Journal of Computational Mathematics, Vol.21, No.6, 2003, 791–800.

OPTIMALITY CONDITIONS OF A CLASS OF SPECIAL NONSMOOTH PROGRAMMING *1)

Song-bai Sheng
(*Colleges of Science, Nanjing University of Aeronautics and Astronautics, Nanjing 210016, China*)

Hui-fu Xu
(*Australian Graduate School of Management, the University of New South Wales, Sydney, Australia*)

Abstract

In this paper, we investigate the optimality conditions of a class of special nonsmooth programming min $F(x) = \sum_{i=1}^{m} |max\{f_i(x), c_i\}|$ which arises from L_1-norm optimization, where $c_i \in R$ is constant and $f_i \in C^1, i = 1, 2, \cdots, m$. These conditions can easily be tested by computer.

Key words: Generalized gradient, Directional derivative, Optimality conditions, Nonsmooth programming.

1. Introduction

Consider a class of special nonsmooth programming

$$\min_{x \in R^n} F(x) = \sum_{i=1}^{m} |\max\{f_i(x), c_i\}| \tag{1.1}$$

where constant $c_i \in R, f_i \in C^1, i = 1, 2, \cdots, m$, and in general there is at least one $c_i < 0$. The problem (1.1) arises from the L_1 norm optimization. For example, the discrete L_1 linear approximation[2], the L_1 solution of an overdetermined linear systems[3], the censored discrete linear L_1 approximation[7,8]

$$\min_{x \in R^n} F(x) = \sum_{i=1}^{m} |y_i - \max\{a_i^T x, z_i\}| \tag{1.2}$$

and from the L_1 penalty function model of constrained programming[5,6]

$$\min_{x \in R^n} F(x) = f(x) + \lambda \sum_{i=1}^{m} \max\{g_i(x), 0\} \tag{1.3}$$

where $\lambda > 0$ is a penalty coefficient.

The aim of this paper is to investigate the optimality conditions of the problem(1.1). It is well known that for the general nonsmooth function $F(x)$, *i.e.*, $F(x)$ is locally Lipschitz continuous at any x, the necessary condition of a local minimizer x^* of $F(x)$ is $0 \in \partial F(x^*)$. This condition is not easily tested by computer. For the special problem (1.1), we can obtain the necessary conditions and sufficient conditions of a local minimizer x^* of $F(x)$, which can easily be tested by computer.

---

* Received July 31, 2001.
1) The project was supported by NNSFC(No. 19771047) and NSFJS (BK97059).

In the next section, we consider the differential properties of $F(x)$ and establish a characterization of the generalized gradient $\partial F(x)$. In section 3, we discuss the descent direction of $F(x)$ based on the gradient of $f_i(x), i = 1, 2, \cdots, m$. Then we provide necessary conditions and sufficient conditions for a (strict) local minimizer of $F(x)$. In the last section, we provide the optimality conditions of problem(1.2) and (1.3).

2. Differential Properties

The nonlinear and nonconvex function $F(x)$ defined by (1.1) can be written as the sum of smooth functions and nonsmooth functions. To do this, for any given $x \in R^n$, define the index sets $\Gamma_j(x), j = 1, \cdots, 5$, by

$$\Gamma_1(x) = \{i \in [1:m] | f_i(x) > c_i \geq 0 \quad \text{or} \quad f_i(x) > c_i, c_i < 0 \quad \textit{and} \quad f_i(x) \neq 0\},$$

$$\Gamma_2(x) = \{i \in [1:m] | f_i(x) < c_i\}, \qquad \Gamma_3(x) = \{i \in [1:m] | f_i(x) = c_i \geq 0\},$$

$$\Gamma_4(x) = \{i \in [1:m] | f_i(x) = c_i < 0\}, \qquad \Gamma_5(x) = \{i \in [1:m] | f_i(x) = 0 \quad \textit{and} \quad c_i < 0\}.$$

The sets $\{\Gamma_j(x), j = 1, \cdots, 5\}$ form a disjoint partition of $\{1, 2, \cdots, m\}$, that is

$$\bigcup_{j=1}^{5} \Gamma_i(x) = \{1, 2, \cdots, m\}, \quad \forall x \in R^n,$$

$$\Gamma_i(x) \bigcap \Gamma_j(x) = \phi, \quad \forall i \neq j \quad \forall x \in R^n.$$

For the simplicity, let $\Gamma_{ijk} = \Gamma_i(x) \cup \Gamma_j(x) \cup \Gamma_k(x)$, we have

$$F(x) = \sum_{i \in \Gamma_{12}} |\max\{f_i(x), c_i\}| + \sum_{i \in \Gamma_{345}} |\max\{f_i(x), c_i\}| \tag{2.1}$$

For $i \in \Gamma_1$ the component function $|\max\{f_i(x), c_i\}| = \text{sign}(f_i(x)) f_i(x)$, which is smooth in a neighborhood of x with gradient $\text{sign}(f_i(x)) \nabla f_i(x)$. For $i \in \Gamma_2$ the component function $|\max\{f_i(x), c_i\}| = |c_i|$, which is constant and hence smooth in a neighborhood of x. Thus the gradient of the smooth part of $F(x)$ is

$$g(x) = \nabla \left(\sum_{i \in \Gamma_{12}} |\max\{f_i(x), c_i\}| \right) = \sum_{i \in \Gamma_1} \text{sign}(f_i(x)) \nabla f_i(x) \tag{2.2}$$

A definition of the generalized gradient $\partial f(x)$ [4] of a piecewise smooth function at a point x is

$$\begin{aligned} \partial f(x) = \text{co}\{v \in R^n | \exists \ \text{ a sequence } \{x_k\} \text{ such that} \\ x_k \to x, \nabla f(x_k) \text{ exists } \ \forall\, k \text{ and } \nabla f(x_k) \to v \text{ as } k \to +\infty\} \end{aligned} \tag{2.3}$$

where co denotes the convex hull. Furthermore, $\partial f(x)$ is a nonempty compact convex set in R^n.

Now, for $i \in \Gamma_{345}$ the corresponding component functions are piecewise smooth. According to (2.3) the generalized gradients are given by

$$\partial |\max\{f_i(x), c_i\}| = \begin{cases} \text{co}\{0, \nabla f_i(x)\} = \{v \in R^n | v = \lambda_i \nabla f_i(x), 0 \leq \lambda_i \leq 1\}, i \in \Gamma_3; \\ \text{co}\{0, -\nabla f_i(x)\} = \{v \in R^n | v = \lambda_i \nabla f_i(x), -1 \leq \lambda_i \leq 0\}, i \in \Gamma_4; \\ \text{co}\{\nabla f_i(x), -\nabla f_i(x)\} = \{v \in R^n | v = \lambda_i \nabla f_i(x), -1 \leq \lambda_i \leq 1\}, i \in \Gamma_5. \end{cases}$$