Mathematical Programming 2 (1972) 1–18. North-Holland Publishing Company

GEOMETRY OF OPTIMALITY CONDITIONS AND CONSTRAINT QUALIFICATIONS *

F.J. GOULD and Jon W. TOLLE
University of North Carolina, Chapel Hill, North Carolina, U.S.A.

Received 6 November 1970
Revised manuscript received 24 April 1971

Certain types of necessary optimality conditions for mathematical programming problems are equivalent to corresponding regularity conditions on the constraint set. For any problem, a certain natural optimality condition, dependent upon the particular constraint set, is always satisfied. This condition can be strengthened in numerous ways by invoking appropriate regularity assumptions on the constraint set. Results are presented for Euclidean spaces and some extensions to Banach spaces are given.

1. Introduction

Consider the optimization problem
P: maximize $f(x)$, subject to

$$g(x) \leq 0 \tag{1.1}$$

$$x \in D \subseteq R^n \tag{1.2}$$

where the objective function f: $R^n \to R$ and the constraint functions g: $R^n \to R^m$ are assumed continuous on an open set containing D. For simplicity and without loss of generality it can be assumed that any equality constraints of interest have each been rewritten as a pair of inequalities and are thereby included in P. **

* This work was supported in part by the Office of Naval Research, Contract No. N00014–67–A–0321–0003 (NR–047–095).

** This assumption is purely for expository convenience. For purposes of computing it is inefficient to transform equality constraints into pairs of inequalities. For recent algorithmic discussions, see, for example, Abadie [1], Colville [6] and Fletcher [10], and references therein.

Associated with problem P is the constraint set S, defined to be the collection of points satisfying (1.1) and (1.2). It is said that any objective function f, to be optimized over S, has a local constrained maximum at $x_0 \in S$ if x_0 is a local solution to P, i.e., there exists an $\epsilon > 0$ such that for all $x \in S \cap \{x\colon \|x - x_0\| < \epsilon\}$, $f(x) \leq f(x_0)$. This paper deals only with objective functions and constraint functions which are differentiable at some distinguished local solution, x_0.

An optimality criterion for P is a condition on the gradient of f which must hold at points where f has a local constrained maximum. For example, the usual Kuhn-Tucker optimality criterion [19] asserts that if f has a local constrained maximum at x_0 then there is a nonnegative $\lambda \in R^m$ such that

$$\nabla f(x_0) - \sum_{i=1}^{m} \lambda_i \nabla g_i(x_0) = 0 \tag{1.3}$$

$$\langle \lambda, g(x_0) \rangle = 0. \tag{1.4}$$

It is known that without an additional regularity assumption, called a constraint qualification, the above assertion may not be true. As an example, consider the problem

$$\text{maximize } -x_1 - x_2, \text{ subject to}$$

$$g_1(x_1, x_2) = x_2 - x_1^3 \leq 0 \tag{1.5}$$

$$g_2(x_1, x_2) = -x_2 \leq 0$$

and take D to be the entire space R^2. This problem has a solution at $x_0 = (0, 0)$, but there is no nonnegative $\lambda \in R^2$ such that (1.3) and (1.4) hold. To rule out such possible exceptions it has been customary to impose a constraint qualification, which is a regularity condition assumed to be satisfied by the constraints g and the constraint set S at some point $x_0 \in S$. It will be seen that such a condition implies that a specified optimality criterion is satisfied by all objective functions (differentiable at x_0) with a local constrained optimum at x_0.

In their 1951 paper Kuhn and Tucker presented a constraint qualification such that (1.3) and (1.4) are valid when the qualification is satisfied. Since then numerous other papers on constraint qualifications have appeared. The interest has been mainly in determining the weakest possible such qualifications, in determining strong but "easily verifiable" qualifications, and in extending the results to increasingly general problems. See, for example, the papers by Abadie [2], Arrow, Hurwicz, Uzawa [3], Cottle [7], Evans [8], Mangasarian and Fromovitz [21], Slater [24], Ritter [22], Varaiya [25], Guignard [15], Canon, Cullum and Polak [4], Hurwicz [17], Zlobec [27], Gould and Tolle [13], [14], and the books by Hadley [16], Karlin [18], Zangwill [26], and Mangasarian [20].

It is thus apparent that the topic of constraint qualifications has been one of the most extensively researched theoretic areas in mathematical programming, and numerous authors have contributed results, as referenced above. In this paper a discussion of constraint qualifications and necessary optimality criteria is presented which is based entirely on the source papers due to Guignard [15] and the authors [13], [14]. While the paper is expository in nature, it is not intended to be a general survey of the many previous works on constraint qualifications. Rather, our intent is to present a synthesis of the known geometric relations between constraint qualifications and necessary optimality conditions. For comparisons with other works and particularly for more detailed discussions of relations between previous works underlying this paper the interested reader is referred to the sources [13], [14], [15] and to the above mentioned references. It should also be mentioned that the main emphasis here is on Euclidean spaces, though in the last section extensions to a Banach space are given.

In summary, the following exposition describes relations between each member of a family of constraint qualifications and a corresponding optimality condition. It is shown that in a certain sense each optimality condition is equivalent to a condition on the geometry of the constraint set, and that for any problem a certain natural optimality condition, dependent upon the particular constraint set, is always satisfied. This condition can be strengthened in numerous ways by invoking stronger constraint qualifications.

This work, as stated above, is based upon the results of Guignard [15] and previous works of the authors [13] and [14]. Guignard demonstrated the existence of optimality criteria which are different from the usual Kuhn-Tucker conditions (1.3) and (1.4). This allowed the application of optimality theory to problems such as (1.5) which could not be previously treated. The authors have focused on the usual Kuhn-Tucker conditions in [13], and on extensions to Banach space in

[14], and have presented a constraint qualification which is necessary and sufficient for these conditions to hold for all f with a local constrained maximum at the point under consideration. Some open questions and directions for further research are also discussed.

2. Notation and terminology

An arbitrary point $x_0 \in S$ will be distinguished for consideration. In terms of this point, define the following entities.

F_0 = {objective functions f with a local constrained maximum at x_0}

$DF_0 = \{z \in R_n : z = \nabla f(x_0) \text{ for some } f \in F_0\}$

$I_0 = \{i \in \{1, 2, ..., m\} : g_i(x_0) = 0\}$

$C_1 = \{z \in R^n : \langle z, \nabla g_i(x_0)\rangle \leq 0, \text{ all } i \in I_0\}$

$B_0^* = \{z \in R^n : z = \Sigma_{I_0} \lambda_i \nabla g_i(x_0) \text{ for some scalars } \lambda_i \geq 0\}$.

The term "active constraints" is used for those constraints indexed by I_0. The set C_0 is a closed convex cone called the linearizing cone, and B_0^* is a closed convex cone called the cone of gradients.

If A and H are sets in R^n, then $\overline{A}$ will denote the closure of A, A/H the relative complement of H in A, and $A + H$ will be the set of all points of the form $x + y$, where $x \in A, y \in H$.

Definition 1. $x \in R^n$ is said to be in the *polar cone* of A, denoted A', if and only if $\langle x, y\rangle \leq 0$ for all $y \in A$.
Properties of the polar cone relevant to this work are
(i) A' is a closed convex cone
(ii) $A_1 \subseteq A_2 \Rightarrow A_2' \subseteq A_1'$
(iii) $A'' = A \iff A$ is a closed convex cone
(iv) $A' = \overline{[\text{convex hull of } A]}'$
(v) $(A_1 \cap A_2)' = \overline{A_1' + A_2'}$
(vi) If A_1, A_2 are convex polyhedral cones, then $\overline{A_1' + A_2'} = A_1' + A_2'$ and hence, by (v), $(A_1 \cap A_2)' = A_1' + A_2'$.
For a discussion of polar cones and their properties see [12], [15], and [23].

Definition 2. If A is nonempty, then the *cone of tangents* to A at $x_0 \in A$, denoted $T(A, x_0)$, is the set of all $x \in R^n$ such that there exists a sequence $\{x_n\} \in A$, converging to x_0, and a nonnegative sequence $\{\lambda_n\} \in R$ such that $\{\lambda_n(x_n - x_0)\}$ converges to x. The set $T(A, x_0)$ is a nonempty closed cone determined by the geometry of A. It need not be convex, but if A is convex then $T(A, x_0)$ is also convex. The cone of tangents has been previously described by Abadie [2] and has been employed in numerous papers by a variety of authors. It will be seen that the set $T'(S, x_0)$ plays a key role in the following exposition.

3. Optimality conditions in Euclidean space

It can be shown that if f is an objective function with a local constrained maximum at x_0 then $\nabla f(x_0) \in T'(S, x_0)$ [2], [15], [25]. In terms of the notation introduced above, $DF_0 \subseteq T'(S, x_0)$. In [13] it was shown that this result could be strengthened to

$$DF_0 = T'(S, x_0) . \tag{3.1}$$

That is, every vector in $T'(S, x_0)$ is actually a gradient of some objective function which has a local constrained maximum at x_0. It will be seen that the equivalence of these two sets is of interest in describing the relationship between constraint qualifications and optimality conditions. Another useful relation which follows from a result of Abadie [2], is that $T(S, x_0) \subseteq C_0$. Then, since $S \subseteq D$, it must be true that $T(S, x_0) \subseteq T(D, x_0)$, and hence $T(S, x_0) \subseteq C_0 \cap T(D, x_0)$. Now using properties (ii) and (v) of polar cones one obtains

$$\overline{C_0' + T'(S, x_0)} = [C_0 \cap T(D, x_0)]' \subseteq T'(S, x_0) . \tag{3.2}$$

A final relation of immediate interest is

$$B_0^* = C_0' \tag{3.3}$$

which has been shown [2], [13], by an application of Farkas' lemma [9].

It is worthwhile to note that the cone C_0 is a local linearization at the point x_0 of that portion of the constraint set S determined by the active constraint functions g, ignoring the set D. As an example, con-

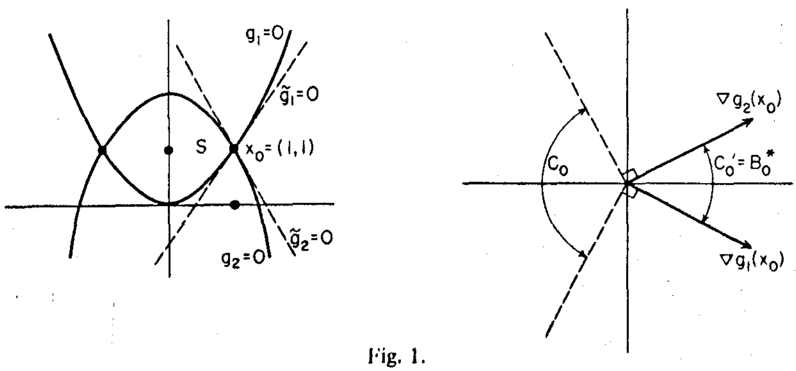

Fig. 1.

sider fig. 1, where for notational convenience it will be assumed that $D = R^2$. The constraint functions are given by

$$g_1(x_1, x_2) = x_1^2 - x_2 \leq 0$$

$$g_2(x_1, x_2) = x_1^2 + x_2 - 2 \leq 0 .$$

The linear approximations to g_1 and g_2 at $x_0 = (1, 1)$, denoted $\tilde{g}_1, \tilde{g}_2$, are given by

$$\tilde{g}_1(x) = g_1(x_0) + \langle \nabla g_1(x_0), x - x_0 \rangle = \langle \nabla g_1(x_0), x - x_0 \rangle$$

$$\tilde{g}_2(x) = g_2(x_0) + \langle \nabla g_2(x_0), x - x_0 \rangle = \langle \nabla g_2(x_0), x - x_0 \rangle .$$

At the point x_0, the constraint set S, given by $\{x\colon g(x) \leq 0\}$, is approximated by $x\colon \tilde{g}(x) \leq 0$, which is the intersection of the two half-spaces $\tilde{g}_1(x) \leq 0, \tilde{g}_2(x) \leq 0$. The linearizing cone C_0 is simply the translation of this approximating set to the origin, and (3.3) indicates that the polar of C_0 is B_0^*, the polyhedral cone of gradients determined by $\nabla g_1(x_0)$, $\nabla g_2(x_0)$.

It can also be verified that for the example illustrated by fig. 1 the set $T(S, x_0)$ is the same as C_0 and hence $T'(S, x_0) = C_0'$.

By now combining the relations (3.1), (3.2) and (3.3) it is seen that the following scheme is justified.

$$\begin{array}{ccc} & DF_0 = T'(S, x_0) \supseteq [C_0 \cap T(D, x_0)]' \\ & \qquad\qquad\qquad\qquad \| \\ B_0^* + T'(D, x_0) = C_0' + T'(D, x_0) & \subseteq \overline{C_0' + T'(D, x_0)} \, . \end{array} \tag{3.4}$$

From this it is immediate that the reduced scheme

$$\begin{array}{c} DF_0 = T'(S, x_0) \\ \cup| \\ B_0^* = C_0' \end{array} \tag{3.5}$$

holds. Now consider the Kuhn-Tucker optimality conditions (1.3) and (1.4), which can be conveniently rewritten as

$$\exists\, \lambda_i \geq 0, i \in I_0, \ni \nabla f(x_0) - \sum_{I_0} \lambda_i \nabla g_i(x_0) = 0. \tag{3.6}$$

In terms of the above notation, (3.6) is equivalent to the assertion that $\nabla f(x_0)$ is in the cone of gradients B_0^*. Hence, (3.6) holds for all objective functions with a local constrained maximum at x_0 if and only if

$$DF_0 \subseteq B_0^* . \tag{3.7}$$

But from (3.5) it is seen that (3.7) holds if and only if

$$T'(S, x_0) = C_0' . \tag{3.8}$$

Note that (3.8) is a condition which refers to the constraint set S and the constraint functions g, but it is independent of the objective function f in P. Actually the left side of (3.8) depends only upon the geometry of S, whilst the right side is determined by the analytic specification of the problem. Since, by (3.3), C_0' is the polyhedral cone determined by the gradients of the active constraints, enlarging the number of constraints can only enlarge the set C_0'. If the constraint set S satisfies a regularity condition of being "well enough specified" by the constraints g, then (3.8) will hold. Examples are easily constructed to show that by adding redundant constraints (those which do not change S) a situation where C_0' is a proper subset of $T'(S, x_0)$ can be converted to one in which (3.8) holds. For instance, in example (1.5) C_0' is a proper subset of $T'(S, x_0)$. If the additional constraint $-x_1 \leq 0$ is added, then S remains unchanged but C_0' is enlarged to the extent that (3.8) holds. In this sense it may appear that the validity of (3.8) is merely a question of "proper specification," or "proper problem formulation." However, it is not known whether every nonlinear program can be "regularized" in the sense of forcing (3.8) to hold by the addition of a finite number of redundant constraints. This notion might be a topic for fruitful research. The idea of "regularizing" constraint sets (for purposes other than those of concern here) has been useful in linear programming [5].

A relation such as (3.8) is called a constraint qualification. All constraint qualifications have the property that the objective function in P can be changed without influencing whether or not they hold. As discussed above, a constraint qualification is an assumption about the relation between the geometry and the analytic specification of the constraint set.

The above remarks now lead to the following theorem which indicates the importance of the qualification (3.8).

Theorem 1. *The optimality condition* (3.6) *is valid for all* $f \in F_0$ *if and only is the constraint qualification* (3.8) *is satisfied.*

Those problems for which $T'(S, x_0) = C_0'$ enjoy in the sense of theorem 1, the best of all known situations, because the Kuhn-Tucker optimality criterion (3.6) is the strongest known necessary condition. However, with references to the scheme (3.5), it is possible to have problems for which C_0' is a proper subset of $T'(S, x_0)$ and (3.8) is not true. Then there will exist objective functions with a local constrained maximum at x_0 for which the Kuhn-Tucker optimality conditions do not hold. The example (1.5) illustrates this possibility. Until recently there has been the uncomfortable necessity of dismissing such bizarre cases without much discussion. However, it will be shown that it is indeed possible in such cases to find new optimality criteria which will hold for all f. * In obtaining these new conditions the requirement that C_0' be as large as $T'(S, x_0)$ is relaxed, and accordingly it will be seen that the new optimality criteria are weaker and hence less satisfactory as necessary conditions.

Considering, then, the possibility that C_0' is a proper subset of $T'(S, x_0)$, it is noted from (3.4) that the following reduced scheme is always valid

* There are other known conditions, called the Fritz John conditions [11], which are known to apply to all problems regardless of whether or not (3.8) is true. However, the Fritz John conditions appear to be of substantial interest only when (3.8) holds, in which case they reduce to the Kuhn-Tucker conditions. For further discussion see [13], [22].

$$\begin{array}{ccc} DF_0 & = & T'(S, x_0) \\ \cup\!\mid & & \cup\!\mid \\ B_0^* + T'(D, x_0) & = & C_0' + T'(D, x_0) \,. \end{array} \tag{3.9}$$

Now consider the new optimality criterion

$$\exists \lambda_i \geq 0, i \in I_0, \ni \nabla f(x_0) - \sum_{I_0} \lambda_i \nabla g_i(x_0) \in T'(D, x_0) \,. \tag{3.10}$$

The condition (3.10) holds for all $f \in F_0$ if and only if

$$DF_0 \subseteq B_0^* + T'(D, x_0) \,. \tag{3.11}$$

But from (3.9) it is seen that (3.11) holds if and only if

$$T'(S, x_0) = C_0' + T'(D, x_0) \,. \tag{3.12}$$

The condition (3.12) can also be considered to be a constraint qualification and it is clear that the following result is true.

Theorem 2. *The optimality condition* (3.10) *is valid for all* $f \in F_0$ *if and only if the constraint qualification* (3.12) *is satisfied.*

Note that if (3.8) holds then (3.12) holds also, but for a problem where (3.8) does not hold it may be possible that the less restrictive condition (3.12) holds, in which case the optimality condition (3.10) is valid for all f with a local constrained maximum at x_0. As an example, consider the problem

$$\begin{aligned} &\max -x_1 - x_2, \text{ subject to} \\ &g_1(x_1, x_2) = x_2 - x_1^3 \leq 0 \\ &g_2(x_1, x_2) = -x_1 \leq 0 \\ &x \in D \,, \end{aligned}$$

where D is the complement of $\{(x_1, x_2): x_1^2 + (x_2 + 1)^2 < 1\}$. The quantities of interest for this problem are sketched in fig. 2. Although the objective function can be made arbitrarily large on S, there is actually a local constrained maximum at the point $x_0 = (0, 0)$. It is easily veri-

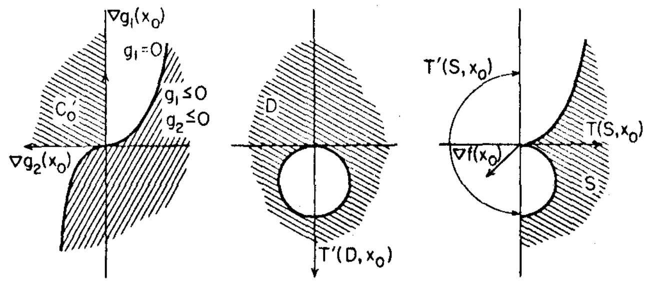

Fig. 2.

fied that C_0' is the second quadrant, $T'(S, x_0)$ consists of the second and third quadrants, and hence C_0' is a proper subset of $T'(S, x_0)$. It is also easily verified that the Kuhn-Tucker conditions do not hold at x_0. However, $T'(D, x_0)$ is the nonpositive portion of the x_2 axis, and hence $T'(S, x_0) = C_0' + T'(D, x_0)$. Consequently, by theorem 2, (3.10) must be valid. For this problem, then, the following optimality condition must hold: there exist nonnegative scalars λ_1, λ_2 such that

$$\frac{\partial f}{\partial x_1}(x_0) - \lambda_1 \frac{\partial g_1}{\partial x_1}(x_0) - \lambda_2 \frac{\partial g_2}{\partial x_1}(x_0) = 0$$

$$\frac{\partial f}{\partial x_2}(x_0) - \lambda_1 \frac{\partial g_1}{\partial x_2}(x_0) - \lambda_2 \frac{\partial g_2}{\partial x_2}(x_0) \le 0 .$$

As mentioned earlier, the Kuhn-Tucker criterion is the strongest known necessary condition for optimality. This means that the criterion (3.6) is "better than" (3.10) in the sense that it specifies a smaller set of candidates for optimality. On the other hand, the validity of (3.10) can be guaranteed for a larger class of problems.

The above discussion will now be generalized to show that for *any* problem P a nontrivial optimality criterion can be found which will be valid for all objective functions with a local constrained maximum at x_0. The criterion will be dependent upon the constraint geometry of the particular problem and the validity of the criterion for all $f \in F_0$ will be equivalent to a constraint qualification which will always be satisfied.

The following lemma will be employed.

Lemma. *Suppose A, H, and K are nonempty cones, with K convex, and $A \cup H = K$. Then $K = A + H$.*

Proof. Suppose $x \in K$. Then $x \in A$ or $x \in H$. Since A and H each contain the origin, $x \in A + H$. Hence $K \subseteq A + H$. Suppose $x \in A + H$. Then $x = a + h$, $a \in A$, $h \in H$. But $A \subseteq K$ and $H \subseteq K$, hence $x = a + h$, $a \in K$, $h \in K$. Since K is a convex cone, $K + K = K$. Hence $a + h \in K$, $x \in K$, and $A + H \subseteq K$.

To employ this result, observe, from (3.5), that $C_0' \subseteq T'(S, x_0)$ and also $T'(S, x_0)/C_0' \subseteq T'(S, x_0)$. We recall that $T'(S, x_0)/C_0'$ denotes the relative complement of C_0' in $T'(S, x_0)$. The set $T'(S, x_0)/C_0'$ does not contain the origin and hence is not a cone. However, letting $\{0\}$ denote the set containing the origin, $T'(S, x_0)/C_0' \cup \{0\}$ is easily seen to be a cone. Since $T'(S, x_0)$ is convex and since $T'(S, x_0) = C_0' \cup T'(S, x_0)/C_0' \cup \{0\}$, it follows from the above lemma that

$$T'(S, x_0) = C_0' + (T'(S, x_0)/C_0' \cup \{0\}) . \tag{3.13}$$

From (3.13), (3.1) and (3.3) it follows that the scheme

$$\begin{array}{ccc} DF_0 & = & T'(S, x_0) \\ \| & & \| \\ B_0^* + (T'(S, x_0)/C_0' \cup \{0\}) & = & C_0' + (T'(S, x_0)/C_0' \cup \{0\}) \end{array} \tag{3.14}$$

is always true. Now consider the optimality criterion

$$\exists \lambda_i \geq 0, i \in I_0, \ni \nabla f(x_0) - \sum_{I_0} \lambda_i \nabla g_i(x_0)$$
$$\in T'(S, x_0)/C_0' \cup \{0\} . \tag{3.15}$$

It follows from (3.14) that

Theorem 3. *Given any problem,* P, *the optimality condition* (3.15) *is valid for all $f \in F_0$.*

In the special case of problems for which (3.8) is satisfied, then $T'(S, x_0)/C_0'$ is empty and (3.15) reduces to the Kuhn-Tucker conditions. However, considering problems for which (3.8) is not satisfied, theorem 3 guarantees that there exists a set of nonnegative multipliers and a cone $G \subseteq T'(S, x_0)$ such that $\nabla f(x_0) - \Sigma_{I_0} \lambda_i \nabla g_i(x_0) \in G$ if

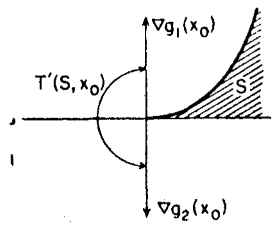

Fig. 3.

$f \in F_0$. In particular one can take any cone G which satisfies the relation $T'(S, x_0) = C_0' + G$. It should be noted that this does not imply that $G = T'(S, x_0)/C_0' \cup \{0\}$, as in certain cases numerous cones $G \subseteq T'(S, x_0)$ will satisfy

$$T'(S, x_0) = C_0' + (T'(S, x_0)/C_0' \cup \{0\}) = C_0' + G\,. \tag{3.16}$$

For example, consider again problem (1.5) for which (3.8) is not satisfied and the Kuhn-Tucker conditions do not hold. As indicated in fig. 3, the cone C_0 is the x_1 axis, C_0' is the x_2 axis, $T(S, x_0)$ is the nonnegative x_1 axis, and $T'(S, x_0)/C_0' \cup \{0\}$ is $\{(x_1, x_2): x_1 < 0\} \cup \{(0, 0)\}$. In this case, if G is taken to be any half ray leading to the left from the origin (i.e. into $T'(S, x_0)/C_0'$) then (3.16) will hold. In particular, as illustrated in [15] by Guignard, G can be taken as the nonpositive part of the x_1 axis to obtain the optimality conditions: there exist nonnegative scalars λ_1, λ_2 such that

$$\frac{\partial f}{\partial x_1}(x_0) - \lambda_1 \frac{\partial g_1}{\partial x_1}(x_0) - \lambda_2 \frac{\partial g_2}{\partial x_1}(x_0) \leq 0 \tag{3.17}$$

$$\frac{\partial f}{\partial x_2}(x_0) - \lambda_1 \frac{\partial g_1}{\partial x_2}(x_0) - \lambda_2 \frac{\partial g_2}{\partial x_2}(x_0) = 0 \tag{3.18}$$

which are valid for all $f \in F_0$.

It is important to note that the conditions (3.17) and (3.18) are stronger than (3.15). The latter would yield (3.17) without (3.18). It is obviously desirable to obtain the strongest possible optimality conditions for a given problem P. Accordingly, as pointed out by Guignard in [15a], in such cases as that illustrated by the above example, it would be desirable to have a tighter theory in order to find the "smallest sets" G, smaller than $T'(S, x_0)/C_0' \cup \{0\}$ which satisfy (3.16). This appears to be another interesting direction for further research.

In general, then, it has been shown that independent of any special regularity assumptions the optimality condition (3.15) holds for all $f \in F_0$. Furthermore, it is seen that if G is any cone contained in $T'(S, x_0)$ such that

$$T'(S, x_0) = C_0' + G \tag{3.19}$$

holds, then the optimality condition (3.20)

$$\begin{gathered}\text{there exist } \lambda_i \geq 0, i \in I_0, \text{ such that} \\ \nabla f(x_0) - \sum_{I_0} \lambda_i \nabla g_i(x_0) \in G\end{gathered} \tag{3.20}$$

is valid for all $f \in F_0$. Conversely, if (3.20) is valid for all $f \in F_0$ then the constraint qualification (3.19) must be satisfied. To obtain the strongest known conditions, those of Kuhn-Tucker, the cone G is taken to be $\{0\}$, in which case the qualification (3.19) reduces to (3.8). In the other extreme, to obtain (nontrivial) weak conditions which are always valid, take G to be the cone $T'(S, x_0)/C_0' \cup \{0\}$. In between these two extremes a family of natural optimality conditions can be stated by varying the choice of G. Each such condition is equivalent to a corresponding constraint qualification given by (3.19) in the sense that it is valid for all $f \in F_0$ if and only if (3.19) is satisfied.

Thus, if the geometry of a given problem is well understood then a "best" optimality condition for that problem can be stated. This is illustrated by the conditions given in equations (3.17) and (3.18). If on the other hand it is desired to consider all problems for which a specified optimality condition is valid, then it must be assumed that the geometry satisfies certain corresponding regularity conditions.

Finally, it is noted that the Kuhn-Tucker conditions (3.6) are entirely algebraic and are independent of the particular problem geometry. The geometric dependence is removed by the assumption (3.8). The more general conditions (3.20) are not entirely algebraic, since the set G explicitly depends upon the particular constraint set S and the functions g. Thus the underlying relations between optimality conditions and geometric considerations are somewhat more explicit in the criterion (3.20).

4. Results in Banach space

In this section we briefly describe an extension of the above presentation to Banach space. All of this material is based upon results presented in [14] and [15]. (Ritter [22] has done some related work in a partially ordered Banach space, and Zlobec [27] has presented some of the following structure in his weakening of Guignard's sufficient conditions presented in [15]). Let X, Y denote Banach spaces, each with the norm topology, and assume X is reflexive. Suppose $A_x \subseteq X$, $A_y \subseteq Y$ and assume $g: X \to Y$. Define the constraint set

$$S = A_x \cap g^{-1}(A_y)\,. \tag{4.1}$$

Suppose $f: X \to R$. Then the optimization problem of interest is

$$\begin{aligned}&\max f(x),\text{ subject to}\\ &\quad x \in S\,.\end{aligned} \tag{4.2}$$

An arbitrary point $x_0 \in S$ is distinguished for consideration and f and g are assumed to be differentiable at x_0. The sets F_0 and DF_0 are analogous to the corresponding sets defined in the previous section. Denote the topological duals of X, Y as X^*, Y^*, respectively, and for any set $N^* \subseteq X^*$ let $\overline{N}^*$ denote the closure of N^* in the weak $*$ topology. Let M be an arbitrary subset of X.

Definition 3. The *polar cone of M* denoted M', is the subset of X^* given by

$$M' = \{x^* \in X^*\colon x^*(m) \leq 0\ \forall m \in M\}\,.$$

Properties (i) thru (v) of polar cones in Euclidean spaces, given in section 2, are also true in Banach spaces [25], [14], [15].

Definition 4. The *cone of tangents* to M at $x_0 \in M$, denoted $T(M, x_0)$, is the set of all $x \in X$ such that there exist sequences $\{\lambda_n\} \in R$, $\lambda_n \geq 0$, all n, and $\{x_n\} \in M$ such that

(i) $x_n \to x_0$

(ii) $\lambda_n(x_n - x_0) \to x$.

Definition 5. The *weak cone of tangents* to M at x_0, denoted $T_w(M, x_0)$, is defined the same as $T(M, x_0)$ except that (ii) is replaced by

(ii)′ $\lambda_n(x_n - x_0) \to x$ weakly.

Definition 6. The *weak cone of pseudotangents* to M at $x_0 \in M$, denoted $P_w(M, x_0)$, is the closure of the convex hull of $T_w(M, x_0)$.

Definition 7. The *weak pseudolinearizing cone* at $x_0 \in S$, denoted $K_w(x_0)$, is the subset of X given by $\{x \in X: Dg(x_0) \in P_w(A_y, g(x_0))\}$, where $Dg(x_0)$ denotes the derivative of g at x_0.

Definition 8. The *weak cone of gradients* at $x_0 \in S$, denoted $B_w^*(x_0)$, is the subset of X^* given by

$$B_w^*(x_0) = \{z^* \in X^*: z^* = \Psi^* \circ Dg(x_0) \text{ for some } \Psi^* \in T'_w(A_y, g(x_0))\} . \tag{4.3}$$

Properties of the above sets are discussed in [14], where it is shown that the following relations hold. *

$$\begin{array}{c} DF_0 = T'_w(S, x_0) \\ \cup| \\ \overline{B_w^*(x_0)} = K'_w(x_0) . \end{array} \tag{4.4}$$

Now consider the optimality condition

$$Df(x_0) \in \overline{B_w^*(x_0)} . \tag{4.5}$$

Since $B_w^*(x_0)$ is not necessarily closed, this condition does not guarantee the existence of a $\Psi^* \in T'_w(A_y, g(x_0))$ such that

$$Df(x_0) = \Psi^* \circ Dg(x_0) .$$

*** This can be simplified when A_y is convex. In this case it is demonstrated in [14] that $T_w(A_y, g(x_0)) = T(A_y, g(x_0)) = P_w(A_y, g(x_0))$.**

However since $B_w^*(x_0)$ is convex and X is reflexive, the weak $*$ closure and the strong closure of $B_w^*(x_0)$ are the same. Hence (4.5) means that there exists a sequence $\{\Psi_n^*\} \in T'_w(A_y, g(x_0))$ such that

$$Df(x_0) = \lim_{n\to\infty} \{\Psi_n^* \circ Dg(x_0)\}$$

Thus condition (4.5) can be viewed as a type of asymptotic Kuhn-Tucker condition (for a discussion of asymptotic sufficient conditions see Zlobec [27]). Conditions under which $B_w^*(x_0)$ is closed are given in [14].

It is apparent that condition (4.5) is valid for all $f \in F_0$ if and only if $DF_0 \subseteq \overline{B_w^*(x_0)}$. Consequently, the scheme (4.4) implies

Theorem 4. *The optimality condition* (4.5) *is valid for all* $f \in F_0$ *if and only if the constraint qualification* $T'_w(S, x_0) = K'_w(x_0)$ *holds.*

It is also shown in [14] that if reflexivity of X is dropped, then the following scheme holds

$$\begin{array}{c} DF_0 \subseteq T'_w(S, x_0) \\ \cup\!| \\ \overline{B_w^*(x_0)} = K'_w(x_0)\,. \end{array} \tag{4.6}$$

Now let G be any prescribed cone in $T'_w(S, x_0)$ such that the constraint qualification

$$T'_w(S, x_0) = K'_w(x_0) + G \tag{4.7}$$

holds (note that if $T'_w(S, x_0) \subseteq K'_w(x_0) + G$ then the assumption that G is a cone in $T'_w(S, x_0)$, plus the facts that $K'_w(x_0) \subseteq T'_w(S, x_0)$ and $T'_w(S, x_0)$ is convex imply that $K'_w(x_0) + G \subseteq T'_w(S, x_0) + T'_w(S, x_0) = T'_w(S, x_0)$ and hence (4.7) holds). Then it follows from (4.6) that the optimality criterion

$$Df(x_0) \in \overline{B_w^*(x_0)} + G \tag{4.8}$$

is valid for all $f \in F_0$. If X is reflexive then it follows from (4.4) that (4.7) is a necessary and sufficient condition for (4.8) to be valid for all $f \in F_0$.

If for a given optimization problem $K'_w(x_0)$ is a proper subset of $T'_w(S, x_0)$ then (4.5) may not be valid. However, for any problem (4.7) is always satisfied if G is taken to be $T'_w(S, x_0)/K'_w(x_0) \cup \{0\}$. Hence, in this case, the optimality condition

$$Df(x_0) \in B^*_w(x_0) + (T'_w(S, x_0)/K'_w(x_0) \cup \{0\}) \tag{4.9}$$

holds for all $f \in F_0$ without any regularity assumptions on (4.2).

It is also apparent that the cone G in $T'_w(S, x_0)$ can be varied to give various optimality conditions and that the results for Euclidean spaces can be paralleled in higher dimensional spaces.

References

[1] J. Abadie, "Generalization of the Wolfe reduced gradient method to the case of nonlinear constraints," in: *Optimization*, Ed. R. Fletcher (Academic Press, 1969).

[2] J. Abadie, "On the Kuhn-Tucker theorem," in: *Nonlinear Programming*, Ed. J. Abadie (North-Holland Publishing Co., Amsterdam, 1967).

[3] K. Arrow, L. Hurwicz and H. Uzawa, "Constraint qualifications in maximization problems," *Naval Research Logistics Quarterly* 8 (1961) 175–191.

[4] M. Canon, C. Cullum and E. Polak, "Constrained minimization problems in finite-dimensional spaces," *SIAM Journal on Control* 4 (1966) 528–547.

[5] A. Charnes and W. Cooper, *Management models and industrial applications of linear programming* (John Wiley and Sons, New York, 1961).

[6] A. Colville, "A comparative study of nonlinear programming codes," IBM, New York Scientific Center Report 320–2949, NYSC, IBM, 410 East 62nd St., New York, New York, 10021.

[7] R. Cottle, "A theorem of Fritz John in mathematical programming," RAND Memorandum RM–3858–PR (October 1963).

[8] J. Evans, "On constraint qualifications in nonlinear programming," *Naval Research Logistics Quarterly*, to appear.

[9] J. Farkas, "Über die Theorie der einfachen Ungleichungen," *Zeitschrift für die reine und angewandte Mathematik* 124 (1901) 1–27.

[10] R. Fletcher, "A class of methods for nonlinear programming with termination and convergence properties," *Integer and Nonlinear Programming*, Ed. J. Abadie (North-Holland Company, Amsterdam, 1970).

[11] F. John, "Extremum problems with inequalities as subsidiary conditions, Studies and Essays," *Courant Anniversary Volume* (Interscience, New York, 1948).

[12] D. Gale, *The theory of linear economic models* (McGraw-Hill Book Company, Inc., New York, 1960) p. 44.

[13] F. Gould and J. Tolle, "A necessary and sufficient qualification for constrained optimization," *SIAM Journal on Applied Mathematics* 20 (1971) 164–172.

[14] F. Gould and J. Tolle, "Optimality conditions and constraints qualifications in Banach space," Institute of Statistics Mimeo Series No. 707, Department of Statistics, University of North Carolina, Chapel Hill, N.C. 27514.

[15] M. Guignard, "Generalized Kuhn-Tucker conditions for mathematical programming problems in a Banach space," *SIAM Journal on Control* 7 (1969) 232–241.

[15a] M. Guignard, "Cahiers du Bureau Universitaire De Recherche Operationnelle, Cahier No. 14, "Conditions d'Optimalité et Dualité en Programmation Mathématique, Paris, 1970, (Available from Bureau Universitaire de Statistique de l'Université de Paris, Tour 45 55, 9, Quai Saint-Bernard, Paris Ve).

[16] G. Hadley, *Nonlinear and dynamic programming* (Addison-Wesley, Reading, Massachusetts, 1964).

[17] L. Hurwicz, "Programming in linear spaces," in: *Studies in linear and nonlinear programming*, Eds. K. Arrow, L. Hurwicz and H. Uzawa (Stanford University Press, Stanford, California, 1958).

[18] S. Karlin, *Mathematical methods and theory in games, programming, and economics* (Addison-Wesley, Reading, Massachusetts, 1959).

[19] H. Kuhn and A. Tucker, *Nonlinear programming, Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, Ed. J. Neyman (University of California Press, Berkeley, 1951) pp. 481–492.

[20] O. Mangasarian, *Nonlinear programming* (McGraw-Hill Book Co., New York, 1969).

[21] O. Mangasarian and S. Fromovitz, "The Fritz John necessary optimality conditions in the presence of equality and inequality constraints," *Journal of Mathematical Analysis and Applications* 17 (1967) 37–47.

[22] K. Ritter, "Optimization theory in linear spaces," *Mathematische Annalen* 184 (1970) 133–154.

[23] M. Simonnard, *Programmation lineaire* (Dunod, Paris, 1962) p. 388.

[24] M. Slater, "Lagrange multipliers revisited, A contribution to nonlinear programming," Cowles Commission Paper, Math. 403 (1950).

[25] P. Varaiya, "Nonlinear programming in Banach space," *SIAM Journal on Applied Mathematics* 15 (1967) 284–293.

[26] W. Zangwill, *Nonlinear programming, A unified approach* (Prentice Hall Publishers, Englewood Cliffs, N.J., 1969).

[27] S. Zlobec, "Asymptotic Kuhn-Tucker conditions for mathematical programming problems in a Banach space," *SIAM Journal on Control* 8 (1970) 505–512.