JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 40, 609–621 (1972)

Necessary Optimality Criteria in Mathematical Programming in the Presence of Differentiability

M. S. BAZARAA† AND J. J. GOODE‡

Georgia Institute of Technology, Atlanta, Georgia 30332

Submitted by Richard J. Duffin

We consider the problem of minimizing a function over a region defined by an arbitrary set, equality constraints, and constraints of the inequality type defined via a convex cone. Under some moderate convexity assumptions on the arbitrary set we develop optimality criteria of the minimum principle type which generalize the Fritz John optimality conditions. As a consequence of this result necessary optimality criteria of the saddle point type drop out. Here convexity requirements on the functions are relaxed to convexity at the point under investigation. We then present the weakest possible constraint qualification which insures positivity of the lagrangian multiplier corresponding to the objective function.

1. INTRODUCTION

The role of optimality criteria in mathematical programming is important both from theoretical and computational points of view. Perhaps the best known conditions for optimality are the Fritz John and the Kuhn–Tucker conditions.

Consider the problem, minimize $f(x)$ subject to $x \in X$, $g_i(x) \leqslant 0$ $(i = 1, 2, \dots, m)$ and $h_i(x) = 0$ $(i = 1, 2, \dots, k)$. Mangasarian and Fromovitz [1] showed that if $\bar{x} \in \operatorname{int} X$ (the interior of X) solves the above problem then there must exist a nonzero (u_0, u, v) such that

$$\text{(i)} \quad u_0 \nabla f(\bar{x}) + \sum_{i=1}^{m} u_i \nabla g_i(\bar{x}) + \sum_{i=1}^{k} v_i \nabla h_i(\bar{x}) = 0;$$

$$\text{(ii)} \quad u_0 \geqslant 0,\ u_i \geqslant 0,\ i = 1, 2, \dots, m;$$

$$\text{(iii)} \quad u_i g_i(\bar{x}) = 0,\ i = 1, 2, \dots, m.$$

This generalizes the original result of Fritz John without equality constraints. Furthermore, they showed that $u_0 > 0$ under certain additional assumptions and hence obtained the Kuhn–Tucker conditions for this problem.

† School of Industrial & Systems Engineering.
‡ School of Mathematics.

If $\bar{x} \notin \operatorname{int} X$ the above conditions are not valid, however. In this case conditions of the minimum principle type replace the above conditions. See for example, Canon, Cullum, and Polak [2], and Mangasarian [3]. In the infinite dimensional setting one may refer to Halkin and Neustadt [4], Neustadt [5], Varaiya [6], Guignard [7] as well as others.

In Section 2 we generalize the theorem of Mangasarian and Fromovitz (see also Mangasarian's book [3] for an extension). The assumption that $\bar{x} \in \operatorname{int} X$ is not required and we obtain optimality conditions of the minimum principle type with minimal convexity assumptions of the set X. Our necessary conditions are similar, but not equal to, the necessary conditions of Canon et al. [2].

As a result of the criteria developed we show that if $\bar{x}$ solves the minimization problem and if f and g are convex at $\bar{x}$ whereas h is linear then a corresponding saddle value problem possesses a solution. This generalizes a well-known result which requires global convexity of f and g. The reader may wish to refer to Kuhn and Tucker [8], Hurwicz and Uzawa [9], Mangasarian [3], and Neustadt [10].

We then present a constraint qualification which insures positivity of the lagrangian multiplier u_0 associated with f. It turns out that the qualification we construct is the weakest possible as discussed in Section 4. Under a proper specialization of the problem our qualification becomes equivalent to that of Gould and Tolle [11]. For further discussion on constraint qualifications, see Arrow, Hurwicz, and Uzawa [12], Mangasarian [3], Bazaraa et al. [13, 14], and Guignard [7].

2. The Main Result

Throughout the study we will use the following notation and terminology. Let E_n be the n-dimensional Euclidean space. If $x, y \in E_n$ then $x^t y$ gives the inner product of x and y where x^t denotes the transpose of x.

Let $f: Y \to E_m$ where Y is a nonempty open set in E_n. f is said to be differentiable at $\bar{x} \in Y$ if there exists an $n \times m$ matrix $\nabla f(\bar{x})$ and an m vector function o such that $f(\bar{x} + x) = f(\bar{x}) + x^t \nabla f(\bar{x}) + o(x)$ for each x with $\bar{x} + x \in Y$ where $o(x)/\| x \| \to 0$ as $\| x \| \to 0$.

If X is a nonempty subset of E_n then X^* denotes its polar, i.e., X^* is the set of all ξ such that $x^t \xi \leqslant 0$ for each $x \in X$. If $X = \emptyset$ then we will interpret X^* as E_n. It is obvious that X^* is a closed convex cone.

Let X be a nonempty convex set in E_n and $\bar{x} \in \operatorname{Cl} X$, the closure of X. In this study we will make use of the following two cones: the cone of interior directions for X at $\bar{x}$ which is denoted by $D(X, \bar{x})$, and the cone of tangents to X at $\bar{x}$ denoted by $T(X, \bar{x})$. These cones are defined below.

$$D(X, \bar{x}) = \{\xi : \text{there is a neighborhood } N \text{ of } \xi \text{ and a } \delta_0 > 0 \text{ such that } y \in N \text{ and } \delta \in (0, \delta_0) \text{ imply that } \bar{x} + \delta y \in X\};$$

$$T(X, \bar{x}) = \{\xi : \xi = \lim \lambda_n(x_n - \bar{x}), x_n \in X, \lambda_n > 0, x_n \to \bar{x}\}.$$

The cone of interior directions has been used by Dubovitskii and Milyutin [15] and was referred to as the set of permissible variations. The cone of tangents was introduced by Abadie [16] in the process of developing some constraint qualification for a nonlinear programming problem. The reader may refer to Bazaraa, Goode and Nashed [17] for discussion and properties of the cone of tangents.

It is obvious that $D(X, \bar{x})$ is an open cone and that $D(X, \bar{x}) \subset T(X, \bar{x})$. It is also obvious that both $D(X, \bar{x})$ and $T(X, \bar{x})$ are E_n if $\bar{x} \in \operatorname{int} X$. As will be seen later, the main result is put in terms of the polar of the cone of interior directions. In order to reduce our result to more familiar optimality criteria, we need the following proposition which characterizes the cone of interior directions when X is a convex set.

PROPOSITION. *Suppose that X is a convex set such that $D(X, \bar{x}) \neq \emptyset$. Then* $\operatorname{Cl} D(X, \bar{x}) = \operatorname{Cl} C(X, \bar{x})$ *where*

$$C(X, \bar{x}) = \{\xi : \textit{there is } \lambda > 0 \textit{ with } \bar{x} + \lambda\xi \in X\}.$$

Proof. It is obvious that $D(X, \bar{x}) \subset C(X, \bar{x})$. It can be easily shown that $C(X, \bar{x})$ is the minimal cone that contains $X - \bar{x}$. Furthermore, since $D(X, \bar{x})$ is open and nonempty by assumption then $\operatorname{int} C(X, \bar{x}) \neq \emptyset$ and hence $\operatorname{int} X \neq \emptyset$. We will show that $\operatorname{int} X - \bar{x} \subset D(X, \bar{x})$. Now let $x \in \operatorname{int} X$ then there is a neighborhood N of 0 such that $x + N \subset \operatorname{int} X$. If $y \in x - \bar{x} + N$ and $\delta \in (0, 1)$ we have

$$\bar{x} + \delta y = (1 - \delta)\,\bar{x} + \delta(\bar{x} + y).$$

But $\bar{x} + y \in x + N \subset \operatorname{int} X$ and $\bar{x} \in \operatorname{Cl} X$ and so $\bar{x} + \delta y \in \operatorname{int} X$ for each $\delta \in (0, 1)$ since X is convex. This shows that $x - \bar{x} \in D(X, \bar{x})$, i.e., $\operatorname{int} X - \bar{x} \subset D(X, \bar{x})$. By convexity of X and since $\operatorname{int} X \neq \emptyset$ then $X - \bar{x} \subset \operatorname{Cl} X - \bar{x} \subset \operatorname{Cl} D(X, \bar{x})$. But since $C(X, \bar{x})$ is the minimal cone that contains $X - \bar{x}$ then we have $C(X, \bar{x}) \subset \operatorname{Cl} D(X, \bar{x})$. This together with the fact that $D(X, \bar{x}) \subset C(X, \bar{x})$ implies that $\operatorname{Cl} C(X, \bar{x}) = \operatorname{Cl} D(X, \bar{x})$.

We may add that if X is convex then $T(X, \bar{x}) = \operatorname{Cl} C(X, \bar{x})$. See, for example, [17]. Hence in view of the above proposition it is clear that $T(X, \bar{x}) = \operatorname{Cl} D(X, \bar{x})$ when X is convex as long as $D(X, \bar{x})$ is not empty. Since convexity of X implies convexity of $C(X, \bar{x})$ then it also implies convexity of $\operatorname{Cl} D(X, \bar{x})$ and $T(X, \bar{x})$. It can be shown that if X is convex then $D(X, \bar{x})$ is also convex.

Consider the problem P : Minimize $\{f(x) : x \in X, g(x) \in \operatorname{Cl} C, h(x) = 0\}$. Here f, g, and h are functions defined on an open set containing X and taking values, respectively, in E_1, E_m, and E_k. C is a nonempty open cone in E_m. The following theorem gives necessary conditions for $\bar{x}$ to solve the above problem. We will postpone the proof of the theorem until the next section. It is worthwhile mentioning that our theorems hold if $\bar{x}$ is a local optimal rather than a global optimal. We choose not to take this point of view in an explicit form in order to simplify notation.

THEOREM 1. *Suppose that $\bar{x}$ solves problem P and suppose that $D(X, \bar{x})$ is convex. If f and g are differentiable at $\bar{x}$ and h is differentiable in a neighborhood of $\bar{x}$ then there exists a nonzero $(u_0, u, v) \in E_1 \times E_m \times E_k$ such that*

(i) $-[\nabla f(\bar{x})\, u_0 + \nabla g(\bar{x})\, u + \nabla h(\bar{x})\, v] \in D^*(X, \bar{x})$

(ii) $u_0 \geqslant 0$, $u \in C^*$, *and*

(iii) $u^t g(\bar{x}) = 0$.

The conditions above may be viewed as generalized Fritz John conditions for optimality. The following remarks may be helpful for the reader:

1. u_0, u, and v can be viewed as lagrangian multipliers and (iii) can be viewed as complementarity slackness condition. In the special case when C is the nonnegative orthant then (ii) and (iii) together reduce to the following well known fact. If $\bar{x}$ solves the problem then either $g_i(\bar{x}) = 0$ or else $u_i = 0$ where $g = (g_1, g_2, \ldots, g_m)$ and $u = (u_1, u_2, \ldots, u_m)$.

2. In the hypothesis of the theorem we require convexity of $D(X, \bar{x})$. This is a considerably weaker hypothesis than convexity of X. For example, let $X = \{(x, y) : y \geqslant x^3\}$ then $D(X, \bar{x}) = \{(x, y) : y > 0\}$. Indeed X is not convex whereas $D(X, \bar{x})$ is. If X is convex and has a nonempty interior then by the proposition it is clear that $D(X, \bar{x})$ is convex and not empty and the hypotheses of the theorem hold. Also note that no convexity-concavity requirements of any of the functions are needed.

3. If X is convex and has nonempty interior then, in view of the proposition, condition (i) reduces to: $(x - \bar{x})^t\, [\nabla f(\bar{x})\, u_0 + \nabla g(\bar{x})\, u + \nabla h(\bar{x})\, v] \geqslant 0$ for each $x \in X$. This inequality may be viewed as a condition of the minimum principle type which generalizes that of Mangasarian [3]. Conditions of this type can be found in Halkin and Neustadt [4], Canon et al. [2], and Neustadt [5] as well as others.

4. If $\bar{x} \in \operatorname{int} X$ then $D(X, \bar{x}) = E_n$ and $D^*(X, \bar{x}) = 0$. Then condition (i) becomes $\nabla f(\bar{x})\, u_0 + \nabla g(\bar{x})\, u + \nabla h(\bar{x})\, v = 0$. This with conditions (ii) and (iii) give a generalized form of the Fritz John conditions. If C is the nonnegative orthant in E_m then we precisely get the Fritz John conditions discussed in the introduction.

5. In condition (i) of Theorem 1 if we can replace $D(X, \bar{x})$ by the larger cone $T(X, \bar{x})$, i.e., if condition (i) reads

$$-[\nabla f(\bar{x})\, u_0 + \nabla g(\bar{x})\, u + \nabla h(\bar{x})\, v] \in T^*(X, \bar{x}),$$

then the theorem will be sharper. However, this sharper result does not hold in general. For example consider the problem: minimize $\{f(x) : x \in X, h(x) = 0\}$ where $X = \{(x_1, x_2) : x_1 \text{ and } x_2 \text{ are rational}\}$, $f(x_1, x_2) = x_2$, and $h(x_1, x_2) = x_2 - \sqrt{2}\, x_1$. It is clear that $X \cap \{x : h(x) = 0\} = \{(0, 0)\}$ so the only admissible point is the origin and hence $\bar{x} = (0, 0)$ solves the above problem. It is clear that $T(X, \bar{x}) = E_2$ and hence $T^*(X, \bar{x}) = \{(0, 0)\}$. Condition (i), however, does not hold for a nonzero (u_0, v). This shows that we cannot strengthen the conclusion of the theorem by replacing $D(X, \bar{x})$ with $T(X, \bar{x})$. One should note that in the above example $D(X, \bar{x})$ is empty and hence the theorem holds trivially.

6. The lagrangian multiplier u_0 is not necessarily positive. In many cases one would like to insure positivity of u_0 and hence some additional assumptions are required. These are called constraint qualifications and we present the weakest possible qualification in Section 4.

7. As a result of Theorem 1 we will be able to develop a generalization of the saddle point optimality criteria with weaker assumptions. The saddle value problem SP can be stated as follows. Find $(\bar{x}, u_0, \bar{u}, \bar{v})$ such that $\psi(\bar{x}, u_0, u, v) \leqslant \psi(\bar{x}, u_0, \bar{u}, \bar{v}) \leqslant \psi(x, u_0, \bar{u}, \bar{v})$ for each $x \in X$, $u \in C^*$, $v \in E_k$ where $\psi(x, u_0, u, v) = u_0 f(x) + u^t g(x) + v^t h(x)$. Here $\bar{x} \in X$, u_0 is a nonnegative scalar, $\bar{u} \in C^*$, $\bar{v} \in E_k$ and $(u_0, \bar{u}, \bar{v}) \neq 0$. This problem has been considered by different authors in the case when C is the nonnegative orthant in E_m. See, for example, Mangasarian [3]. The problem has also been considered in an infinite dimensional setting using cones by Hurwicz and Uzawa [9] and Neustadt [10]. It is well known (see the mentioned references) that under convexity of f, C-convexity of g, and linearity of f one can claim that if $\bar{x}$ solves the minimization problem P then there must exist $(u_0, \bar{u}, \bar{v})$ such that $(\bar{x}, u_0, \bar{u}, \bar{v})$ solves the saddle value problem SP. The method of proving the above result hinges about the existence of separating hyperplanes between nonintersecting convex sets. We will make use of the above theorem in partially relaxing convexity of f and C-convexity of g. This is given by Theorem 2 below. First we need the following definition of convexity at a point due to Mangasarian [3] and C-convexity at a point which generalizes the notion of C-convexity on a convex set. See, for example, Eisenberg [18].

DEFINITION. Let $\alpha : X \to E_m$ where X is a nonempty convex set in E_n. Let K be a convex cone in E_m. Then α is said to be *K-convex at* $\bar{x} \in X$ if $x \in X$ and $\lambda \in (0, 1)$ implies that $\lambda\alpha(x) + (1 - \lambda)\, \alpha(\bar{x}) - \alpha(\lambda x + (1 - \lambda)\, \bar{x}) \in \operatorname{Cl} K$.

α is said to be K-convex if α is K-convex for each $\bar{x} \in X$. If $m = 1$ and K is the set of nonnegative reals then we get definitions of convexity of α at $\bar{x}$ and convexity of α on X.

THEOREM 2. *Suppose that $\bar{x}$ solves problem P and suppose that X is a convex set with nonempty interior. Further suppose that f is convex at $\bar{x}$, g is C-convex at $\bar{x}$, and h is linear. Then there exists a nonzero $(u_0, \bar{u}, \bar{v})$ with $u_0 \geqslant 0$, $\bar{u} \in C^*$ such that $(\bar{x}, u_0, \bar{u}, \bar{v})$ solves* SP.

Proof. From Theorem 1 and following remarks there exists a nonzero $(u_0, \bar{u}, \bar{v})$ such that $u_0 \geqslant 0$, $\bar{u} \in C^*$, $\bar{u}^t g(\bar{x}) = 0$ and

$$(x - \bar{x})^t [\nabla f(\bar{x}) u_0 + \nabla g(\bar{x}) \bar{u} + \nabla h(\bar{x}) \bar{v}] \geqslant 0$$

for each $x \in X$. Since $u_0 \geqslant 0$ then by convexity of f at $\bar{x}$ we get

$$u_0 f(x) \geqslant u_0 f(\bar{x}) + u_0(x - \bar{x})^t \nabla f(\bar{x}).$$

By C-convexity of g at $\bar{x}$ it follows that

$$g(\lambda\bar{x} + (1 - \lambda) x) - \lambda g(\bar{x}) - (1 - \lambda) g(x) \in \operatorname{Cl} C$$

for each $x \in X$ and each $\lambda \in (0, 1)$. Since $\bar{u} \in C^*$ it then follows that

$$\bar{u}^t g(\bar{x} + (1 - \lambda)(x - \bar{x})) - \bar{u}^t g(\bar{x}) \leqslant (1 - \lambda) \bar{u}^t (g(x) - g(\bar{x}))$$

for each $x \in X$ and $\lambda \in (0, 1)$. Dividing by $1 - \lambda$ and letting $\lambda \to 1$ we get $(x - \bar{x})^t \nabla g(\bar{x}) \bar{u} \leqslant \bar{u}^t (g(x) - g(\bar{x}))$ for each $x \in X$. Finally by linearity of h we get $h(x) = h(\bar{x}) + \nabla^t h(\bar{x})(x - \bar{x})$. Combining the inequalities corresponding to f and g, the last equality and the inequality of Theorem 1 we get

$$\begin{aligned} &u_0(f(x) - f(\bar{x})) + \bar{u}^t(g(x) - g(\bar{x})) + \bar{v}^t(h(x) - h(\bar{x})) \\ &\quad \geqslant (x - \bar{x})^t [\nabla f(\bar{x}) u_0 + \nabla g(\bar{x}) \bar{u} + \nabla h(\bar{x}) \bar{v}] \geqslant 0 \end{aligned}$$

for each $x \in X$. Rearranging the terms we get $\psi(x, u_0, \bar{u}, \bar{v}) \geqslant \psi(\bar{x}, u_0, \bar{u}, \bar{v})$ for each $x \in X$ where $\psi(x, u_0, u, v) = u_0 f(x) + u^t g(x) + v^t h(x)$. Noting that $h(\bar{x}) = 0$, $g(\bar{x}) \in \operatorname{Cl} C$, $\bar{u}^t g(\bar{x}) = 0$ it follows that $\psi(\bar{x}, u_0, u, v) \leqslant \psi(\bar{x}, u_0, \bar{u}, \bar{v})$ for each $u \in C^*$, $v \in E_k$. This shows that $(\bar{x}, u_0, \bar{u}, \bar{v})$ solves the saddle value problem and the proof is complete.

We would like to mention that if $(\bar{x}, u_0, \bar{u}, \bar{v})$ solves the saddle value problem and $u_0 > 0$ then $\bar{x}$ solves problem P. No convexity of any kind is required here. The proof of this statement is straight forward and is hence omitted.

3. Proof of Theorem 1

The proof of the theorem is based on the following three lemmas.

LEMMA 1. *Suppose that Y is a convex set in E_n, K is a convex cone in E_m, and A is an $m \times n$ matrix. If the system $Ax \in K$ has no solution $x \in Y$ then there exists a nonzero $u \in K^*$ such that $u^t Ax \geqslant 0$ for each $x \in Y$.*

Proof. The set $A(Y) = \{Ax : x \in Y\}$ is convex and by hypothesis $A(Y) \cap K$ is empty. By the fundamental separation theorem for convex sets (see, for example, [3]) there is a nonzero u such that $u^t Ax \geqslant u^t k$ for each $x \in Y$ and $k \in \mathrm{Cl}\, K$. Since $0 \in \mathrm{Cl}\, K$ then $u^t Ax \geqslant 0$ for each $x \in Y$. It is also clear that $u^t k \leqslant 0$ since otherwise $u^t Ax \geqslant \lambda u^t k_0$ for each $\lambda > 0$ where $u^t k_0 > 0$, which is impossible. This completes the proof.

As will be seen in Lemma 3 we will show that if $\bar{x}$ solves the minimization problem P then a certain system does not have a solution. This system can be represented in the form $Ax \in K$, $x \in Y$. We then apply the above lemma to obtain the theorem. The following lemma will enable us to obtain the complementarity slackness condition (iii) of Theorem 1. First we need the following definition.

DEFINITION. Let K be an open convex cone in E_m and let $b \in \mathrm{Cl}\, K$. Then $K_b = \{c - \lambda b : c \in K, \lambda \geqslant 0\}$.

Note that K_b is an open convex cone and that $K \subset K_b$. Also the line $\{\lambda b : \lambda \text{ is real}\} \subset \mathrm{Cl}\, K_b$. This very fact gives us the complementarity slackness condition. Neustadt [5] used this cone in developing his general necessary conditions for optimality. The following lemma essentially says that K_b is the cone of interior directions for K at b.

LEMMA 2. *If $a \in K_b$ then $b + \delta a + o(\delta) \in K$ for $\delta > 0$ sufficiently small, where $o(\delta)/\delta \to 0$ as $\delta \to 0$.*

Proof. Let $a = c - \lambda b \in K_b$ where $c \in K$ and $\lambda \geqslant 0$. Then

$$b + \delta a + o(\delta) = (1 - \lambda\delta)\, b + \delta \left(c + \frac{o(\delta)}{\delta}\right).$$

Since K is open and $c \in K$ then for δ sufficiently small we have $c + o(\delta)/\delta \in K$. Also for δ sufficiently small $(1 - \lambda\delta) = \mu \geqslant 0$. Hence for small enough $\delta > 0$,

$$b + \delta a + o(\delta) = (1 - \lambda\delta)\, b + \delta \left(c + \frac{o(\delta)}{\delta}\right) \in \mathrm{Cl}\, K + K.$$

But $\mathrm{Cl}\, K + K = K$ and hence $b + \delta a + o(\delta) \in K$, and the proof is complete.

The following lemma can be viewed as a generalized linearization lemma. The lemma states that if a nonlinear system has no solution then there is a corresponding linear system which also has no solution. We can then make use of Lemma 1. The lemma below is a generalization of a result of Mangasarian and Fromovitz [1] (see also [3]). The method of proof is essentially based on the implicit function theorem as in [3].

LEMMA 3. *Let X be a subset of E_n and K be an open convex cone in E_m. Let α and h be defined on an open set containing X and taking values in E_m and E_k, respectively. Suppose that $\bar{x} \in X$ such that $\alpha(\bar{x}) \in \operatorname{Cl} K$ and $h(\bar{x}) = 0$. Further suppose that α and h are differentiable at $\bar{x}$ and $\nabla h(\bar{x})$ has rank k. If $\alpha(x) \in K$, $h(x) = 0$ has no solution in X then $\nabla^t \alpha(\bar{x})\, \xi \in K_{\alpha(\bar{x})}$, $\nabla^t h(\bar{x})\, \xi = 0$ has no solution in $D(X, \bar{x})$.*

Proof. Since $\nabla h(\bar{x})$ has rank k then $k \leqslant n$. The case where $k = 0$ corresponds to the absence of equality constraints and is easy. If $k = n$ then the only solution to $\nabla^t h(\bar{x})\, \xi = 0$ is $\xi = 0$. Since $D(X, \bar{x})$ is open, then $0 \notin D(X, \bar{x})$ and the result is immediate. Thus we assume that $0 < k < n$. By the implicit function theorem (see, for example, [19]) the following exist:

(i) a partition of $x^t = (x_1{}^t, x_2{}^t)$ with $x_1 \in E_{n-k}$, $x_2 \in E_k$ and $\nabla^t h(\bar{x}) = (\nabla_1{}^t h(\bar{x}), \nabla_2{}^t h(\bar{x}))$ where $\nabla_2 h(\bar{x})$ is nonsingular;

(ii) an open set Ω in E_{n-k} containing $\bar{x}_1$ where $\bar{x} = (\bar{x}_1, \bar{x}_2)$;

(iii) a unique function e on Ω with values in E_k such that $\bar{x}_2 = e(\bar{x}_1)$, e is differentiable on Ω and $h(x_1, e(x_1)) = 0$ for each $x_1 \in \Omega$.

We will prove the lemma by showing that if $\nabla^t \alpha(\bar{x})\, \xi \in K_{\alpha(\bar{x})}$, $\nabla^t h(\bar{x})\, \xi = 0$ has a solution in $D(X, \bar{x})$ then $\alpha(x) \in K$, $h(x) = 0$ has a solution in X. Let ξ be solution to the first system and let $\xi^t = (\xi_1{}^t, \xi_2{}^t)$ be the partition of (i) above. Then $\nabla_1{}^t h(\bar{x})\, \xi_1 + \nabla_2{}^t h(\bar{x})\, \xi_2 = 0$. Since $h(x_1, e(x_1)) = 0$ for each $x_1 \in \Omega$ and $\bar{x}_1 \in \Omega$ then $\nabla_1 h(\bar{x}) + \nabla e(\bar{x}_1)\, \nabla_2 h(\bar{x}) = 0$ and hence $\nabla_1{}^t h(\bar{x})\, \xi_1 + \nabla_2{}^t h(\bar{x})\, \nabla^t e(\bar{x}_1)\, \xi_1 = 0$. Since $\nabla_2 h(\bar{x})$ is nonsingular then the last equation and $\nabla_1{}^t h(\bar{x})\, \xi_1 + \nabla_2{}^t h(\bar{x})\, \xi_2 = 0$ imply that $\xi_2 = \nabla^t e(\bar{x}_1)\, \xi_1$. By differentiability of e at $\bar{x}_1$ we have

$$e(\bar{x}_1 + \delta\xi_1) = e(\bar{x}_1) + \delta \nabla^t e(\bar{x}_1)\, \xi_1 + \delta \cdot \epsilon_1(\delta).$$

Here $\epsilon_1(\delta)$ is a vector-valued function of δ such that $\epsilon_1(\delta) \to 0$ as $\delta \to 0$. Thus

$$e(\bar{x}_1 + \delta\xi_1) = \bar{x}_2 + \delta\xi_2 + \delta \cdot \epsilon_1(\delta).$$

Then

$$(\bar{x}_1 + \delta\xi_1, e(\bar{x}_1 + \delta\xi_1)) = \bar{x} + \delta\xi + \delta \cdot \epsilon_2(\delta)$$

where $\epsilon_2{}^t(\delta) = (0, \epsilon_1{}^t(\delta))$. Now since Ω is open and $\bar{x}_1 \in \Omega$, then for δ sufficiently small we have $\bar{x}_1 + \delta\xi_1 \in \Omega$ and hence $h(\bar{x}_1 + \delta\xi_1, e(\bar{x}_1 + \delta\xi_1)) = 0$. Since α is differentiable at $\bar{x}$ we have

$$\begin{aligned}\alpha(\bar{x}_1 + \delta\xi_1, e(\bar{x}_1 + \delta\xi_1)) &= \alpha(\bar{x} + \delta\xi + \delta\epsilon_2(\delta)) \\ &= \alpha(\bar{x}) + \delta\nabla^t\alpha(\bar{x})\,\xi + \delta\cdot\epsilon_3(\delta),\end{aligned}$$

where $\epsilon_3(\delta) \to 0$ as $\delta \to 0$. Since by assumption $\nabla^t\alpha(\bar{x})\,\xi \in K_{\alpha(\bar{x})}$, we conclude by Lemma 2 that $\alpha(\bar{x}_1 + \delta\xi_1, e(\bar{x}_1 + \delta\xi_1)) \in K$ for δ sufficiently small. So far we have shown that $h(\bar{x} + \delta\xi + \delta\cdot\epsilon_2(\delta)) = 0$ and $\alpha(\bar{x} + \delta\xi + \delta\cdot\epsilon_2(\delta)) \in K$ for δ sufficiently small. Finally since $\zeta \in D(X, \bar{x})$, then for $\delta > 0$ small enough we have $\bar{x} + \delta\xi + \delta\cdot\epsilon_2(\delta) \in X$. Therefore it follows that $(\bar{x}_1 + \delta\xi_1, e(\bar{x}_1 + \delta\xi_1))$ solves the system $\alpha(x) \in K$, $h(x) = 0$, and $x \in X$ for $\delta > 0$ sufficiently small and the proof is complete.

Proof of Theorem 1. If $D(X, \bar{x})$ is empty, then $D^*(X, \bar{x}) = E_n$ and the theorem holds trivially. Without loss of generality assume that the rank of $\nabla h(\bar{x})$ is k, since otherwise the theorem holds trivially. Now suppose that $\bar{x}$ solves P, i.e., $x \in X$, $g(x) \in \mathrm{Cl}\, C$ and $h(x) = 0$ imply that $f(x) \geqslant f(\bar{x})$. Let $\alpha(x) = (f(x) - f(\bar{x}), g(x))$ and $K = R_- \times C$ where R_- is the set of negative real numbers. Note that $\bar{x} \in X$, $\alpha(\bar{x}) \in \mathrm{Cl}\, K$ and $h(\bar{x}) = 0$. Also note that the system $\alpha(x) \in K$, $h(x) = 0$ has no solution in X. Therefore by Lemma 3 the system $\nabla^t\alpha(\bar{x})\,\xi \in K_{\alpha(\bar{x})}$, $\nabla^t h(\bar{x})\,\xi = 0$ has no solution in $D(X, \bar{x})$. By Lemma 1 we conclude that there is a nonzero $(q, v) \in K^*_{\alpha(\bar{x})} \times E_k$ such that $q\nabla^t\alpha(\bar{x})\,\xi + v\nabla^t h(\bar{x})\,\xi \geqslant 0$ for each $\xi \in D(X, \bar{x})$. Therefore

$$-[\nabla\alpha(\bar{x})\,q + \nabla h(\bar{x})\,v] \in D^*(X, \bar{x}).$$

Noting that $K_{\alpha(\bar{x})} \supset K$ then $K^*_{\alpha(\bar{x})} \subset K^*$ and so $q \in K^*$, i.e., $q = (u_0, u)$ with $u_0 \geqslant 0$ and $u \in C^*$. Finally since both $\alpha(\bar{x})$ and $-\alpha(\bar{x})$ belong to $K_{\alpha(\bar{x})}$ and $q \in K^*_{\alpha(\bar{x})}$ then $u^t g(\bar{x}) = 0$. To summarize there exist a nonzero (u_0, u, v) such that $u_0 \geqslant 0$, $u \in C^*$, $u^t g(\bar{x}) = 0$, and

$$-[\nabla f(\bar{x})\,u_0 + \nabla g(\bar{x})\,u + \nabla h(\bar{x})\,v] \in D^*(X, \bar{x})$$

and the proof is complete.

4. Constraint Qualification

In reference to Theorem 1 of Section 2, one may note that u_0, the lagrangian multiplier associated with the objective function, is not necessarily positive. In many cases, however, one would like to obtain positivity of u_0.

For instance, if u_0 is positive, X is convex, f is convex at $\bar{x}$, g is C-convex at $\bar{x}$, and h is linear, then the conditions of Theorem 1 assure that $\bar{x}$ solves problem P. This fact is clear from Theorem 2 and the discussion following it.

If $u_0 > 0$ then without loss of generality u_0 can be assumed equal to one, and Theorem 1 then gives a generalized form of the Kuhn–Tucker conditions. In order to insure that $u_0 > 0$ we need some additional assumptions. These assumptions are referred to in the literature as constraint qualifications since they do not involve the objective function.

We would like to emphasize an important fact which is sometimes ignored. When one speaks of a constraint qualification, one means a condition which insures positivity of the lagrangian multiplier associated with any objective function having a constrained minimum at the point under investigation. One may find a condition that will work for a particular function. However, such condition is not regarded as a constraint qualification. This point has been emphasized by Gould and Tolle [11]. To make our discussion more precise we give the following definition which extends the notion of Lagrange regularity given in [11].

DEFINITION. The constraint set $S = \{x \in X : g(x) \in \operatorname{Cl} C, h(x) = 0\}$ is said to be *Lagrange regular at* $\bar{x} \in S$ if for every objective function f having a minimum over S at $\bar{x}$, there exist (u, v) such that

(i) $-[\nabla f(x_0) + \nabla g(\bar{x})\, u + \nabla h(\bar{x})\, v] \in D^*(X, \bar{x})$, and

(ii) $u \in C^*$, $u^t g(\bar{x}) = 0$.

Note that these conditions are precisely those of Theorem 1 with $u_0 = 1$. As mentioned earlier, the above criteria can be viewed as a generalization of the Kuhn–Tucker conditions.

We will present a constraint qualification which is both necessary and sufficient for Lagrange regularity, i.e., a qualification which is both necessary and sufficient for the validation of the Kuhn–Tucker criteria of the above definition. We would like to emphasize that one will get different necessary and sufficient qualifications according to the type of Lagrange regularity (or Kuhn–Tucker conditions) one poses. See Gould and Tolle [11] and Bazaraa et al. [14] for various necessary and sufficient constraint qualifications. Now consider the following constraint qualification:

$$T^*(S, \bar{x}) \subset D^*(X, \bar{x}) + C_g + C_h,$$

where

$$C_g = \{\xi : \xi = \nabla g(\bar{x})\, u, u \in C^*, u^t g(\bar{x}) = 0\},$$

and

$$C_h = \{\xi : \xi = \nabla h(\bar{x})\, v, v \in E_k\}.$$

It is clear that C_g and C_h are nonempty closed convex cones. C_h is in fact a subspace, namely, the range space of $\nabla h(\bar{x})$. Since the inequality constraints are defined via a cone, we cannot talk about the binding constraints corresponding to the point $\bar{x}$. In some sense the requirement that $u^t g(\bar{x}) = 0$ replaces this notion here. To carry the discussion further, suppose that C is the negative orthant in E_m and let $I = \{i : g_i(\bar{x}) = 0\}$, i.e., I is the set of binding constraints. Then

$$C_g = \left\{ \xi : \xi = \sum_{i \in I} u_i \nabla g_i(\bar{x}),\ u_i \geqslant 0 \text{ for each } i \in I \right\}.$$

Note that in this case C_g is the polyhedral cone generated by the gradients of the binding constraints.

In order to develop the main result of this section we need the following two remarks. For a proof of Remark 1 one may refer to Varaiya [6]. The reader may refer to Gould and Tolle [11] or Bazaraa et al. [14] for a proof of Remark 2.

Remark 1. If $\bar{x}$ solves the problem: minimize $f(x)$ subject to $x \in S$ then $-\nabla f(\bar{x}) \in T^*(S, \bar{x})$.

Remark 2. Given a nonzero $y \in T^*(S, \bar{x})$ there is a function f which is differentiable at $\bar{x}$ with $y = -\nabla f(\bar{x})$ and such that f has a minimum over S at $\bar{x}$.

THEOREM 3. *S is Lagrange regular at $\bar{x}$ if and only if the constraint qualification holds.*

Proof. Suppose that the constraint qualification holds, i.e.,

$$T^*(S, \bar{x}) \subset D^*(X, \bar{x}) + C_g + C_h .$$

Let f be differentiable at $\bar{x} \in S$ and having a minimum over S at $\bar{x}$. Then by Remark 1 and the constraint qualification it follows that

$$-\nabla f(\bar{x}) \in T^*(S, \bar{x}) \subset D^*(X, \bar{x}) + C_g + C_h .$$

In other words there exists (u, v) such that $u \in C^*$, $u^t g(\bar{x}) = 0$ and $-[\nabla f(\bar{x}) + \nabla g(\bar{x})\, u + \nabla h(\bar{x})\, v] \in D^*(X, \bar{x})$. This shows that S is Lagrange regular at $\bar{x}$. Conversely, suppose that S is Lagrange regular at $\bar{x}$ and let $y \in T^*(S, \bar{x})$. By Remark 2 we can find a function f having a minimum over S at $\bar{x}$ and $-\nabla f(\bar{x}) = y$. Then by Lagrange regularity we get

$$y = -\nabla f(\bar{x}) \in D^*(X, \bar{x}) + C_g + C_h ,$$

i.e., $T^*(S, \bar{x}) \subset D^*(X, \bar{x}) + C_g + C_h$ and the proof is complete.

The above theorem may be reworded as follows: The constraint qualification $T^*(S, \bar{x}) \subset D^*(X, \bar{x}) + C_g + C_h$ is the weakest possible for the validation of the Kuhn–Tucker criteria of the type given in the above definition. In particular, if $\bar{x} \in \operatorname{int} X$ and C is the negative orthant in E_m then $D^*(X, \bar{x}) = \{0\}$ and the constraint qualification reduces to $T^*(S, \bar{x}) \subset C_g + C_h$. This is precisely the qualification introduced by Gould and Tolle [11]. This shows that each of the known constraint qualifications implies the mentioned condition. As a matter of fact, this is not hard to show in an explicit manner. In the absence of equality constraints the reader may refer to Bazaraa et al. [13] for the implications of existing qualifications and their relationships with the weakest possible.

References

1. O. L. Mangasarian and S. Fromovitz, The Fritz John necessary optimality conditions in the presence of equality and inequality constraints, *J. Math. Anal. Appl.* **17** (1967), 37–47.
2. M. D. Canon, C. D. Cullum, and E. Polak, "Theory of Optimal Control and Mathematical Programming," McGraw–Hill, New York, 1970.
3. O. L. Mangasarian, "Nonlinear Programming," McGraw–Hill, New York, 1969.
4. H. Halkin and L. W. Neustadt, General Necessary Conditions for Optimization Problems, *Proc. Nat. Acad. Sci. U.S.A.* **56** (1966), 1066–1071.
5. L. W. Neustadt, A general theory of extremals, *J. Comput. Syst. Sci.* **3** (1969), 57–92.
6. P. P. Varaiya, Nonlinear programming in Banach space, *SIAM J. Appl. Math.* **15** (1967), 284–293.
7. M. Guignard, Generalized Kuhn–Tucker conditions for mathematical programming problems in a Banach space, *SIAM J. Control* **7** (1969), 232–241.
8. H. W. Kuhn and A. W. Tucker, Nonlinear programming, *in* "Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability," Vol. 5, University of California Press, Berkeley, CA, 1951.
9. L. Hurwicz and H. Uzawa, A note on the Lagrangian saddle points, *in* "Studies in Linear and Nonlinear Programming" (K. J. Arrow, L. Hurwicz, and H. Uzawa, Eds.), Stanford University Press, Stanford, CA, 1958.
10. L. W. Neustadt, Sufficiency conditions and a duality theory for mathematical programming problems in arbitrary linear spaces, *in* "Nonlinear Programming" (J. B. Rosen, O. L. Mangasarian, and K. Ritter, Eds.), Academic Press, New York, 1970.
11. F. J. Gould and J. W. Tolle, A necessary and sufficient qualification for constrained optimization, *SIAM J. Appl. Math.* **20** (1971), 164–172.
12. K. J. Arrow, L. Hurwicz, and H. Uzawa, Constraint qualifications in maximization problems, *Naval Res. Logistics Quart.* **8** (1961), 175–191.
13. M. S. Bazaraa, J. J. Goode, and C. M. Shetty, Constraint qualifications revisited, *Management Sci.* **18** (1972), 567–573.
14. M. S. Bazaraa, J. J. Goode, M. Z. Nashed, and C. M. Shetty, Nonlinear Programming without Differentiability in Banach Spaces: Necessary and Sufficient Constraint Qualification, in *J. Applicable Anal.*, to appear.

15. A. Ya. Dubovitskii and A. A. Milyutin, Extremum problems in the presence of constraints, *Zh. Vychisl. Mat. i Mat. Fiz.* **5** (1965), 395–453.
16. J. Abadie, On the Kuhn–Tucker theorem, *in* "Nonlinear Programming" (J. Abadie, Ed.), North Holland, Amsterdam, 1967.
17. M. S. Bazaraa, J. J. Goode, and M. Z. Nashed, On the cones of tangents with applications, *J. Optimization Theory Appl.*, to appear.
18. E. Eisenberg, On cone functions, *in* "Recent Advances in Mathematical Programming" (R. L. Graves and P. Wolfe, Eds.), McGraw–Hill, New York, 1963.
19. W. Rudin, "Principles of Mathematical Analysis," McGraw–Hill, New York, 1964.