CONTRIBUTION TO DUBOVITSKIY AND MILYUTIN'S OPTIMIZATION FORMALISM

Ludmila Rigby
Department of Computing and Control
Imperial College of Science and Technology
London SW7 2BZ
United Kingdom

ABSTRACT

This paper is a contribution to the unified approach of Halkin, Neustadt, Gamkrelidze and others to the theory of necessary conditions for general optimization problems.

The basic problem is formulated in terms of real linear topological spaces, mappings between them and a partial ordering determined by a proper convex cone. It includes, therefore, problems with both scalar- and vector-valued optimality criteria.

Optimality conditions are developed in terms of Gâteaux and Fréchet differentials of given mappings and linear continuous functionals on the spaces concerned, making use of the Dubovitskiy and Milyutin's formalism.

INTRODUCTION

We develop necessary and sufficient optimality conditions for a Mathematical Programming Problem, employing the Dubovitskiy and Milyutin's formalism [1,2]. This is built round a form of the separation principle for finite families of convex sets with empty intersection, each set corresponding to an approximation to a constraint in the problem.

The Mathematical Programming Problem is of considerable generality and covers a wide range of applications. We take the underlying set to be a real linear topological space and allow for vector-valued objective function as well as for infinitely many equality and inequality constraints.

Our findings differ in minor respects from known results in the literature, principally over weakening of differentiability requirements on the constraint and objective function in obtaining the necessary conditions. We feel though that the main contribution is in presenting a simple, complete proof of the results.

1. FORMULATION OF THE BASIC PROBLEM

Take X, Y_1, Y_2, Y_3 real linear topological spaces, mappings $F : X \to Y_1$, $G : X \to Y_2$, $H : X \to Y_3$, set $Q \subseteq X$, proper convex cones $C \subseteq Y_1$, $S \subseteq Y_2$ with non-empty interiors, and a fixed element $e \in Y_3$. The set Ω of *feasible points* is defined as follows:

$$\Omega = \{x\in X; x\in Q, G(x)\in S, H(x)=e\}.$$

We introduce the Basic Problem:

Basic Problem

Find $x \in \Omega$ such that

$$\{z\in\Omega; F(z)-F(x)\in C\setminus\{\theta\}\} = \phi.$$

Every such element is called *optimal*. (Here, and subsequently, θ denotes the zero element in the space concerned.)

This will be recognised as a mathematical programming problem over a real linear topological space with multivalued objective function in the presence of constraints, of a form similar to that studied in [3].

The set Q comprises the implicit constraints, while the set $\{z\in X; H(z)=e, G(z)\in S\}$ defines the explicit (equality and inequality) constraints. We make no assumptions concerning finite dimensionality of the ranges of the constraint functions.

2. BASIC THEORY

Let X be a real linear topological space. (Throughout this paper we shall suppose all topologies to be Hausdorff.)

The real linear space of all linear continuous functionals on X is denoted by X^*.

2.1 We begin with definitions of "conical approximation" and "polar cone", as these are the two basic concepts in formulating and deriving the necessary conditions for optimality.

A list of useful properties is given below. Verification of these properties is routine and so no proofs are given.

Let Q be a subset of X, $x \in X$ and N a base of neighbourhoods of zero. The following sets

$$K(x,Q) = \{z\in X;\ \exists\,\varepsilon>0\ \forall\,\alpha\in(0,\varepsilon)\ \ x+\alpha z\in Q\}$$

$$C(x,Q) = \{z\in X;\ \forall\,\varepsilon>0\ \ \exists\,\alpha\in(0,\varepsilon)\ \ x+\alpha z\in Q\}$$

$$N(x,Q) = \{z\in X;\ \exists\,\varepsilon>0\ \ \exists\,U\in\mathcal{N}\ \forall\,\alpha\in(0,\varepsilon)\ \forall\,\omega\in z+U\ \ x+\alpha\omega\in Q\}$$

$$M(x,Q) = \{z\in X;\ \forall\,\varepsilon>0\ \forall\,U\in\mathcal{N}\ \exists\,\alpha\in(0,\varepsilon)\ \ \exists\,\omega\in z+U\ \ x+\alpha\omega\in Q\}$$

are called *conical approximations* of Q with respect to x.

The following terminology is usually attached:

K,C - cones of approximations

N - cone of internal directions

M - cone of tangents

Precise definitions of these cones and notation vary in the literature; our notation is as used in [4].

Proposition 1

Let P and Q be subsets of X, $x \in X$. Then

(a) $N(x,Q) \subseteq K(x,Q) \subseteq C(x,Q) \subseteq M(x,Q)$;

(b) $K(x,Q)$ and $C(x,Q)$ are cones;
$N(x,Q)$ is an open cone;
$M(x,Q)$ is a closed cone;

(c) $N(x,P) \cap N(x,Q) = N(x,P\cap Q)$;
$K(x,P) \cap K(x,Q) = K(x,P\cap Q)$;
$N(x,P) \cap M(x,Q) \subseteq M(x,P\cap Q)$;
$K(x,P) \cap C(x,Q) \subseteq C(x,P\cap Q)$;

(d) $N(x,X) = X$; $M(x,\phi) = \phi$;

(e) $N(x,Q) = N(x, \operatorname{int} Q)$;
$M(x,Q) = M(x,\bar{Q})$;

(f) if $P \subseteq Q$ then
$N(x,P) \subseteq N(x,Q)$, $K(x,P) \subseteq K(x,Q)$,
$C(x,P) \subseteq C(x,Q)$, $M(x,P) \subseteq M(x,Q)$;

(g) if Q is cone then
$K(\theta,Q) = Q$, $N(\theta,Q) = \operatorname{int} Q$;

(h) if Q is convex then
$K(x,Q) = C(x,Q)$, $N(x,Q) = K(x, \operatorname{int} Q)$

$N(x,Q)$, $K(x,Q)$, $C(x,Q)$ and $M(x,Q)$ are convex

(i) if Q is convex, $\mathrm{int}\, Q \neq \phi$ then
$N(x,Q)^{-} = M(x,Q)$, $N(x,Q) = \mathrm{int}\, M(x,Q)$.

Let Q be a subset of X. The set

$$Q^{+} = \{f \in X^{*};\ f(x) \geq 0 \text{ for all } x \in Q\}$$

is called the *polar cone* of Q.

Proposition 2

Let P and Q be subsets of X. Then

(a) Q^{+} is a convex cone in X^{*};

(b) if $P \subseteq Q$ then $Q^{+} \subseteq P^{+}$;

(c) $X^{+} = \{\theta\}$, $\{\theta\}^{+} = X^{*}$;

(d) $Q^{+} = (\bar{Q})^{+}$;

(e) if Q is convex, $\mathrm{int}\, Q \neq \phi$, then $Q^{+} = (\mathrm{int}\, Q)^{+}$;

(f) if Q is a subspace in X then
$Q^{+} = \{f \in X^{*};\ f(x)=0 \text{ for all } x \in Q\}$.

Proposition 3

Take Q a convex subset of X, $f \in X^{*}$. Consider the assertions:

(a) $f(x) \leq f(y)$ for all $y \in Q$;

(b) $f \in K(x,Q)^{+}$;

(c) $f \in N(x,Q)^{+}$;

Then if $x \in Q$, (a) $\Leftrightarrow$ (b), and if $x \in \bar{Q}$, $\mathrm{int}\, Q \neq \phi$, (a) $\Leftrightarrow$ (b) $\Leftrightarrow$ (c).

Proposition 4

Take Q a cone in X, $x \in \bar{Q}$ and $f \in X^{*}$. If $f(x) \leq f(y)$ for all $y \in Q$, then $f(x) = 0$ and $f \in Q^{+}$.

2.2 We make essential use of the following consequence of the Hahn-Banach Theorem:

Separation Theorem

Let $Q_0, Q_1, \ldots, Q_n$ $(n \geq 1)$ be non-empty convex sets in X, $Q_1, \ldots, Q_n$ open and $\theta \in \bar{Q}_i$, $i = 0, 1, \ldots, n$. Then $\bigcap_{i=0}^{n} Q_i = \phi$ if and only if there exist linear continuous functionals $f_0, f_1, \ldots, f_n$ on X, not all zero, such that

(a) $f_i \in Q_i^{+}$, $i = 0, 1, \ldots, n$;

(b) $\sum_{i=0}^{n} f_i = \theta$.

If $\bigcap_{i=0}^{n} Q_i = Q \neq \phi$, then $Q^+ = \sum_{i=0}^{n} Q_i^+$.

This result was originally stated in [1]. The proof given in [2] for X normed space, carries over to this more general setting unaltered.

In the rest of this section let X and Y be real linear topological spaces and $C \subseteq Y$ a convex cone with non-empty interior.

We say that mapping $T : X \to Y$ is C-*convex*, if

$$T(\lambda x+\mu y) - \lambda T(x) - \mu T(y) \in C$$

for all $x, y \in X$ and $\lambda, \mu \in (0,1)$, $\lambda + \mu = 1$.

Proposition 5

Suppose that $T : X \to Y$ is C-convex and $T(\theta) = \theta$.

(a) If $T^{-1}(\text{int } C) \neq \phi$ and $g \in [T^{-1}(C)]^+$, then there exists $f \in C^+$ such that $f(T(x)) \leq g(x)$ for all $x \in X$.

(b) If $T^{-1}(\text{int } C) = \phi$, then there exists non-zero functional $f \in C^+$ such that $f(T(x)) \leq 0$ for all $x \in X$.

Proof

Let us denote

$$Q_0 = \{(T(x)-\omega,-g(x));\ x \in X,\ \omega \in \bar{C}\},$$

$$Q_1 = \{(y,z);\ y \in \text{int } C,\ z>0\}.$$

Then, evidently, Q_0 is a convex set in $Y \times R$, $(\theta,0) \in Q_0$, and Q_1 is an open convex cone in $Y \times R$. It is easily verified that $Q_0 \cap Q_1 = \phi$, for if it was not so, then there would exist $x \in X$ and $\omega \in \bar{C}$ such that $T(x) - \omega \in \text{int } C$ and $g(x) < 0$. But $T(x) \in C$ because $\omega \in \bar{C}$, and so $g(x) \geq 0$, which is a contradiction.

From the Separation Theorem we conclude, that there exists a non-zero functional $(h,\alpha) \in Y^* \times R$, where $(h,\alpha) \in Q_1^+$ and $-(h,\alpha) \in Q_0^+$. Since $T^{-1}(\text{int } C) \neq \phi$ it follows that $\alpha > 0$ and $h(T(x)-\omega) - \alpha g(x) \leq 0$ for all $x \in X$ and $\omega \in \bar{C}$. We set $f = h/\alpha$. Putting $x = \theta$ and $\omega = \theta$, respectively, we in turn get that $f \in C^+$ and $f(T(x)) \leq g(x)$ for all $x \in X$. To conclude the proof we notice, that if $T^{-1}(\text{int } C) = \phi$, $\text{int } C \cap \{T(x)-\omega;\ x \in X,\ \omega \in \bar{C}\} = \phi$. Applying the Separation Theorem again we obtain a non-zero functional $f \in Y^*$ such that $f(T(x)-\omega) \leq 0$ for all $x \in X$ and $\omega \in \bar{C}$. Therefore $f \in C^+$ and $f(T(x)) \leq 0$ for all $x \in X$.

If $T : X \to Y$ is linear and continuous, then T^* denotes the dual mapping $T^* : Y^* \to X^*$ defined by

$$T^*(f) = f \circ T \text{ for all } f \in Y^*.$$

Corollary

Suppose that $T : X \to Y$ is linear and continuous.

(a) If $T^{-1}(\text{int } C) \neq \phi$, then $[T^{-1}(C)]^+ = T^*(C^+)$.

(b) If $T^{-1}(\text{int } C) = \phi$, then there exists a non-zero functional $f \in C^+$ such that $f \circ T = \theta$.

We take note also of the following standard result.

Proposition 6

Let X, Y be Banach spaces and $T : X \to Y$ linear and continuous projection. Then $[T^{-1}(\{\theta\})]^+ = T^*(Y^*)$.

2.3 To state the results of this sub-section, we need to recall ([5]) the definitions of Fréchet and Gâteaux differentiability.

The mapping $T : X \to Y$ is said to be *Gâteaux differentiable* (briefly G-differentiable) at $x \in X$, if the limit

$$\lim_{t \to 0^+} \frac{T(x+th)-T(x)}{t}$$

written $DT(x,h)$, exists for all $h \in X$. Then $DT(x)$ denotes the mapping $DT(x,\cdot) : X \to Y$.

We say that T is *strongly G-differentiable* at $x \in X$, if for all $h \in X$

$$DT(x,h) = \lim_{\substack{t \to 0^+ \\ k \to h}} \frac{T(x+tk)-T(x)}{t}$$

The mapping $T : X \to Y$, where X and Y are normed spaces, is said to be *Fréchet differentiable* (briefly F-differentiable) at $x \in X$ with F-derivative $DT(x) : X \to Y$, if $DT(x)$ is linear and continuous, and for all $h \in X$

$$\lim_{h \to \theta} \frac{||T(x+h)-T(x)-DT(x)(h)||}{||h||} = 0$$

If moreover, for all $h, k \in X$

$$\lim_{\substack{(h,k)\to(\theta,\theta)\\ h\neq k}} \frac{||T(x+h)-T(x+k)-DT(x)(h-k)||}{||h-k||} = 0$$

then T is said to be *strongly F-differentiable* at x.

The next three propositions follow directly from the above definitions.

Proposition 7

Suppose that T is G-differentiable at $x \in X$. Then

(a) $DT(x)^{-1}\ (K(T(x),\ \text{int}\ C)) \subseteq K(x,T^{-1}(\text{int}\ C))$;

(b) $DT(x)^{-1}\ (\text{int}\ C) \subseteq K(x,T^{-1}(T(x)+\ \text{int}\ C))$.

Proposition 8

Let T be linear mapping and $x \in X$. Then

$$\{h\in X;\ DT(x,h)=\theta\} = K(x,T^{-1}(\{T(x)\})).$$

Proposition 9

Suppose that T is strongly G-differentiable at $x \in X$. Then

(a) $DT(x)^{-1}(N(T(x),\ \text{int}\ C)) \subseteq N(x,T^{-1}(\text{int}\ C))$;

(b) $DT(x)^{-1}(\text{int}\ C) \subseteq N(x,T^{-1}(T(x)+\ \text{int}\ C))$.

Proposition 10

Let X, Y be Banach spaces and suppose that T is strongly F-differentiable at $x \in X$, $DT(x) : X \to Y$ is a projection. Then

$$\{h\in X;\ DT(x)(h)=\theta\} = M(x,T^{-1}(\{T(x)\})).$$

Proof

Let us denote $K = \{h\in X;\ DT(x)(h)=\theta\}$. Since every strongly F-differentiable mapping is strongly G-differentiable, it follows immediately that $M(x,T^{-1}(\{T(x)\})) \subseteq K$.

Conversely, let $k \in K$ and $\varepsilon > 0$ be given. Suppose that $Y \neq \{\theta\}$, for if $Y = \{\theta\}$, then $K = X = M(x,X)$. K is a closed subspace in X, so X/K is also a Banach space and there exists a linear continuous bijection $A : X/K \to Y$, $A \neq \theta$, such that $DT(x) = A \circ f$, where $f : X \to X/K$ is the canonical projection $f(x) = x + K$. By the open-mapping theorem A^{-1} is continuous, hence A is an isomorphism.

Mapping T is strongly F-differentiable, therefore there exists $\delta > 0$ such that

(1) $$||h_1|| < \delta, ||h_2|| < \delta \Rightarrow ||T(x+h_1)-T(x+h_2)-DT(x)(h_1-h_2)|| < c||h_1-h_2||,$$

where $c = \dfrac{1}{4||A^{-1}||}$

Further there exists $\alpha \in (0,\varepsilon)$ such that

(2) $$||h-k|| < \varepsilon \Rightarrow \alpha||h|| < \delta \text{ and}$$

(3) $$||T(x+\alpha k)-T(x)|| < \alpha\varepsilon c/2.$$

We now define sequences $\{k_n\}$ in X and $\{t_n\}$ in X/K in the following manner:

$$k_0 = k, \quad t_0 = K \quad \text{and}$$

(4) $$t_{n+1} = t_n - 1/\alpha\, A^{-1}[T(x+\alpha k_n)-T(x)],$$

(5) $$k_{n+1} \in t_{n+1} \text{ such that } ||k_{n+1}-k_n|| < 2||t_{n+1}-t_n||,\ n \geq 0.$$

Using (3) we note that $||t_1||<\varepsilon/8$, and $||k_1-k|| < \varepsilon/4$. We show by induction that the sequence $\{k_n\}$ has the following properties:

(6) $$||k_{n+1}-k_n|| < 2^{-n}||k_1-k||,$$

(7) $$||k_{n+1} - k|| < \varepsilon/2$$

This would imply that $\{k_n\}$ and $\{t_n\}$ are Cauchy sequences in X and X/K, respectively, and therefore convergent. If $h = \lim k_n$, $s = \lim t_n$, then $h \in s$ and using (4) we conclude that

$$s = s - 1/\alpha\, A^{-1}[T(x+h)-T(x)], \text{ i.e.}$$

$$T(x+\alpha h) = T(x) \text{ and } ||h-k|| < \varepsilon,\ \alpha \in (0,\varepsilon)$$

in other words $k \in M(x,T^{-1}(\{T(x)\}))$.

We saw that (6) and (7) are valid for $n = 0$. Supposing their validity for $m \leq n-1$, we show that they hold for n. Since $k_n \in t_n$, we have $A(t_n) = DT(x)(k_n)$, and

$$t_{n+1} = -A^{-1}\left[\frac{T(x+\alpha k_n)-T(x)}{\alpha} - DT(x)(k_n)\right].$$

Then

$$||t_{n+1}-t_n|| = \frac{||A^{-1}||}{\alpha} . ||T(x+\alpha k_n)-T(x+\alpha k_{n-1})-DT(x)(k_n-k_{n-1})||$$

$$< \frac{||A^{-1}||}{\alpha} .c\alpha||k_n-k_{n-1}|| = 1/4||k_n-k_{n-1}||.$$

Hence

$$||k_{n+1}-k_n|| < 2||t_{n+1}-t_n|| < 1/2||k_n-k_{n-1}|| < 2^{-n}||k_1-k||,$$

and

$$||k_{n+1}-k\ || \leq \sum_{i=0}^{n} ||k_{i+1}-k_i|| \leq 2||k_1-k|| < \varepsilon/2$$

which completes the proof.

3. THE MAIN RESULTS

In this section we consider the Basic Problem. First, under the assumption X and Y_3 are Banach spaces, we give necessary conditions for optimality, in the presence of equality and inequality constraints:

Theorem 1

Let us suppose that either

(a) there exists a non-empty open convex cone $K \subseteq N(x,Q)$,

(b) mappings F and G are strongly G-differentiable at x, mapping H is strongly F-differentiable at x;

or

(a') there exists a non-empty open convex cone $K \subseteq C(x,Q)$,

(b') mappings F anf G are G-differentiable at x, mapping H is linear and continuous,

and

(c) DF(x) and DG(x) are C-convex and S-convex, respectively, and continuous, DH(x) has a closed range.

If x is optimal, then there exist functionals $\alpha \in X^*$, $\lambda \in Y_1^*$, $\mu \in Y_2^*$ and $\nu \in Y_3^*$ such that

(I) $[\alpha+\lambda\circ DF(x)+\mu\circ DG(x)+\nu\circ DH(x)](h) \leq 0$ for all $h \in X$;

(II) $\lambda \neq \theta$ or $\mu \neq \theta$ or $\nu \neq \theta$;

(III) $\alpha \in K^+$, $\lambda \in C^+$, $\mu \in S^+$ and $\mu(G(x)) = 0$.

In the absence of equality constraints, we may dispense with the assumption that X be a Banach space:

Theorem 2

Let us suppose that either

(a) there exists a non-empty convex cone $K \subseteq M(x,Q)$,

(b) mappings F and G are strongly G-differentiable at x;

or

(a') there exists a non-empty convex cone $K \subseteq C(x,Q)$,

(b') mappings F and G are G-differentiable at x;

and

(c) DF(x) and DG(x) are C-convex and S-convex, respectively, and continuous.

If x is optimal, then there exist functionals $\alpha \in X^*$, $\lambda \in Y_1^*$ and $\mu \in Y_2^*$ such that

(I) $[\alpha+\lambda\circ DF(x)+\mu\circ DG(x)](h) \leq 0$ for all $h \in X$;

(II) $\lambda \neq \theta$ or $\mu \neq \theta$;

(III) $\alpha \in K^+$, $\lambda \in C^+$, $\mu \in S^+$ and $\mu(G(x)) = 0$.

The following results, Theorems 3 and 4, consider the Basic Problem when the objective function is single-valued. Under certain convexity assumptions, the necessary conditions in the above theorems become sufficient for optimality, if the multiplier associated with the objective function is non-zero (in particular, if the appropriate Slater's condition holds). As above, we may develop our results in a more general framework where equality constraints are absent.

Theorem 3

Let us suppose that

(a) x is a feasible point;

(b) Q is a convex set with non-empty interior;

(c) mappings F and G are C-convex and S-convex, respectively, and G-differentiable at x, where DF(x) and DG(x) are continuous, mapping H is linear and continuous;

(d) there exist functionals $\alpha \in X^*$, $\mu \in Y_2^*$, $\nu \in Y_3^*$ and a real number λ such that

(I) $[\alpha+\lambda DF(x)+\mu\circ DG(x)+\nu\circ H](h) \leq 0$ for all $h \in X$;

(II) $\lambda \neq 0$ or $\mu \neq \theta$ or $\nu \neq \theta$,

(III) $\alpha(x) \leq \alpha(y)$ for all $y \in Q$,

(IV) $\lambda \in C^+$, $\mu \in S^+$ and $\mu(G(x)) = 0$.

Then

(1) If $\lambda \neq 0$, x is optimal.

(2) If H is a projection and there exists $z \in \text{int } Q$ such that $G(z) \in \text{int } S$, $H(z) = e$, then $\lambda \neq 0$.

Theorem 4

Let us suppose that

(a) x is a feasible point;

(b) Q is convex;

(c) mappings F and G are C-convex and S-convex respectively, and G-differentiable, where $DF(x)$ and $DG(x)$ are continuous;

(d) there exist functionals $\alpha \in X^*$, $\mu \in Y_2^*$ and a real number λ such that

(I) $[\alpha+\lambda DF(x)+\mu \circ DG(x)](h) \leq 0$ for all $h \in X$;

(II) $\lambda \neq 0$ or $\mu \neq \theta$,

(III) $\alpha(x) \leq \alpha(y)$ for all $y \in Q$,

(IV) $\lambda \in C^+$, $\mu \in S^+$ and $\mu(G(x)) = 0$.

Then

(1) If $\lambda \neq 0$, x is optimal.

(2) If there exists $z \in Q$ such that $G(z) \in \text{int } S$, then $\lambda \neq 0$.

4. PROOF OF THE MAIN RESULTS

Here we prove the theorems of Section 3. We shall see that these results follow simply from the Separation Theorem using the properties of conical approximations and polar cones developed in Section 2.

Proof of Theorem 1

Let us suppose that x is optimal and assumptions (a), (b) and (c) hold.

We define

$$P = \{z\in X;\ F(z)-F(x)\in C\setminus\{\theta\}\},$$
$$K_1 = \{h\in X;\ DF(x,h) \in \text{int } C\},$$
$$K_2 = \{h\in X;\ DG(x,h)\in N(G(x),\ \text{int } S)\}, \text{ and}$$
$$K_3 = \{h\in X;\ DH(x,h)=\theta\}.$$

Note that K_1 and K_2 are open convex sets, $\theta \in \bar{K}_1$, $\theta \in \bar{K}_2$ and K_3 is a subspace in X.

We first show that the conclusion of the theorem follows trivially if either $K_1 = \phi$ or $K_2 = \phi$ or $DH(x)$ is not a projection.

Suppose that $K_1 = \phi$. Then using Propositions 2 and 5 we conclude that there is a non-zero functional $\lambda \in (\text{int } C)^+ = C^+$ such that

$$\lambda \circ DF(x,h) \leq 0 \text{ for all } h \in X.$$

Now suppose that $K_2 = \phi$. Using Proposition 5 again it follows that there is a non-zero functional $\mu \in N(G(x),\ \text{int } S)^+$ such that

$$\mu \circ DG(x,h) \leq 0 \text{ for all } h \in X.$$

Propositions 2, 3 and 4 imply that $\mu \in S^+$ and $\mu(G(x)) = 0$.

If $DH(x)$ is not a projection onto Y_3, then the range of $DH(x)$ is a proper closed subspace of Y_3, so by Hahn-Banach theorem there exists a non-zero functional $\nu \in Y_3^*$ such that $\nu \circ DH(x) = \theta$.

Now suppose that both K_1 and K_2 are non-empty and $DH(x)$ is a projection. It follows immediately from Propositions 8 and 10 that

$$\begin{aligned} &K_1 \subseteq N(x,F^{-1}(F(x)+ \text{int } C)) \subseteq N(x,P), \\ &K_2 \subseteq N(x,G^{-1}(\text{int } S)) \subseteq N(x,G^{-1}(S)), \text{ and} \\ &K_3 = M(x,H^{-1}(\{e\})). \end{aligned}$$

We point out that

$$Q \cap P \cap G^{-1}(S) \cap H^{-1}(\{e\}) = \phi \tag{1}$$

by virtue of x being optimal, and consequently using Proposition 1, we conclude that

$$N(x,Q) \cap N(x,P) \cap N(x,G^{-1}(S)) \cap M(x,H^{-1}(\{e\})) = \phi \tag{2}$$

and so

$$K \cap K_1 \cap K_2 \cap K_3 = \phi \tag{3}$$

By the Separation Theorem there exist functionals $f \in K^+$, $f_1 \in K_1^+$, $f_2 \in K_2^+$, $f_3 \in K_3^+$, not all zero, such that

$$f + f_1 + f_2 + f_3 = \theta.$$

From Propositions 5 and 6 it follows that there are functionals $\lambda \in (\text{int } C)^+$, $\mu \in N(G(x), \text{int } S)^+$ and $\nu \in Y_3^*$, not all zero, such that for all $h \in X$

$$\begin{aligned} &\lambda \circ DF(x,h) \leq f_1(h), \\ &\mu \circ DG(x,h) \leq f_2(h), \text{ and} \\ &\nu \circ DH(x) = f_3. \end{aligned}$$

We set $\alpha = f$. Consequently for all $h \in X$

$$[\alpha+\lambda\circ DF(x)+\mu\circ DG(x)+\nu\circ DH(x)](h) \leq 0.$$

As before we observe that $\lambda \in (\text{int } C)^+$ implies $\lambda \in C^+$ and $\mu \in N(G(x),\text{int } S)^+$ implies that $\mu \in S^+$ and $\mu(G(x)) = 0$.

Now let us suppose that (a'), (b') and (c) hold. Using Propositions 7 and 9 we get the following inclusions:

$$K_1 \subseteq K(x,F^{-1}(F(x)+\text{int }C)) \subseteq K(x,P)$$
$$K_2 \subseteq K(x,G^{-1}(\text{int }S)) \subseteq K(x,G^{-1}(S)), \text{ and}$$
$$K_3 = K(x,H^{-1}(\{e\})).$$

Finally we observe that, reasoning almost exactly as before, but with

$$C(x,Q) \cap K(x,P) \cap K(x,G^{-1}(S)) \cap K(x,H^{-1}(\{e\})) = \phi \tag{4}$$

instead of (2) and with

$$K_2 = \{h\in X,\ DG(x,h)\in K(G(x),\ \text{int }S)\}$$

possibly replacing the earlier definition of K_2, we can draw the same conclusions.

We confine ourselves to proving Theorems 1 and 3 only, in as much as virtually the same arguments are used to prove Theorems 2 and 4.

In proving Theorem 2, however, we do not make use of Propositions 6 and 10, and so we need not require X to be a Banach space.

<u>Proof of Theorem 3</u>

Firstly suppose that $\lambda \neq 0$ and let $z \in X$, $z \neq x$, be another feasible point, i.e.

$$z \in Q,\ G(z) \in S \text{ and } H(z) = e.$$

We write $h = z - x$. In view of the convexity of F and G

$$DF(x,h) \in F(z) - F(x) + \bar{C},$$
$$DG(x,h) \in G(z) - G(x) + \bar{S},$$

from which it follows that

$$\lambda DF(x,h) \geq \lambda[F(z)-F(x)],$$
$$\mu \circ DG(x,h) \geq 0.$$

Moreover we have

$$\nu \circ H(h) = \nu(H(z)-H(x)) = 0, \text{ and}$$
$$\alpha(h) = \alpha(z) - \alpha(x) \geq 0.$$

Therefore

$$\lambda[F(z)-F(x)] \leq \alpha(h) + \lambda DF(x,h) + \mu \circ DG(x,h) + \nu \circ H(h) \leq 0.$$

If $C = [0,\infty)$, then $\lambda > 0$ and $F(z) - F(x) \leq 0$, i.e. $F(z) - F(x) \notin C\setminus\{\theta\}$. If $C = (-\infty,0]$, then $\lambda < 0$ and $F(z) - F(x) \geq 0$, so again $F(z) - F(x) \notin C\setminus\{\theta\}$. Hence $\{z\in\Omega,\ F(z)-F(x)\in C\setminus\{0\}\} = \phi$, which proves that x is optimal.

Now let H be a projection and suppose that there exists some $z \in \text{int } Q$ such that $G(z) \in \text{int } S$ and $H(z) = e$. Writing $h = z - x$ and $K = K(x,Q)$, we use

$$K_2 = \{h\in X;\ DG(x,h)\in K(G(x),\ \text{int } S)\}, \text{ and}$$
$$K_3 = \{h\in X;\ H(h)=\theta\}$$

as before.

By hypothesis $h \in K(x,Q)$, since $x, z \in Q$ and Q is convex, moreover $G(x) + \alpha DG(x,h) \in (1-\alpha)G(x) + \alpha G(z) + \bar{S} \subseteq \text{int } S$ for all $\alpha \in (0,1]$, so $h \in K_2$, and $H(h) = H(z) - H(x) = \theta$, i.e. $h \in K_3$. Therefore $K \cap K_2 \cap K_3 \neq \phi$.

Now suppose that $\lambda = 0$. We write

$$f = \alpha$$
$$f_2 = -\alpha - \nu \circ H$$
$$f_3 = \nu \circ H.$$

Then

$$f + f_2 + f_3 = \theta.$$

Proposition 3 implies that $f \in K^+$. Since H is a projection, $f_3 \neq \theta$ if $\nu \neq \theta$, and $f_3 \in K_3^+$. Further we note that $\mu \in S^+$ and $\mu(G(x)) = 0$ imply that $\mu \in K(G(x),\ \text{int } S)^+$; in view of the inequality $\mu \circ DG(x,k) \leq f_2(k)$ for all $k \in X$ we see that $f_2 \in K_2^+$. Moreover $f_2 \neq \theta$ if $\mu \neq \theta$, for in that case $0 < \mu \circ DG(x,h) \leq f_2(h)$. Because $f + f_2 + f_3 = \theta$, and either $f_2 \neq \theta$ or $f_3 \neq \theta$, we conclude from the Separation Theorem that

$$K \cap K_2 \cap K_3 = \phi$$

which is a contradiction.

5. CONCLUDING REMARKS

As emphasised in the introduction, our main concern in this paper has been to give a simple, complete derivation of the optimality conditions. We conclude, however, by drawing attention to some minor differences with available results. Research into necessary conditions for optimality has been limited largely to the case when the range of the equality constraint function is finite dimensional [1,2,3,6,7]. Only recently has attention been given to the more general situation studied here, where this finite dimensionality requirement is disposed with [8,9,10]. Theorem 1 is in a sense complementary to [3, theorem 6.1]. The Theorem gives necessary conditions for optimality under different differentiability assumptions on F and G (Gâteaux differentiability to a continuous C-convex, resp. S-convex function) as compared with [3] (strong Gâteaux differentiability to a C-convex, resp. S-convex function). This is achieved at the cost of expressing the necessary conditions with respect to a smaller "convex approximation" to the underlying set Q.

It does not appear possible, retaining the present level of generality, to remove the hypothesis in Theorems 1 and 2, that the G-differentials are continuous; that it can be dispensed with in [3] leans heavily on the finite dimensionality of the range of H.

Finally we mention that we slightly generalise results in [8] to the extent that the development here is in real linear topological spaces and also in that their strongest result [8, theorem 2.3] is stated for H continuously Fréchet differentiable, while only strong Fréchet differentiability is here required.

REFERENCES

1. Dubovitskiy A.Ya., and Milyutin A.A., Extremum Problems in the Presence of Restrictions, (USSR) Journal of Computational Mathematics and Mathematical Physics, Vol. 5, No. 1, 1965.

2. Halkin H., A Satisfactory Treatment of Equality and Operator Constraints in the Dubovitskii-Milyutin Optimizational Formalism, Journal of Optimization Theory and Applications, Vol. 6, No. 2, 1970.

3. Neustadt L.W., A General Theory of Extremals, Journal of Computer and System Sciences, Vol. 3, No. 1, 1969.

4. Vlach M., On Necessary Conditions of Optimality in Linear Spaces, Comment. Math. Univ. Carolinae, Vol. 11, No. 3, 1970.

5. Nonlinear Functional Analysis, Proceedings of an Advanced Seminar, Ed. L.B. Rall, Academic Press, 1971.

6. Neustadt L.W., An Abstract Variational Theory with Applications to a Broad Class of Optimization Problems, I General Theory, SIAM Journal on Control, Vol. 4, No. 3, 1966.

7. Neustadt L.W., An Abstract Variational Theory with Applications to a Broad Class of Optimization Problems, II Applications, SIAM Journal on Control, Vol. 5, No. 1, 1967.

8. Bazaraa M.S. and Goode J.J., Necessary Optimality Criteria in Mathematical Programming in Normed Linear Spaces, Journal of Optimization Theory and Applications, Vol. 11, No. 3, 1973.

9. Craven B.D., Nonlinear Programming in Locally Convex Spaces, Journal of Optimization Theory and Applications, Vol. 10, No. 4, 1972.

10. Virsan C., Necessary Conditions for Optimization Problems with Operational Constraints, SIAM Journal on Control, Vol. 8, No. 4, 1970.