseudomonotone and $u_n \rightharpoonup u_o$, $Tu_n \rightharpoonup w_o$ one may conclude that $w_o \in Tu_o$ and that

$$(Tu_n, u_n) \to (Tu_o, u_o).$$

Finally $(Tu_o, u_o) \geq 0$ because $u_o \in C$ and $Tu_o \in C^+$, but because $(Tu_n, u_n) \leqslant 0$ one has actually that

$$(Tu_o, u_o) = 0 \quad u_o \in C, \quad Tu_o \in C^+. \blacksquare$$

[51] Remarks:

(1) This result with $D(T) = X$ is contained in [35] because one can show that a demicontinuous everywhere defined type (P) mapping satisfying the condition $(Tu, u) \geqslant -k \|u\|$ is regular.

(2) For [50] to add any new result it must be ascertained that mappings can satisfy [50] without being pseudomonotone. This would appear to be possible since one can envisage examples in which $u \underset{j}{\rightharpoonup} u_o$ but $\{Tu_j\}$ is unbounded. Note that it is only in this case that the requirement that T be type (P) on C is not implied by generalised pseudomonotonicity on C.

(3) Conditions (4) and (5) could be combined into the weaker condition that when $u \in C$ and $(Tu, u) \leqslant 0$ one has $\|u\| \leqslant M$ and $\|Tu\| \leqslant K(M)$. Trivially any strictly monotone map with

$T(0) = 0$ satisfies this condition.

It seems worth emphasising that any continuous finite dimensional mapping is pseudomonotone so that [43] holds for all these mappings. Karamardian (1972) has shown that the complementarity problem is solved in R^n for any continuous f satisfying $(x,f(x)) = 0$ when $x \in R^n \cap \{x \mid \|x\| = M\}$ while [39], [43] include any continuous map satisfying $(x,f(x)) > 0$ when $x \in C$ and $\|x\| > M$.

Since lower semicontinuous convex functions have maximal monotone subgradients one might ask if weaker convex type functions have associated with them any monotone type operators. In any case where, for instance, the derivative satisfies some monotonicity requirement one immediately has the whole of monotone operator theory as an adjunct for proving optimization results. A tentative start in this direction is provided by the following two results.

[52] Proposition: Let X be a reflexive Banach space and let $f:X \to R$ be a compactly differentiable quasiconvex mapping. Suppose that $f'(x)$ is bounded as a function of x and is completely continuous at any local minimum. Then $f'(x)$ is type (P).

Proof: Suppose $x_n \rightharpoonup x_o$. In the case that x_o is a local minimum

$$f'(x_n) \to f'(x_o) \text{ by hypothesis so that}$$

$$\lim(f'(x_n),x_n-x_o) = 0.$$

Suppose now that x_o is not a local minimum. Since f is continuous and quasiconvex f is lower semicontinuous in the weak topology and $\liminf f(x_n) \geq f(x_o)$. Equivalently,

if $\quad n \geq n_o, \qquad f(x_o) - \epsilon_n \leq f(x_n)$

where $\{\epsilon_n\}$ is a sequence of positive numbers converging to zero. Because x_o is not a local minimum one can find, for $\epsilon_n < \epsilon_{n_o}$, a sequence $\{y_n\}$ with $y_n \to x_o$ and such that for $n \geq n_1$

$$f(x_n) \geq f(x_o) - \epsilon_n \geq f(y_n).$$

Since f is quasiconvex and differentiable one has for $n \geq n_1$

$$(f'(x_n), x_n - y_n) \geq 0.$$

This in turn leads to

$$(f'(x_n), x_n - x_o) = (f'(x_n), x_n - y_n) + (f'(x_n, y_n - x_o) \geq (f'(x_n), y_n - x_o).$$

This last term converges to zero because $y_n \to x_o$ and since $\{f'(x_n)\}$ being the image of a bounded set is bounded by hypothesis. Thus $\liminf (f'(x_n), x_n - x_o) \geq 0$ and f is type (P). ∎

As an application of this result one can prove that solutions exist for the following kind of variational inequality.

$$(Tu_o . u - u_o) \geq 0 \quad \text{if} \quad f(u) \leq f(u_o).$$

[53] Theorem: Let X be a reflexive Banach Space and suppose $f: X \to R$ is quasiconvex with a demicontinuous type (P) derivative. Let $T: X \to X'$ be an everywhere defined hemicontinuous monotone mapping satisfying $(T u - T v, u - v) \geq B(\|u-v\|)\|u-v\|$ for some strictly increasing B with $B(0) = 0$; $B(\infty) = \infty$. Suppose that $(f'(x), x) \geq -k\|x\|$ for some $k > 0$ and $\|x\| > M$, then there is a solution to

$$(Tu_o - \overline{w}, u - u_o) \geq 0 \;\; \forall\, u \in L(u_o) = \{u \mid f(u) \leq f(u_o)\}.$$

Proof: Let $T_1 = T + f'$. Since T is hemicontinuous monotone with $D(T) = X$, T is demicontinuous and T_1 is demicontinuous. Suppose $x_n \rightharpoonup x_o$ and $\limsup (T_1 x, x_n - x_o) \leq 0$. Since f' is assumed type (P) one must have $\limsup (Tx_n, x_n - x_o) \leq 0$.

But $(Tx_n, x_n - x_o) \geq (Tx_o, x_n - x_o) + B(\| x_n - x_o \|)(\| x_n - x_o \|)$ and hence $\limsup B(\| x_n - x_o \|) \| x_n - x_o \| = 0$. Suppose that $x_n \not\to x_o$ then for some subsequence $\{x'_n\}$ $\| x'_n - x_o \| \geq \epsilon > 0$ which in turn means that $\limsup B(\| x_n - x_o \|) \| x_n - x_o \| \geq B(\epsilon)\epsilon > 0$. Thus $x_n \to x_o$ and as in previous results T_1 must be pseudomonotone. Since T is B- monotone and $(f'(x), x) \geq -k\|x\|$, one has

$$(T_1 u, u) \geq (T\, o, u) + (f'(u), u) + B(\| u \|) \| u \|$$

$\geq (B(\| u \|) - \| To \| - k) \| u \|$ if $\|u\| \geq M$ and T_1 is coercive.

Applying [43] with C = X one sees that there is a solution to

$$T_1 u_o = \vec{w} \qquad \forall\, \vec{w} \in X'.$$

Let $u \in L(u_o)$ then $(Tu_o - w, u - u_o) = -(f'(u_o), u - u_o) \geq 0$ since f is quasiconvex. ∎

The condition $(f'(u), u) \geq -k \| u \|$ is always guaranteed if there is a global minimum for f at o as one then has $(f'(u), u - 0) \geq 0 \quad \forall\, u \in \{u \mid f(u) \geq f(0)\} = X$. It is clear that the condition can be satisfied without any such minimum existing.

In more general terms the theorem opens up the question of when one can find a solution $u_o \in C(u_o)$

$(Tu_o, u - u_o) \geq 0 \quad \forall\, u \in C(u_o)$. (That is, when $C(u)$ is a convex set which varies with u.)

The last remarks of this chapter convern the solution of generalised variational inequalities. Specifically, suppose $T: X \to B[X, Y]$ and that S is a closed convex cone with interior in X. One can ask for solutions to

(1) $(Tu_o)(u - u_o) \not\in -S^o \quad \forall\, u \in C$ or to

(2) $(Tu_o)(u - u_o) \in S \quad \forall\, u \in C$.

Conditions for solution of (1) are easily obtained but (2), which appears more interesting, also appears much less tractable.

A particular example of conditions for (1) to be solvable is given by requiring that for some nonzero $u^+ \in S^+$, u^+T satisfies the conditions of [43]. Stronger results can be proved by using the natural generalisations of the concepts of this section.

BIBLIOGRAPHY

Bibliography

1. ABADIE, J. (ed.), (1967), Nonlinear programming, North Holland Co., Amsterdam.

2. BAIR, J., JONGMAN, F., (1972), La séperation vraie dans in espace véctoriel, Bull. Soc. Roy. Sc., Liége, vol.1.41, 163-170.

3. BAZARAA, M.S., GOODE, J.J., SHETTY, C.M., (1971), Optimality criteria in nonlinear programming without differentiability, Operations Research, vol.19, 77-86.

4. BAZARAA, M.S., GOODE, J.J., NASHED, M.Z., (1972), A nonlinear complementary problem in mathematical programming in Banach space, Proc. A.M.S. 35, vol.1, 165-170.

5. BEN-ISRAEL, A., CHARNES, A., (1963), Contributions to the theory of generalised inverses, SIAM J. Appl. Math., vol.11, 667-699.

6. BERGE, C., (1959), Espaces topologiques, fonctions multivoque, Paris, Dunod.

7. BREZIS, H., (1968), Équations et inéquations non-linéaires dans les espace véctoriels en dualité, Ann. Inst., Fourier (Grenoble), vol.18, 115-176.

8. BROWDER, F.E., (1966,a), On the unification of the calculus of variations and the theory of monotone nonlinear operations in Banach spaces, Proc. N.A.Sc., U.S.A., vol.56, 419-425.

9. (1966,b), Problémes non-linéaires, Univ. of Montreal Press.

10. (1968), The fixed point theory of multivalued mappings in topological vector spaces, Math. Annalen, vol.177, 283-301.

11. (1969), Nonlinear variational inequalities and maximal monotone mappings in Banach spaces, Math. Annalen, vol.183, 213-231.

12. BROWDER, F.E., HESS, P., (1972), Nonlinear mappings of monotone type in Banach spaces. J.Funct. Anal., vol.11, 251-294.

13. COTTLE, R., FÉRLAND, J., (1971), On pseudoconvex functions of non negative variables, Math. Prog., vol.1, 95-101.

14. CRAVEN, B.D., (1970), A generalization of Lagrange multipliers, Bull. Aust. Math. Soc., vol.3, 353-362.

15. (1972), Nonlinear programming in locally convex spaces, J.O.T.A., vol.10, 197-210.

16. (Submitted), Converse duality in Banach space.

17. (Preprint), Mathematical programming and Pontryagin's principle.

18. CRAVEN, B.D., MOND, B., (1971), On converse duality in nonlinear

19. (1972), Converse and symmetric duality in complex nonlinear programming, J. Math. Ann and Appl., vol.37, 617-626.

20. (1973), Transposition theorems for cone-convex functions, SIAM J. Appl. Math., vol.24, 603-612.

21. DANIEL, J.W., (1971), The approximate minimization of functionals, Prentice Hall Series in Automatic Computation, New Jersey.

22. DEMPSTER, M.A.H., (1971), Elements of optimization, (preliminary copy), Balliol College, Oxford.

23. DUNFORD, N., SCHWARTZ, J.T., (1966), Linear operators, part 1, Interscience Publishers, New York.

24. EGGLESTON, H.G., (1969), Convexity, Cambridge University Press.

25. FIACCO, A.V., (1968), Second order sufficient conditions for weak and strict constrained minima, SIAM J. Appl. Math., vol.16, 105-108.

26. FLETT, T.M., (1967), On differentiation in normed vector spaces, J. London.Math. Soc., vol. 42, 523-533.

27. GUIGNARD, M., (1969), Generalized Kuhn-Tucker conditions for mathematical programming problems in Banach space, SIAM J. Control, vol.7, 232-241.

28. HALKIN, H., (1972,a), Aninverse function theorem for continuous functions which are assumed to be differentiable at one point only, C.O.R.E. Discussion Paper No. 7201.

29. (1972,b), Implicit functions and optimization problems without continuous differentiability of the data, C.O.R.E. Discussion Paper no. 7206.

30. HESS, P., (1971), A variational approach to a class of nonlinear eigenvalue problems, Proc. A.M.S., vol.77, 272-276.

31. (1972), On nonlinear mappings of monotone type homotopic to odd operators, J. Funct. Anal., vol.11, 138-167.

32. HILLE, E., PHILLIPS, R.S., (1957), Functional analysis and semigroups, A.M.S. Colloq., vol.31 (revised edition), New York.

33. JOHN, F., (1948), Extremum problems with inequalities as subsidiary conditions, Studies and Essays, Courant Anniversary Volume, New York, 187-204.

34. KARAMARDIAN, S., (1969), The nonlinear complementarity problem with applications, Articles I and II, J.O.T.A., vol.4, 87-98, 167-181.

35. (1972), The complementarity problem, Math. Prog., vol.2, 107-129.

36. KATO, T., (1969), Demicontinuity, hemicontinuity and monotonicity I, Bull. A.M.S., vol.70, 548-550.

37. (1967), Demicontinuity, hemicontinuity and monotonicity II, Bull. A.M.S., vol.73, 886-889.

38. KELLEY, J.L., (1964), General topology, D. van Nostrand Co. Inc., New York.

39. KELLEY, J.L., NAMIOKA, I., (1963), Linear topological spaces, D. van Nostrand Co. Inc., New York.

40. KUHN, H.W., TUCKER, A.W., (1952), Nonlinear programming, Proc. 2nd. Berkeley Symp. on Math. Statistics and Probability, vol.5, Univ. of California Press, Berkeley, 481-492.

41. LUENBERGER, D.G., (1968), Quasiconvex programming, SIAM J. Appl. Math., vol.16, 1090-1095.

42. (1969), Optimization by vector space methods, John Wiley, New York.

43. MANGASARIAN, O.L., (1969), Nonlinear programming, McGraw-Hill Series in Systems Science, New York.

44. MANGASARIAN, O.L., FROMOVITZ, S., (1967), The Fritz John necessary optimality conditions in the presence of equality and inequality constraints, J. Math. Ann. and Appl., vol.17, 37-47.

45. MARTOS, B., (1971), Quadratic programming with a quasiconvex objective function, Operations Research, vol.19, 87-97.

46. MASSAM, H., (1973), Mathematical programming in locally convex topological spaces, M.Sc. Thesis, McGill, Montreal.

47. McCORMICK, G.P., (1967), Second order conditions for constrained minima, SIAM J. Appl. Math., vol.15, 641-652.

48. McCORMICK, G.P., TAPIA, R.A., (1972), The gradient projection method under mild differentiability conditions, SIAM J. Control, vol.10, 93-98.

49. MINTY, G.J., (1962), Monotone (nonlinear) operators in Hilbert space, Duke Math. J., vol.29, 341-346.

50. MOSCO, U., (1969), Convergence of convex sets and of solutions of variational inequalities, Advances in Math., vol.3, 510-585.

51. NAGAHISA, Y., SAKAWA, Y., (1969), Nonlinear programming in Banach spaces, J.O.T.A., vol.4, 182-190.

52. NASHED, M. Z., (1971), Differentiability and related properties of nonlinear operators: some aspects of the role of differentials in nonlinear functional analysis. In: Nonlinear Functional Analysis and Applications, L.B.Rall, (ed.), Academic Press, New York, 103-301.

53. NORRIS, D.O., (1971), A generalized Lagrange multiplier rule for equality constraints in normed linear spaces, SIAM J. Control, vol.9, 561-567.

54. PONSTEIN, J., (1967), Seven kinds of convexity, SIAM Review, vol.2, 115-119.

55. RISSANEN, J., (1967), On duality without convexity, J. Math. Ann.and

56. RITTER, K., (1969), Optimization in linear spaces I, Math. Annalen, vol.182, 189-206.

57. (1969,b), Optimization in linear spaces II, Math. Annalen, vol.183, 169-180.

58. (1970), Optimization in linear spaces III, Math. Annalen, vol.184, 133-154.

59. ROBERTSON, A.P., ROBERTSON, W.J., (1964), Topological vector spaces, Cambridge University Press.

60. ROCKAFELLAR, R.T., (1966), Characterization of the subdifferentials of convex functions, Pac. J. Math., vol.17, 497-510.

61. (1969), Local boundedness of nonlinear monotone operators, Mich. Math. J., vol.16, 397-407.

62. (1970,a), Convex analysis, Princeton University Press.

63. (1970,b), On the maximality of sums of nonlinear maximal monotone operators, Trans. A.M.S., vol.14, 75-88.

64. (1970,c.), Monotone operators associated with saddle functions and minimax problems. In: Proceedings of Symposia in Pure Mathematics, vol.18, part I, F.E. Browder (ed.), 241-250.

65. ROSEN, J.B., MANGASARIAN, O.L., RITTER, K., (1970), Nonlinear programming, Academic Press, Wisconsin.

66. SION, M., (1958), On general minimax theorems, Pac. J. Math., vol.8, 171-176.

67. STOER, J., WITZGALL, C., (1970), Convexity and optimization in finite dimensions, part I, Springer Verlag, Berlin.

68. TAPIA, R.A., (1971), Integration and differentiation of nonlinear operators. In: Non Linear Functional Analysis and Applications, L.B. Rall (ed.), Academic Press, New York.

69. TAYLOR, A., (1958), Introduction to functional analysis, John Wiley, New York.

70. TURNOVEC, F., (1970), Generalized saddle points and dual equilibrium problems, C.O.R.E. Discussion Paper no. 7026.

71. VAINBERG, M.M., (1964), Variational methods for the study of nonlinear operators, Holden Day Inc., London.

72. VAJADA, S., (1974), Theory of linear and nonlinear programming, Longman Group Ltd., London.

73. VALENTINE, F.A., (1964), Convex sets, McGraw-Hill, New York.

74. VARAIYA, P.P.,(1967), Nonlinear programming in Banach spaces, SIAM J. Appl. Math., vol.19, 239-244.

75. WONG, Y.-C., NG, K.-F., (1973), Partially ordered topological vector spaces, Oxford Mathematical Monographs, Oxford University Press.

76. ZARANTONELLO, E.H., (1971), Projections on convex sets in Hilbert space and spectral theory. In: Contributions to Nonlinear Functional Analysis, E.H. Zarantonello (ed.), Academic Press, New York, 234-424.

77. (1973), Dense single-valuedness of monotone operators, Israel J. Math., vol. , 158-166.

78. ZLOBEC, S., (1970), Asymptotic Kuhn-Tucker conditions for mathematical programming problems in a Banach space, SIAM J. Control, vol.8, 505-512.

79. ZLOBEC, S., MASSAM, H., (1973), Various definitions of the derivative in mathematical programming. Presented at VIII International Symposium on Mathematical Programming, Stanford.