Journal of Computational Mathematics, Vol.9, No.3, 1991, 222-228.

A GLOBALLY CONVERGENT ALGORITHM FOR A LOCALLY LIPSCHITZ FUNCTION*1)

Zhang Lian-sheng Tian Wei-wen
(*Shanghai University of Science and Technology, Shanghai, China*)

Abstract

In this paper, an algorithm of global convergence is proposed for a locally Lipschitz function, which is strictly differentiable at almost all differentiable points, and several examples are computed on an IBM PC.

§1. Introduction

A great deal of effort has been devoted to nondifferentiable optimization in recent years. Many researches have gone into implementable algorithms mainly by Wolfe, Lemarechal, Zowe, Polak, Kiwiel, et al., besides various definitions of subgradient and corresponding first-order optimality conditions. However, in these algorithms the functions are required to be convex or semi-smooth in order to guarantee global convergence of the algorithms. It seems that there is no globally convergent algorithm for the locally Lipschitz function without additional condition. In this paper, we propose an algorithm of global convergence for a locally Lipschitz function, which is strictly differentiable at almost all differentiable points. In addition, we describe some concepts and properties of the locally Lipschitz function.

Definition 1.1. Let $f : R^n \to R$ be locally Lipschitz continuous. The generalized gradient of f at x is defined by

$$\partial f(x) = \text{co}\{\lim_{v_i \to 0} \nabla f(x + v_i)\}$$

where $\nabla f(x)$ denotes the gradient of f at x, co denotes the convex hull of a set, and v_i are such that $\nabla f(x+v_i)$ exists and $\lim\limits_{v_i \to 0} \nabla f(x+v_i)$ exists. We recall that a locally Lipschitz function $f(x), x \in R^n$, is differentiable almost everywhere.

Definition 1.2. Let $f : R^n \to R$ be locally Lipschitz continuous. The generalized directional derivative of f at x in the direction h is defined by

$$f^0(x; h) = \overline{\lim_{\substack{y \to 0 \\ t \to 0}}} \frac{f(x + y + th) - f(x + y)}{t}.$$

Proposition. Let $f : R^n \to R$ be locally Lipschitz continuous. Then
1. $\partial f(x)$ exists and is compact at all $x \in R^n$.
2. $\partial f(x)$ is bounded on bounded sets.

* Received April 12, 1988.
1) The Project Supported by National Natural Science Foundation of China.

3. $\partial f(x)$ is upper semi-continuous (u.s.c.) in the sense that

$$\{x_i \to \hat{x}, y_i \in \partial f(x_i) \text{ and } y_i \to \hat{y}\} \Rightarrow \{\hat{y} \in \partial f(\hat{x})\}.$$

4. $f^0(x;h)$ exists for all $x, h \in R^n$, and

$$f^0(x;h) = \max_{\xi \in \partial f(x)} \langle \xi, h \rangle.$$

Definition 1.3. For any $\varepsilon > 0$, the ε-smeared generalized gradient is defined by

$$\partial_\varepsilon f(x) = \text{co}\Big\{ \bigcup_{x' \in x + \varepsilon B(\theta,1)} \partial f(x') \}, \quad B(\theta, 1) = \{y : \|y\| \le 1\Big\}.$$

$\partial_\varepsilon f(x)$ has properties 1–3 above.

§2. Several Lemmas

In this section, we demonstrate several lemmas concerning the algorithm.
Let $f : R^n \to R$ be locally Lipschitzian. For any $\varepsilon > 0$, we define

$$h_\varepsilon(x) = -\text{Nr}(\partial_\varepsilon f(x)) = -\text{Argmin}\{\|h\| : h \in \partial_\varepsilon f(x)\} \tag{2.1}$$

and $\varepsilon : R^n \to R$ by

$$\varepsilon(x) = \max\{\varepsilon \in \mathcal{E} : \|h_\varepsilon(x)\|^2 \ge \delta\varepsilon\} \tag{2.2}$$

where

$$\mathcal{E} = \{\varepsilon : \varepsilon = \varepsilon_0 v^k, \ k \in N^+\} U\{0\} \tag{2.3}$$

and $v \in (0,1), \varepsilon_0 > 0, \delta > 0$ are given.

Lemma 1. *The function $\|h_\varepsilon(x)\|^2$ defined by (2.1) is lower semi-continuous with respect to x.*

Proof. We know that the point-set mapping $\partial_\varepsilon f(x)$ is u.s.c. and the set $\bigcup_{x \in A} \partial_\varepsilon f(x)$ is bounded on the bounded set A. Since $h_\varepsilon(x) = -$ Argmin $\{\|h\| : h \in \partial_\varepsilon f(x)\}$ implies that $-h_\varepsilon(x) \in \partial_\varepsilon f(x)$, it follows that the set $\{-h_\varepsilon(x)\}$ is bounded. Hence, we can suppose that $-h_\varepsilon(x) \xrightarrow[x \to x_0]{} h_0$ and $h_0 \in \partial_\varepsilon f(x_0)$ holds by the upper semi-continuity of $\partial_\varepsilon f(x)$. Next, since $-h_\varepsilon(x_0) \in \partial_\varepsilon f(x_0)$ and $-h_\varepsilon(x_0) =$ Argmin $\{\|h\| : h \in \partial_\varepsilon f(x_0)\}$, it follows that $\|h_0\| \ge \|-h_\varepsilon(x_0)\| = \|h_\varepsilon(x_0)\|$. Moreover, since $\lim_{x \to x_0} -h_\varepsilon(x) = h_0$ implies $\lim_{x \to x_0} \|h_\varepsilon(x)\| = \|h_0\| \ge \|h_\varepsilon(x_0\|$, it follows that $\lim_{x \to x_0} \|h_\varepsilon(x)\|^2 \ge \|h_\varepsilon(x_0)\|^2$. Consequently, $\lim_{x \to x_0} \inf \|h_\varepsilon(x)\|^2 \ge \|h_\varepsilon(x_0)\|^2$, that is, $\|h_\varepsilon(x)\|^2$ is l.s.c..

Lemma 2. *For every $\bar{x} \in R^n$ such that $\theta \bar{\in} \partial f(\bar{x})$, there exists a $\rho(\bar{x}) > 0$ such that*

$$\varepsilon(x_i) \ge v\varepsilon(\bar{x}) > 0, \qquad \textit{for all } x_i \in B(\bar{x}, \rho(\bar{x}))$$

where $B(\bar{x}, \rho(\bar{x})) = \{x : \|x - \bar{x}\| \le \rho(\bar{x})\}$.