JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS **80**, 545–550 (1981)

On Sufficiency of the Kuhn–Tucker Conditions

MORGAN A. HANSON

Department of Statistics, Florida State University, Tallahassee, Florida 32306; and Department of Pure Mathematics, La Trobe University, Bundoora 3083, Australia

Submitted by E. S. Lee

1. INTRODUCTION

Let $f_0(x), f_1(x),..., f_m(x)$ be differentiable functions defined on a set $C \subset E^n$, and let $f(x)$ denote the vector $(f_1(x),..., f_m(x))'$. Consider the problem

$$\min_{x \in C} f_0(x) \tag{1.1}$$

$$\text{subject to} \quad f(x) \leqslant 0. \tag{1.2}$$

According to the Kuhn–Tucker theorem it is necessary, under certain constraint qualifications, that for x_0 to be minimal in this problem, there exists a vector $v_0 \in E^m$ such that

$$\nabla f_0(x_0) + \nabla v_0' f(x_0) = 0, \tag{1.3}$$

$$v_0' f(x_0) = 0, \tag{1.4}$$

and

$$v_0 \geqslant 0. \tag{1.5}$$

If all the functions $f_0(x), f_1(x),..., f_m(x)$ are convex on C then these conditions are also sufficient.

Various classes of functions have been defined for the purpose of weakening this limitation of convexity in mathematical programming. Mangasarian [1] has speculated that pseudo-convexity of $f_0(x)$ and quasi-convexity of $f(x)$ are the weakest conditions that can be imposed so that the above conditions are sufficient for optimality. It will be shown that there are other wide classes of functions for which the conditions are sufficient.

0022-247X/81/040545–06$02.00/0

409/80/2-17

A dominant feature in the use of convexity is that local optimality implies global optimality; and consequently it may appear that the local nature of the differential calculus may be an inhibiting factor in generalizing too far away from convexity. Another much used property of convex functions is that they are always bounded on one side by their tangent hyperplanes at any point, which facilitates the use of linear bounds and approximations.

The use of nonlinear bounds will be considered here. From the definition of a convex differentiable function $\phi(x)$ on C,

$$\phi(x_1) - \phi(x_2) \geqslant (x_1 - x_2)' \nabla\phi(x_2) \qquad \text{for all} \quad x_1, x_2 \in C,$$

it is suggested by the generalization of Taylor's expansion by Burmann [see 2] (being an expansion of a function in powers of another given function, rather than simply in powers of x) that we consider a class of functions, with prescribed nonlinear bounds, given by the following definition:

$$\phi(x_1) - \phi(x_2) \geqslant \eta'(x_1, x_2) \nabla\phi(x_2) \qquad \text{for all} \quad x_1, x_2 \in C, \tag{1.6}$$

for some arbitrary given vector function $\eta(x_1, x_2)$ defined on $C \times C$. Note that the sum of any number of funtions satisfying (1.6) also satisfies (1.6), a property that is lacking in quasi-convex and pseudo-convex functions.

2. Sufficiency

THEOREM 2.1. *Let $f_0(x), f_1(x), \ldots, f_m(x)$ be differentiable functions on $C \subset E^n$ satisfying* (1.6) *for some $\eta(x_1, x_2)$ on $C \times C$. If there exist $x_0 \in C$ and $v_0 \in E^m$ satisfying the Kuhn–Tucker conditions* (1.3)–(1.5) *then* $f_0(x_0) = \min_{x \in C}\{f_0(x) \mid f(x) \leqslant 0\}$.

Proof. For any $x \in C$ satisfying $f(x) \leqslant 0$ we have

$$\begin{aligned} f_0(x) - f_0(x_0) &\geqslant \eta'(x, x_0) \nabla f_0(x_0) \\ &= -\eta'(x, x_0) \nabla v_0' f(x_0) && \text{by (1.3)}, \\ &\geqslant -v_0'(f(x) - f(x_0)) && \text{by (1.6) and (1.5)}, \\ &= -v_0' f(x) && \text{by (1.4)}, \\ &\geqslant 0 && \text{by (1.2) and (1.5)}. \end{aligned}$$

So x_0 is minimal, which proves the theorem.

Note, more generally, that the proof of Theorem 2.1 requires that definition (1.6) apply only at the points (x, x_0), where x is any point in C and x_0 is the minimal point.

Consider the following example involving a mixture of trigonometric, linear, and quadratic functions:

$$\begin{aligned} \text{minimize} \quad & f_0(x) \equiv x_1 - \sin x_2 \\ \text{subject to} \quad & f_1(x) \equiv \sin x_1 - 4 \sin x_2 \leqslant 0, \\ & f_2(x) \equiv 2 \sin x_1 + 7 \sin x_2 + x_1 - 6 \leqslant 0, \\ & f_3(x) \equiv 2x_1 + 2x_2 - 3 \leqslant 0, \\ & f_4(x) \equiv 4x_1^2 + 4x_2^2 - 9 \leqslant 0, \\ & f_5(x) \equiv -\sin x_1 \leqslant 0, \\ & f_6(x) \equiv -\sin x_2 \leqslant 0, \end{aligned}$$

and C is the constraint set defined by $f(x) \leqslant 0$.

Note the non-convex nature of the constraint set and of the objective function. All of these functions satisfy definition (1.6) with

$$\eta(x, u) = \begin{pmatrix} \dfrac{\sin x_1 - \sin u_1}{\cos u_1} \\[2ex] \dfrac{\sin x_2 - \sin u_2}{\cos u_2} \end{pmatrix},$$

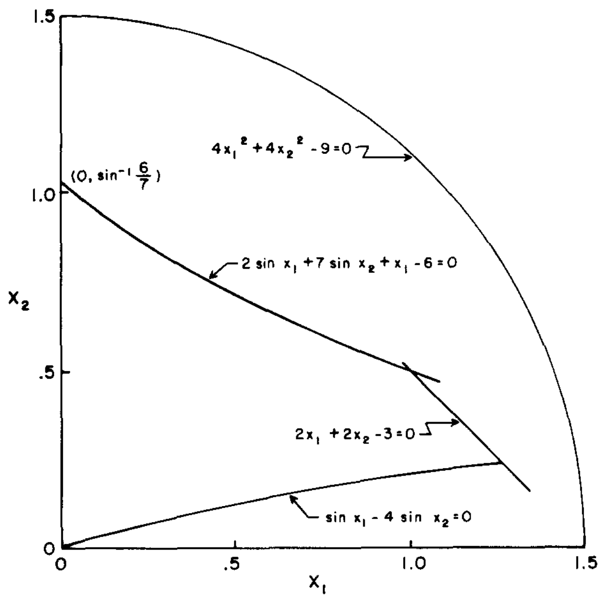

FIGURE 1

where

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad \text{and} \quad u = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}.$$

Here the Lagrangian function $L(x, v) \equiv f_0(x) + v'f(x)$ is

$$\begin{aligned} L = x_1 &- \sin x_2 \\ &+ v_1(\sin x_1 - 4 \sin x_2) \\ &+ v_2(2 \sin x_1 + 7 \sin x_2 + x_1 - 6) \\ &+ v_3(2x_1 + 2x_2 - 3) \\ &+ v_4(4x_1^2 + 4x_2^2 - 9) \\ &+ v_5(-\sin x_1) \\ &+ v_6(-\sin x_2), \end{aligned}$$

and it is easily verified that the Kuhn–Tucker conditions (1.3)–(1.5), are satisfied by

$$v_0 = [0, 1/7, 0, 0, 10/7, 0]',$$

and

$$x_0 = [0, \sin^{-1}(6/7)]'.$$

Note that in order to apply Theorem 2.1 it is not necessary to know the function $\eta(x_1, x_2)$. It is sufficient to know that it exists.

3. Existence of the Function $\eta(x_1, x_2)$

It would be desirable to establish general criteria for determining in any given case whether there exists a function $\eta(x_1, x_2)$ satisfying the system of inequalities

$$\begin{aligned} f_0(x_1) - f_0(x_2) &\geqslant \eta'(x_1, x_2)\, \nabla f_0(x_2), \\ f_1(x_1) - f_1(x_2) &\geqslant \eta'(x_1, x_2)\, \nabla f_1(x_2), \\ &\vdots \\ f_m(x_1) - f_m(x_2) &\geqslant \eta'(x_1, x_2)\, \nabla f_m(x_2), \end{aligned}$$

or, in matrix notation

$$A\eta \leqslant C, \tag{3.1}$$

where A is the Jacobian matrix $[\nabla f_i(x_2)]$ and C is the vector $[f_i(x_1) - f_i(x_2)]$.

From Gale's theorem of the alternative [3] either the system $A\eta \leqslant C$ has a solution η, or the system $A'y = 0$, $C'y = -1$, $y \geqslant 0$, has a solution y, but not both. In a given situation one could in principle determine the existence of η in the former system by determining the nonexistence of y in the latter system.

4. Duality

We consider the following pair of problems defined on C:

Primal Problem.

$$\text{minimize} \quad f_0(x) \tag{4.1}$$

$$\text{subject to} \quad f(x) \leqslant 0 \tag{4.2}$$

Dual Problem.

$$\underset{(u,v)}{\text{maximize}} f_0(u) + v'f(u) \tag{4.3}$$

$$\text{subject to} \quad \nabla f_0(u) + \nabla v'f(u) = 0 \tag{4.4}$$

and

$$v \geqslant 0 \tag{4.5}$$

THEOREM 4.1. *Under the conditions of the Kuhn–Tucker theorem, if x_0 is minimal in the primal problem, then (x_0, v_0) is maximal in the dual problem, where v_0 is given by the Kuhn–Tucker conditions, and $f_i(x)$, $i = 0$, 1,..., m, satisfy* (1.6); *and the extremal values are equal in the two problems.*

Proof. Let (u, v) be any vector satisfying constraints (4.4) and (4.5) of the dual problem.

Then,

$$\begin{aligned}
&[f_0(x_0) + v_0'f(x_0)] - [f_0(u) + v'f(u)] \\
&\quad = f_0(x_0) - f_0(u) - v'f(u) && \text{by (1.4)}, \\
&\quad \geqslant \eta'(x_0, u)\nabla f_0(u) - v'f(u) && \text{by (1.6)}, \\
&\quad = -\eta'(x_0, u)\nabla f(u) - v'f(u) && \text{by (4.4)}, \\
&\quad = -v'f(x_0) && \text{by (1.6)}, \\
&\quad \geqslant 0 && \text{by (4.2) and (4.5)}.
\end{aligned}$$

So (x_0, v_0) is maximal in the dual problem, and since $v_0' f(x_0) = 0$, the extrema of the two problems are equal.

5. Further Generalization

Corresponding to the definitions of quasi-convexity and pseudo-convexity we can define two more general classes of functions, respectively:

$$\phi(x_1) \leqslant \phi(x_2) \Rightarrow \eta'(x_1, x_2) \nabla\phi(x_2) \leqslant 0, \tag{5.1}$$

$$\eta'(x_1, x_2) \nabla\phi(x_2) \geqslant 0 \Rightarrow \phi(x_1) \geqslant \phi(x_2) \tag{5.2}$$

for all $x_1, x_2 \in C$.

It can readily be shown that in problem (1.1), (1.2), if $f_0(x)$ satisfies (5.2) and $f_i(x)$, $i = 1, ..., m$, satisfy (5.1), the Kuhn–Tucker conditions are sufficient.

Comment. It is apparent that the classes of functions introduced by definitions (1.6), (5.1), and (5.2) will replace convex and generalized convex functions in most of the theory and applications of mathematical programming and cognate topics in their more general ramifications.

References

1. O. L. Mangasarian, Pseudo-convex functions, *SIAM J. Control* 3 (1965), 281–290.
2. E. T. Whittaker and G. N. Watson, "A Course of Modern Analysis," p. 129, Cambridge Univ. Press, Cambridge, England, 1963.
3. O. L. Mangasarian, "Nonlinear Programming," p. 34, McGraw–Hill, New York, 1969.