Mathematical Programming 22 (1982) 227–230
North-Holland Publishing Company

SHORT COMMUNICATION

ON MINIMAX OPTIMIZATION PROBLEMS

Zvi DREZNER

School of Management, The University of Michigan-Dearborn, Dearborn, MI, U.S.A.

Received 23 June 1980
Revised manuscript received 23 April 1981

We give a short proof that in a convex minimax optimization problem in k dimensions there exist a subset of $k+1$ functions such that a solution to the minimax problem with those $k+1$ functions is a solution to the minimax problem with all functions. We show that convexity is necessary, and prove a similar theorem for stationary points when the functions are not necessarily convex but the gradient exists for each function.

Key words: Minimax, Nonconvex Optimization, Stationary Points.

1. Introduction

We generalize a minimax theorem for convex functions to non-convex differentiable functions. We prove two theorems, and present two examples. One example shows that convexity is necessary for the first theorem, and the second shows that the second theorem must be formulated with non-negative stationary points rather than with local minimum points.

2. The convex case

In this section we assume that the functions involved are convex. Let $f_i(x)$ for $i = 1, \dots, n$ be convex functions. Consider the following minimax optimization problem.

$$\underset{x}{\text{minimize}} \left\{ \max_{1 \le i \le n} \{f_i(x)\} \right\} \tag{1}$$

where

$$x = (x_1, \dots, x_k).$$

Let f^* be the minimal value of the objective function in problem (1). By the convexity of $f_i(x)$ the set $\{x \mid f_i(x) \le f_0\}$ is either empty or a convex set. Let $N = \{1, \dots, n\}$ and let $I \subset N$. Let $F_I(f_0)$ be the intersection of all sets $\{x \mid f_i(x) \le f_0\}$ for $i \in I$. Note that f^* is the minimal value of f_0 such that $F_N(f_0) \ne \emptyset$.

The following theorem is proved in [3] via a transformation to a regular mathematical programming problem. We present here a short and direct proof.

Theorem 1. *There exists a subgroup $I \subset N$ of cardinality less than or equal to $k+1$ such that the problem*

$$\underset{x}{\text{minimize}} \left\{ \max_{i \in I} \{ f_i(x) \} \right\} \tag{2}$$

has an optimal value of f^*. Furthermore, at least one of the solution points to problem (2) *is also a solution point to problem* (1).

Proof. The case $n \le k+1$ is trivial, so let us assume $n > k+1$. Consider all possible sets of $k+1$ members out of N. Let them be $I_1, \dots, I_r$ where $r = \binom{n}{k+1}$. Calculate

$$F_{I_j} = \min_x \left\{ \max_{i \in I_j} \{ f_i(x) \} \right\}. \tag{3}$$

Since $F_N(f^*) \ne \emptyset$, $F_{Ij}(f^*) \ne \emptyset$. Therefore, $f_{I_j} \le f^*$. Let $f^m = \max_j \{ f_{I_j} \}$. Since $f_{I_j} \le f^*$, then $f^m \le f^*$. Since $f_{I_j} \le f^m$, for every I_j, $F_{I_j}(f^m) \ne \emptyset$. By Helly's Theorem [2, 4], $F_N(f^m) \ne 0$ and therefore $f^* \le f^m$. Therefore, $f^m = f^*$. If $f^m = f^*$, then there exists I_j such that $f_{I_j} = f^*$. Therefore, the solution to problem (2) with $I = I_j$ has the optimal value f^*. Now, since $F_N(f^*) \subset F_{I_j}(f^*)$, at least one of the solution points to problem (2) is also a solution point to problem (1).

Note that if $f_i(x)$ are strictly convex, then the solution point to problem (1) is unique. Therefore, the unique solution for group I in the theorem must be the unique solution to problem (1).

3. The nonconvex case

One may suggest that Theorem 1 can be true for nonconvex functions and local minima. In this section we show that such modification of the theorem is false. However, we present another generalization to Theorem 1 for nonconvex differentiable functions. In this section we assume that $f_i(x)$ are not necessarily convex.

To show that convexity is essential for the theorem, consider the following example:

$$\begin{aligned} &f_1(x, y) = 2 - |x-1| - |y|, && f_2(x, y) = 2 - |x+1| - |y|, \\ &f_3(x, y) = 2 - |x| - |y-1|, && f_4(x, y) = 2 - |x| - |y+1|, \\ &f_5(x, y) = |x| + |y|. \end{aligned} \tag{3}$$

A simple check shows that $(0, 0)$ is a global minimum. However, removal of any of the first four functions creates a descent direction. For example, by removing $f_1(x, y)$ from (3) the objective function decreases as x increases. Therefore, there is no set of three functions with a local minimum at $(0, 0)$.

Geometrically, every $f_i(x, y)$ for $i = 1, \dots, 4$ 'covers' a section of angle $\pi/2$ inside which $f_i(x, y)$ is increasing. We have arranged the four functions to cover all possible directions, Removing one function creates a direction in which the objective function is decreasing. One can construct functions that increase inside a section of angle $2\pi/K$ for $K \geq 4$ getting a counterexample for any $K \geq 4$.

When $f_i(x)$ are convex the angle of such sections must be at least π. Note that when the gradient of $f_i(x)$ exists, that angle must be exactly π. This observation leads to the following theorem.

Let us first present a new concept similar to definitions in [1]. Let us have a function $f(x)$ with directional derivatives at every point. A stationary point is a point such that the directional derivative is zero in every direction. Let us define a *non-negative stationary point* as a point such that the directional derivative is non-negative in every direction. Similarly, a non-positive stationary point possesses non-positive directional derivatives.

Two trivial properties of the new concepts:

(i) If a point is a non-negative stationary point and a non-positive stationary point, then it must be a stationary point.

(ii) If the gradient of $f(x)$ exists at a certain point, then non-negative stationary point, non-positive stationary point, and stationary point are equivalent at that point.

Let $f_I(x) = \max_{i \in I}\{f_i(x)\}$. Let x^S be a non-negative stationary point of $f_N(x)$, and let $S = \{i \mid f_i(x^S) = f_N(x^S)\}$. Note that $f_i(x)$ are not necessarily convex.

Theorem 2. *If the gradient of $f_i(x)$ at x^S exists for $i \in S$, then there exists a subset $I \subset N$ of cardinality less than or equal to $k + 1$ such that x^S is a non-negative stationary point of $f_I(x)$.*

Proof. The case $n \leq k + 1$ is trivial, so let us assume $n > k + 1$. Let θ be a direction vector. Since x^S is a non-negative stationary point of $f_N(x)$, there exist $i \in S$ such that $\nabla f_i(x^S) \cdot \theta \geq 0$ in this direction. Let $g_i(x) = \nabla f_i(x^S) \cdot (x - x^S)$ for $i \in S$. $g_i(x) = 0$ is the tangent hyperplane to $f_i(x)$ at $x = x^S$. Since for any direction θ there exists an $i \in S$ for which $g_i(x + \lambda\theta) \geq 0$ for $\lambda \geq 0$, x^S is the optimal point for the problem:

$$\min_x \left\{ \max_{i \in S} \{g_i(x)\} \right\}.$$

Since $g_i(x)$ are convex for $i \in S$, there exist by Theorem 1 a subset I, $I \subset S \subset N$, of cardinality less than or equal to $k + 1$ such that x^S is optimal for the problem:

$$\min_x \left\{ \max_{i \in I} \{g_i(x)\} \right\}.$$

Therefore, for any direction θ there exists an $i \in I$ for which $g_i(x+\lambda\theta) \geq 0$, which means $\nabla f_i(x^S) \cdot \theta \geq 0$. Therefore, x^S is a non-negative stationary point of $f_I(x)$.

In the following example x^S is the global minimum of $f_N(x)$, but for each possible I, $f_I(x)$ has no local minimum point at $x = x^S$.

$$\begin{aligned} &f_1(x,y) = 2 - [(x-1)^2 + y^2]^{1/2}, \qquad f_2(x,y) = 2 - [(x+1)^2 + y^2]^{1/2} \\ &f_3(x,y) = 2 - [x^2 + (y-1)^2]^{1/2}, \qquad f_4(x,y) = 2 - [x^2 + (y+1)^2]^{1/2} \\ &f_5(x,y) = (x^2 + y^2)^{1/2}. \end{aligned} \tag{4}$$

A simple check shows that (0, 0) is the global minimum. The gradient of $f_i(x, y)$ for $i = 1, \dots, 4$ exists at $(0, 0)$. However, removal of any of the first four functions creates one descent direction while the directional derivative in that direction is equal to zero. For example, by removing $f_1(x, y)$ the objective function decreases as x increases. The derivative along that direction is zero.

Acknowledgment

I would like to thank the associate editor for his useful remarks, and for turning my attention to references [1] and [3].

References

[1] F.H. Clarke, "Generalized gradients and applications", *Transactions of the American Mathematical Society* 205 (1975) 247–262.

[2] H.G. Eggleston, *Convexity* (Cambridge University Press, Cambridge, 1969).

[3] V.L. Levin, "Application of E. Helly's theorem to convex programming problems of best approximation and related questions", (English translation) *Mathematics of the USSR Sbornik* 8 (1969) 235–247.

[4] J. Radon, "Mengen konvexer Körper, die einen gemeinsamen Punkt enthalten", *Mathematische Annalen* 83 (1921) 113–115.