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

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

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

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

Let

$$G(x) = \{v \in R^n | v = g(x) + \sum_{i \in \Gamma_{345}} \lambda_i \nabla f_i(x)\} \tag{2.4}$$

where $\lambda_i \in [0,1]$ for $i \in \Gamma_3$, $\lambda_i \in [-1,0]$ for $i \in \Gamma_4$, and $\lambda_i \in [-1,1]$ for $i \in \Gamma_5$. Then $G(x)$ is a nonempty compact and convex polytope. On the other hand, as generalized gradient satisfies $\partial(f_1(x) + f_2(x)) \subseteq \partial f_1(x) + \partial f_2(x)$, one has $\partial F(x) \subseteq G(x)$. Because $F(x)$ may not be convex, the inclusion $\partial F(x) \subseteq G(x)$ can be strict in general. The following result ensures the coincidence of the sets $\partial F(x)$ and $G(x)$.

Let $\widetilde{\lambda}_i = 0$ or 1 for $i \in \Gamma_3$, $\widetilde{\lambda}_i = -1$ or 0 for $i \in \Gamma_4$, and $\widetilde{\lambda}_i = -1$ or 1 for $i \in \Gamma_5$. Let $\widetilde{\lambda} = (\widetilde{\lambda}_i, i \in \Gamma_{345})$ be a vector. Then it is obvious that every vertex of set $G(x)$ corresponds to a given vector $\widetilde{\lambda}$. Define the set $X(\widetilde{\lambda})$ by

$$X(\widetilde{\lambda}) = \left\{ y \in R^n \;\middle|\; \begin{array}{l} f_i(y) \geq c_i \text{ for } \; i \in \Gamma_3 \text{ with } \widetilde{\lambda}_i = 1, \\ f_i(y) \leq c_i \text{ for } \; i \in \Gamma_3 \text{ with } \widetilde{\lambda}_i = 0; \\ f_i(y) \geq c_i \text{ for } \; i \in \Gamma_4 \text{ with } \widetilde{\lambda}_i = -1, \\ f_i(y) \leq c_i \text{ for } \; i \in \Gamma_4 \text{ with } \widetilde{\lambda}_i = 0; \\ f_i(y) \geq 0 \text{ for } \; i \in \Gamma_5 \text{ with } \widetilde{\lambda}_i = 1, \\ f_i(y) \leq 0 \text{ for } \; i \in \Gamma_5 \text{ with } \widetilde{\lambda}_i = -1 \end{array} \right\}$$

and set $\Lambda(x) = \{\widetilde{\lambda}, \widetilde{\lambda}$ corresponding to vertex of $G(x)\}$.

Theorem 2.1. *$\partial F(x) = G(x)$ at $x \in R^n$ if and only if int $X(\widetilde{\lambda}) \neq \phi$ for every $\widetilde{\lambda} \in \Lambda(x)$.*

Proof. Assume int $X(\widetilde{\lambda})$ is nonempty for every $\widetilde{\lambda} \in \Lambda(x)$. Since $\partial F(x) \subseteq G(x)$, it is suffices to show $G(x) \subseteq \partial F(x)$. As both $G(x)$ and $\partial F(x)$ are convex and nonempty, we only need to prove that any vertex of $G(x)$ belongs to $\partial F(x)$. Now, suppose that

$$u = g(x) + \sum_{i \in \Gamma_{345}} \widetilde{\lambda}_i \nabla f_i(x) \tag{2.5}$$

is a vertex of $G(x)$, which corresponds a vector $\widetilde{\lambda} \in \Lambda(x)$. Note that the given $x \in X(\widetilde{\lambda})$ for all $\widetilde{\lambda} \in \Lambda(x)$. Then there exists a sequence $\{x_k\} \subseteq \text{int}X(\widetilde{\lambda})$ such that $\{x_k\}$ converges to x. As x_k is an interior point, the function $|\max\{f_i(x), c_i\}|, i \in \Gamma_{345}$, is differentiable at x_k.Therefore $\nabla F(x_k)$ exists for all k and $\nabla F(x_k) \to u$ as $k \to +\infty$. By the definition of generalized gradient, one has $u \in \partial F(x)$.

Conversely, if $\partial F(x) = G(x)$ then every $\widetilde{\lambda} \in \Lambda(x)$ corresponds to a vertex u of $G(x)$ (see (2.5)), and hence $\partial F(x)$, by the definition of generalized gradient there exists a sequence $\{x_k\}$ such that $x_k \to x, \nabla F(x_k)$ exists for all k and $\nabla F(x_k) \to u$ as $k \to \infty$. Note that every differentiable point x_k is an interior point, and $\partial F(x)$ is convex nonempty, so for k sufficiently large $\{x_k\}$ must belong to $X(\widetilde{\lambda})$ by the definition of $X(\widetilde{\lambda})$, that is int $X(\widetilde{\lambda}) \neq \phi$.

It is not easy to verify int $X(\widetilde{\lambda}) \neq \phi$ for all $\widetilde{\lambda} \in \Lambda(x)$. However, in usual applied problems we have $c_i \geq 0, i = 1, 2, \cdots, m, i.e., \Gamma_4 \cup \Gamma_5 = \phi$. In the following, we shall give sufficient conditions to ensure $\partial F(x) = G(x)$.

The function $f(x)$ is said to be locally convex at x if there exists a neighborhood U of x such that $f(x)$ is convex on U.

Theorem 2.2. *If $c_i \geq 0, i = 1, 2, \cdots, m$ and $f_i(x)$ is locally convex at x for $i \in \Gamma_1 \cup \Gamma_3$. Then $\partial F(x) = G(x)$.*

Proof. Because $|\max\{f_i(x), c_i\}| = \max\{f_i(x), c_i\}$ is locally convex at x for $i \in \Gamma_{123}$ and $\Gamma_{45} = \phi$, $F(x)$ is locally convex at x. According as the Corollary 3 of Proposition 2.3.3 and Proposition 2.3.6(b) in [4], we have $\partial F(x) = \sum_{i=1}^{m} \partial|\max\{f_i(x), c_i\}| = G(x)$.

3. Optimality Conditions

A key tool in the discussion of optimality conditions in nonsmooth optimization is the directional derivative $F'(x,s)$ defined by

$$F'(x,s) = \lim_{t\to 0^+}[F(x+ts) - F(x)]/t, \quad x, s \in R^n, s \neq 0$$

In this paper, since F is a piecewise continuously differentiable, $F'(x,s)$ exists. Let $F_i(x) = |\max\{f_i(x), c_i\}|$.

Theorem 3.1. *For any* $x, s \in R^n, s \neq 0$

$$\begin{aligned} F'(x,s) &= g(x)^T s + \sum_{i\in\Gamma_3} \max\{0, \nabla f_i(x)^T s\} \\ &- \sum_{i\in\Gamma_4} \max\{0, \nabla f_i(x)^T s\} + \sum_{i\in\Gamma_5} |\nabla f_i(x)^T s| \end{aligned} \tag{3.1}$$

where $g(x) = \sum_{i\in\Gamma_1} \text{sign}(f_i(x))\nabla f_i(x)$.

Proof. Since

$$\nabla F_i(x) = \begin{cases} \text{sign}(f_i(x))\nabla f_i(x), & i \in \Gamma_1 \\ 0, & i \in \Gamma_2 \end{cases}$$

we have $\sum_{i\in\Gamma_{12}} F'_i(x,s) = g(x)^T s$. For $i \in \Gamma_3$, as $f_i(x) = c_i > 0, F_i(x) = \max\{f_i(x), c_i\}$. Due to $f_i \in C^1$, it is well known that

$$\begin{aligned} \lim_{t\to 0^+}[F_i(x+ts) - F_i(x)]/t &= \lim_{t\to 0^+}[\max\{f_i(x+ts), c_i\} - c_i]/t \\ &= \begin{cases} \nabla f_i(x)^T s, & \text{if } \nabla f_i(x)^T s \geq 0 \\ 0, & \text{if } \nabla f_i(x)^T s < 0 \end{cases} \end{aligned}$$

Then

$$F_i'(x,s) = \max\{0, \nabla f_i(x)^T s\}, \quad i \in \Gamma_3$$

Similarly

$$F_i'(x,s) = -\max\{0, \nabla f_i(x)^T s\}, \quad i \in \Gamma_4$$
$$F_i'(x,s) = |\nabla f_i(x)^T s|, \quad i \in \Gamma_5.$$

The proof is completed.

The inclusion $\partial F(x) \subseteq G(x)$ implies that the well known necessary condition $0 \in \partial F(x^*)$, for x^* to be a local minimizer of F, carry through to $0 \in G(x^*)$. However, it does not ensure that x^* is a local minimizer of F if $0 \in G(x^*)$, but we can obtain a sequence $\{x_k\}$ approximated to a minimizer x^*. For convenience, let

$$a_i = \nabla f_i(x), \quad i = 1, 2, \cdots, m$$

$$A = A(x) = [a_i, i \in \Gamma_{53}]$$

Theorem 3.2. *Let* x *be a point such that* $0 \in G(x)$ *and rank* $(A) < n$. *Then* $F'(x,s) \leq 0$ *for all* s *satisfying*

$$A^T s = 0 \tag{3.2}$$

Furthermore, if there is $j \in \Gamma_4$ *such that* $a_j \overline{\in} R(A)$, *the range of* A, *then there exists* $\overline{s}$ *such that* $F'(x,\overline{s}) < 0$ *and* $A^T\overline{s} = 0$.

Proof. Due to (2.4)and $0 \in G(x)$, there exist multipliers $\{\lambda_i, i \in \Gamma_{345}\}$ such that

$$g(x) + \sum_{i\in\Gamma_{345}} \lambda_i a_i = 0$$

where $\lambda_i \in [0,1]$ for $i \in \Gamma_3, \lambda_i \in [-1,0]$ for $i \in \Gamma_4$ and $\lambda_i \in [-1,1]$ for $i \in \Gamma_5$. Thus it follows from (3.1) and (3.2)

$$\begin{aligned} F'(x,s) &= g(x)^T s - \sum_{i\in\Gamma_4} \max\{0, a_i^T s\} \\ &= -\sum_{i\in\Gamma_4} [\max\{0, a_i^T s\} + \lambda_i a_i^T s] \le 0 \end{aligned}$$

since $\lambda_i[-1,0]$ for $i \in \Gamma_4$ and

$$\max\{0, a_i^T s\} + \lambda_i a_i^T s = \begin{cases} (1+\lambda_i) a_i^T s \ge 0, & \text{if } \ a_i^T s \ge 0 \\ \lambda_i a_i^T s \ge 0, & \text{if } \ a_i^T s < 0 \end{cases}$$

Furthermore, if there is $j \in \Gamma_4$ such that $a_j \overline{\in} R(A)$, then as rank $(A) < n$ there exists $\overline{s} \in R^n$ such that

$$A^T \overline{s} = 0 \ \text{ but } \ a_j^T \overline{s} \ne 0$$

Therefore

$$\begin{aligned} F'(x,\overline{s}) &= -\sum_{i\in\Gamma_4} [\max\{0, a_i^T \overline{s}\} + \lambda_i a_i^T \overline{s}] \\ &\le -[\max\{0, a_j^T \overline{s}\} + \lambda_j a_j^T \overline{s}] \\ &= \begin{cases} -(1+\lambda_j) a_j^T \overline{s}, & \text{if } \ a_j^T \overline{s} \ge 0 \\ -\lambda_j a_j^T \overline{s}, & \text{if } \ a_j^T \overline{s} < 0 \end{cases} \end{aligned}$$

Now, if $\lambda_j \in [-1,0)$ and $\overline{s}$ is chosen such that $a_j^T \overline{s} < 0$ then $F'(x,\overline{s}) < 0$. Alternatively, if $\lambda_j = 0$ and $\overline{s}$ is chosen such that $a_j^T \overline{s} > 0$ then $F'(x,\overline{s}) < 0$. Completing the proof.

For the simplicity of notation, in the following we shall denote $A^* = A(x^*), g^* = g(x^*), \Gamma_{53}^* = \Gamma_5(x^*) \cup \Gamma_3(x^*)$ and so on, but $a_i = \nabla f_i(x^*), \forall i \in \Gamma_{345}^*$.

Suppose x^* is a point such that $g^* \in R(A^*)$ and $a_j \in R(A^*)$ for all $j \in \Gamma_4^*$. Then there exist multipliers $\{\lambda_i\}$ and $\{u_i^j\}$ satisfying

$$g^* = \sum_{i\in\Gamma_{53}^*} \lambda_i a_i, \quad a_j = \sum_{i\in\Gamma_{53}^*} u_i^j a_i, \forall j \in \Gamma_4^*$$

The multipliers $\{\lambda_i\}$,$\{u_i^j\}$ are not unique in general. Without loss of generality, let

$$\Gamma_5^* = \{1, 2, \cdots, L\}, \quad \Gamma_3^* = \{L+1, \cdots, L+T\},$$

$$\{a_1, \cdots, a_l\} \text{ is a basis of } R(a_1, \cdots, a_L),$$

$$\{a_1, \cdots, a_l, a_{L+1}, \cdots, a_{L+t}\} \text{ is a basis of } R(A^*) \tag{3.3}$$

Clearly $l + t \le n$ and $L + T \le m$. Then there exist the unique multipliers $\{\lambda_i\}$, $\{\alpha_i\}$, $\{\beta_i^j\}$, $\{\delta_i^j\}$, $\{\rho_i^j\}$, $\{u_i^j\}$ and $\{v_i^j\}$ satisfying

$$g^* = \sum_{i=1}^{l} \lambda_i a_i + \sum_{i=L+1}^{L+t} \alpha_i a_i$$

$$a_j = \sum_{i=1}^{l} \beta_i^j a_i, \quad j = l+1, \cdots, L, \tag{3.4}$$

$$a_j = \sum_{i=1}^{l} \delta_i^j a_i + \sum_{i=L+1}^{L+t} \rho_i^j a_i, \quad j = L+t+1, \cdots, L+T, \tag{3.5}$$

$$a_j = \sum_{i=1}^{l} u_i^j a_i + \sum_{i=L+1}^{L+t} v_i^j a_i, \quad \forall j \in \Gamma_4^*.$$

Theorem 3.3. *If x^* is a local minimizer of F and $R(A^*)$ has a basis (3.3), then there exist multipliers $\{\lambda_i\}$, $\{\alpha_i\}$, $\{\beta_i^j\}$, $\{\delta_i^j\}$, $\{\rho_i^j\}$, $\{u_i^j\}$ and $\{v_i^j\}$ satisfying*

$$\text{(i)} \quad g^* = \sum_{i=1}^{l} \lambda_i a_i + \sum_{i=L+1}^{L+t} \alpha_i a_i, \qquad a_j = \sum_{i=1}^{l} u_i^j a_i + \sum_{i=L+1}^{L+t} v_i^j a_i, \qquad \forall j \in \Gamma_4^* \tag{3.6}$$

$$\text{(ii)} \quad \sigma\lambda_i + 1 + \sum_{j=l+1}^{L} |\beta_i^j| + \sum_{j=L+t+1}^{L+T} \max\{0, \sigma\delta_i^j\} - \sum_{j\in\Gamma_4^*} \max\{0, \sigma u_i^j\} \geq 0, \quad i = 1, \cdots, l \tag{3.7}$$

$$\sigma\alpha_i + \max\{0, \sigma\} + \sum_{j=L+t+1}^{L+T} \max\{0, \sigma\rho_i^j\} - \sum_{j\in\Gamma_4^*} \max\{0, \sigma v_i^j\} \geq 0, \quad i = L+1, \cdots, L+t \tag{3.8}$$

where $\sigma = 1$ or -1. Moreover if the point x^* is a strictly local minimizer of F, then the following condition also holds
(iii) The inequalities in (ii) are all strict,and $l + t = n$.

Proof. Suppose x^* is a local minimizer of F. Then $0 \in \partial F(x^*) \subseteq G(x^*)$, and it follows from (2.4) that there exist multipliers $\{\widetilde{\lambda}_i\}$ such that

$$0 = g(x^*) + \sum_{i\in\Gamma_{345}^*} \widetilde{\lambda}_i a_i, \quad \text{or} \quad g^* = \sum_{i\in\Gamma_{345}^*} (-\widetilde{\lambda}_i) a_i$$

As a descent direction does not exist at x^*,Theorem 3.2 implies $a_j \in R(A^*)$ for all $j \in \Gamma_4^*$.Thus it follows from (3.3) that there exist multipliers $\{\lambda_i\}, \{\alpha_i\}, \{u_i^j\}$and $\{v_i^j\}$ satisfying (3.6) uniquely. Futhermore ,there exist multipliers $\{\beta_i^j\}, \{\delta_i^j\}$and $\{\rho_i^j\}$ such that (3.4) and (3.5) are satisfied.

Now suppose that there exists an index $k \in \{1, \cdots, l, L+1, \cdots, L+t\}$ such that

$$\sigma\lambda_k + 1 + \sum_{j=l+1}^{L} |\beta_k^j| + \sum_{j=L+t+1}^{L+T} \max\{0, \sigma\delta_k^j\} - \sum_{j\in\Gamma_4^*} \max\{0, \sigma u_k^j\} < 0, \quad \text{if} \quad k \in \{1, \cdots, l\}$$

or

$$\sigma\alpha_k + \max\{0, \sigma\} + \sum_{j=L+t+1}^{L+T} \max\{0, \sigma\rho_k^j\} - \sum_{j\in\Gamma_4^*} \max\{0, \sigma v_k^j\} < 0, \quad \text{if} \quad k \in \{L+1, \cdots, L+t\}$$

where $\sigma = 1$ or -1. Let s be a solution of the following linear equations

$$\begin{cases} a_k^T s = \sigma \\ a_i^T s = 0, \quad i = 1, \cdots, l, L+1, \cdots L+t, \quad i \neq k \end{cases}$$

Then in view of (3.1) and (3.6), we obtain

$$\begin{aligned}
F'(x^*,s) &= \sum_{i=1}^{l}\lambda_i a_i^T s+\sum_{i=L+1}^{L+t}\alpha_i a_i^T s+\sum_{i=1}^{l}|a_i^T s|+\sum_{j=l+1}^{L}|\sum_{i=1}^{l}\beta_i^j a_i^T s|+\sum_{i=L+1}^{L+t}\max\{0,a_i^T s\}\\
&+\sum_{j=L+t+1}^{L+T}\max\{0,\sum_{i=1}^{l}\delta_i^j a_i^T s+\sum_{i=L+1}^{L+t}\rho_i^j a_i^T s\}-\sum_{j\in\Gamma_4^*}\max\{0,\sum_{i=1}^{l}u_i^j a_i^T s+\sum_{i=L+1}^{L+t}v_i^j a_i^T s\}\\
&=\begin{cases}\sigma\lambda_k+1+\sum_{j=l+1}^{L}|\beta_k^j|+\sum_{j=L+t+1}^{L+T}\max\{0,\sigma\delta_k^j\}-\sum_{j\in\Gamma_4^*}\max\{0,\sigma u_k^j\},\\ \qquad \text{if}\quad k\in\{1,\cdots,l\}\\ \sigma\alpha_k+\max\{0,\sigma\}+\sum_{j=L+t+1}^{L+T}\max\{0,\sigma\rho_k^j\}-\sum_{j\in\Gamma_4^*}\max\{0,\sigma v_k^j\},\\ \qquad \text{if}\quad k\in\{L+1,\cdots,L+t\}\end{cases}\\
&<0
\end{aligned}$$

This contradicts the fact that $x^*$is a local minimizer of F, hence (3.7) and (3.8) hold.

Finally suppose x^* is a strictly local minimizer of F. Then exactly as above we can establish conditions (3.6)-(3.8). Suppose that the condition (iii) does not hold,that is (a): $l+t<n$, or (b): one of the inequalities (3.7) or (3.8) is equality.For case(a),one can choose any nonzero $s\in R(A^*)^{\perp}$. for case (b), there exists an index $k\in\{1,\cdots,l,L+1,\cdots,L+t\}$ which satisfies (3.7) or (3.8) with equality, and a nonzero s such that

$$\begin{cases} a_k^T s=\sigma\\ a_i^T s=0\quad i=1,\cdots,l,L+1,\cdots L+t,\quad i\neq k\end{cases}$$

where $\sigma=1$ or -1. Then following the proof above, we can find a direction s from (a) or (b) such that $F'(x^*,s)=0$, contradicting the fact that x^* is a strictly local minimizer of F. The proof is completed.

Theorem 3.4. *Suppose $R(A^*)$ has a basis (3.3) at a point x^*. Then x^* is a local minimizer of F if there exist unique multipliers $\{\lambda_i\}$, $\{\alpha_i\}$, $\{\beta_i^j\}$, $\{\delta_i^j\}$, $\{\rho_i^j\}$, $\{u_i^j\}$ and $\{v_i^j\}$ satisfying*

(i) $g^*=\sum_{i=1}^{l}\lambda_i a_i+\sum_{i=L+1}^{L+t}\alpha_i a_i,\qquad a_j=\sum_{i=1}^{l}\beta_i^j a_i,\qquad j=l+1,\cdots,L,$

$$a_j=\sum_{i=1}^{l}\delta_i^j a_i+\sum_{L+1}^{L+t}\rho_i^j a_i,\qquad j=L+t+1,\cdots,L+T,$$

$$a_j=\sum_{i=1}^{l}u_i^j a_i+\sum_{i=L+1}^{L+t}v_i^j a_i,\qquad \forall j\in\Gamma_4^*.$$

(ii)

$$\begin{cases}\sigma\lambda_i+1+\sum_{j=l+1}^{L}\sigma\beta_i^j+\sum_{L+t+1}^{L+T}\sigma\delta_i^j-\sum_{j\in\Gamma_4^*}\max\{0,\sigma u_i^j\}\geq 0,\qquad i=1,\cdots,l & (3.9)\\ \sigma\alpha_i+\max\{0,\sigma\}+\sum_{j=L+t+1}^{L+T}\sigma\rho_i^j-\sum_{j\in\Gamma_4^*}\max\{0,\sigma v_i^j\}\geq 0,\qquad i=L+1,\cdots,L+t & (3.10)\end{cases}$$

where $\sigma=1$ or -1. Moreover, the point x^* is a strictly local minimizer of F if the following condition also holds.

(iii) the inequalities (3.9) and (3.10) are all strict, and $l+t=n$.

Proof. It is obvious that x^* is a local minimizer of F if and only if $F'(x^*,s)\geq 0$ for all $s\in R^n$. Furthermore, x^* is a strictly local minimizer of F if and only if $F'(x^*,s)>0$ for all $s\neq 0$. Now using the simple fact that for any $a,b\in R,\max\{0,a\}\geq a$ and $\max\{0,a+b\}\leq$

$\max\{0,a\}+\max\{0,b\}$, as well as condition (i) and (3.1), we have

$$\begin{aligned}
&F'(x^*,s)\\
&= g^{*T}s + \sum_{i\in\Gamma_5^*}|a_i^Ts| + \sum_{i\in\Gamma_3^*}\max\{0,a_i^Ts\} - \sum_{j\in\Gamma_4^*}\max\{0,a_j^Ts\}\\
&= \sum_{i=1}^{l}\lambda_i a_i^Ts + \sum_{i=L+1}^{L+t}\alpha_i a_i^Ts + \sum_{i=1}^{l}|a_i^Ts| + \sum_{j=l+1}^{L}|\sum_{i=1}^{l}\beta_i^j a_i^Ts| + \sum_{i=L+1}^{L+t}\max\{0,a_i^Ts\}\\
&+ \sum_{i=L+t+1}^{L+T}\max\{0,\sum_{i=1}^{l}\delta_i^j a_i^Ts + \sum_{i=L+1}^{L+t}\rho_i^j a_i^Ts\} - \sum_{j\in\Gamma_4^*}\max\{0,\sum_{i=1}^{l}u_i^j a_i^Ts + \sum_{i=L+1}^{L+t}v_i^j a_i^Ts\}\\
&\geq \sum_{i=1}^{l}\lambda_i a_i^Ts + \sum_{i=L+1}^{L+t}\alpha_i a_i^Ts + \sum_{i=1}^{l}|a_i^Ts| + \sum_{l=l+1}^{L}\sum_{i=1}^{l}\beta_i^j a_i^Ts + \sum_{i=L+1}^{L+t}\max\{0,a_i^Ts\}\\
&+ \sum_{i=L+t+1}^{L+T}(\sum_{i=1}^{l}\delta_i^j a_i^Ts + \sum_{i=L+1}^{L+t}\rho_i^j a_i^Ts) - \sum_{j\in\Gamma_4^*}[\sum_{i=1}^{l}\max\{0,u_i^j a_i^Ts\} + \sum_{i=L+1}^{L+t}\max\{0,v_i^j a_i^Ts\}]\\
&= \sum_{i=1}^{l}[\lambda_i\text{sign}(a_i^Ts) + 1 + \text{sign}(a_i^Ts)(\sum_{j=l+1}^{L}\beta_i^j + \sum_{j=L+t+1}^{L+T}\delta_i^j) - \sum_{i\in\Gamma_4^*}\max\{0,\text{sign}(a_i^Ts)u_i^j\}]\\
&\cdot|a_i^Ts| + \sum_{i=L+1}^{L+t}[\alpha_i\text{sign}(a_i^Ts) + \max\{0,\text{sign}(a_i^Ts)\} + \text{sign}(a_i^Ts)\sum_{i=L+t+1}^{L+T}\rho_i^j\\
&- \sum_{j\in\Gamma_4^*}\max\{0,\text{sign}(a_i^Ts)v_i^j\}]|a_i^Ts|\\
&= \sum_{i=1}^{l}\xi_i|a_i^Ts| + \sum_{i=L+1}^{L+t}\eta_i|a_i^Ts| \geq 0, \qquad \forall s\in R^n
\end{aligned}$$

where from condition(ii)

$$\xi_i = \begin{cases} 1, & \text{if } \ a_i^Ts=0,\\ \lambda_i + 1 + \sum_{j=l+1}^{L}\beta_i^j + \sum_{j=L+t+1}^{L+T}\delta_i^j - \sum_{j\in\Gamma_4^*}\max\{0,u_i^j\}\geq 0, & \text{if } \ a_i^Ts>0,\\ -\lambda_i + 1 - \sum_{j=l+1}^{L}\beta_i^j - \sum_{j=L+t+1}^{L+T}\delta_i^j - \sum_{j\in\Gamma_4^*}\max\{0,-u_i^j\}\geq 0, & \text{if } \ a_i^Ts<0,\end{cases}$$
$$i=1,\cdots,l$$

and

$$\eta_i = \begin{cases} 0, & \text{if } \ a_i^Ts=0,\\ \alpha_i + 1 + \sum_{j=L+t+1}^{L+T}\rho_i^j - \sum_{j\in\Gamma_4^*}\max\{0,v_i^j\}\geq 0, & \text{if } \ a_i^Ts>0,\\ -\alpha_i - \sum_{j=L+t+1}^{L+T}\rho_i^j - \sum_{j\in\Gamma_4^*}\max\{0,-v_i^j\}\geq 0, & \text{if } \ a_i^Ts<0,\end{cases}$$
$$i=L+1,\cdots,L+t$$

Thus x^* is a local minimizer of F. Moreover if condition (iii) also holds then A^* is nonsingular and $A^{*T}s\neq 0$ for any $s\neq 0$. Therefore there exists an index i such that $a_i^Ts\neq 0$ and $\xi_i>0$ or $\eta_i>0$, and hence $F'(x^*,s)>0$ for any $s\neq 0$, which implies x^* to be a strictly local minimizer of F. The proof is completed.

4. Corollary and Application

In this section, we shall introduce the regularity to nonsmooth function $F(x)$, and then give the sufficient and necessary conditions at a local minimizer x^* of F for two special problem.

Definition 4.1. *F is said to be regular at x if the vectors $a_i=\nabla f_i(x), i\in\Gamma_{35}$, are linearly independent.*

The following results are only a corollary of Theorem 3.3 and Theorem 3.4.

Corollary 4.1. *If F is regular at x^* then x^* is a local minimizer of F if and only if there exist unique multipliers $\{\lambda_i\}$, $\{\alpha_i\}$, $\{u_i^j\}$ and $\{v_i^j\}$ satisfying*

(i) $g^* = \sum_{i\in\Gamma_5^*} \lambda_i a_i + \sum_{i\in\Gamma_3^*} \alpha_i a_i, \qquad a_j = \sum_{i\in\Gamma_5^*} u_i^j a_i + \sum_{i\in\Gamma_3^*} v_i^j a_i, \qquad \forall j \in \Gamma_4^*,$

(ii) $\begin{cases} \sigma\lambda_i + 1 - \sum_{j\in\Gamma_4^*} \max\{0, \sigma u_i^j\} \geq 0, & \forall i \in \Gamma_5^* \\ \sigma\alpha_i + \max\{0,\sigma\} - \sum_{j\in\Gamma_4^*} \max\{0, \sigma v_i^j\} \geq 0, & \forall i \in \Gamma_3^* \end{cases}$

where $\sigma = 1$ or -1. Moreover, x^* is a strictly local minimizer of F if and only if the following condition also holds.

(iii)$l + t = n$ and inequalities in (ii) are strict.

Proof. As F is regular at x^*, one has $l = L$ and $t = T$. The conclusion is trivial from Theorem 3.3 and Theorem 3.4.

Corollary 4.2. *If F is regular at x^* and $c_i \geq 0, i = 1, \cdots, m$. Then x^* is a local minimizer of F if and only if there exist unique multipliers $\{\alpha_i\}$ satisfying*

(i) $g^* = \sum_{i\in\Gamma_3^*} \alpha_i a_i,$

(ii) $\alpha_i \in [-1, 0], \qquad \forall i \in \Gamma_3^*.$

Moreover, the point x^* is a strictly local minimizer of F if and only if the following condition also holds

(iii) $\alpha_i \neq 0, \forall i \in \Gamma_3^*$, and rank$(A^*) = n$.

Proof. As $c_i \geq 0, i = 1, \cdots, m$, it implies $\Gamma_{45} = \phi$ for any $x \in R^n$, and $F(x) = \sum_{i=1}^{m} \max\{f_i(x), c_i\}$. Moreover, F is regular at x^*, one has $t = T$. Thus the result is a corollary of Theorem 3.3 and Theorem 3.4.

Now consider the censored discrete linear L_1 approximation problem [8]

$$\min F(x) = \sum_{i=1}^{m} |y_i - \max\{z_i, a_i^T x\}| \tag{4.1}$$

where $y_i, z_i \in R$ and $a_i \in R^n, i = 1, \cdots, m$. It is obvious that (4.1) is a special case of (1.1) under conditions of $f_i(x) = a_i^T x - y_i$ and $c_i = z_i - y_i$ for all i. That is to say the conclusions in [8] are only the corollaries of this paper.

For the constrained optimization problem

$$\begin{aligned} &\min f(x) \\ &s.\ t. \quad g_j(x) \leq 0 \quad j = 1, 2, \cdots, m \end{aligned} \tag{4.2}$$

Where $f, g_j \in C^1, j = 1, \cdots, m$, the well-known exact penalty function is

$$\min F(x) = f(x) + \lambda \sum_{j=1}^{m} \max\{g_j(x), 0\} \tag{4.3}$$

where $\lambda > 0$ is a penalty parameter. Clearly, for the fixed λ the optimality conditions of (4.3) are only the special cases of Theorem 3.3, Theorem 3.4 and Corollary 4.2.

References

[1] Aubin J.-P., Optima and Equilibria-An Introduction to Nonlinear Analysis, Springer-Verlag, Berlin, 1993.

[2] Barrodale, I. and Roberts, F. D. K., An improved algorithm for discrete L_1 linear approximation, *SIAM J. Numer. Anal.,* **10** (1973), 839-848.

[3] Bartels, R. H., Conn, A. R. and Sinclair, J. W., Minimization techniques for piecewise differentiable functions: the L_1 solution of an overdetermined linear systems, *SIAM J. Numer. Anal.,* **15** (1978), 224-241.

[4] Clarke, F. H., Optimization and Nonsmooth Analysis, John Wiley & Sons, 1983.

[5] Evtushenko, Yu. G., Numerical Optimization Techniques, Translation Editor: Stoer, J., Optimization Software, Inc., 1985.

[6] Fletcher, R., Practical Methods of Optimization, Vol.2, Constrained Optimization, John Wiley & Sons, 1981.

[7] Powell, J. L., Least absolute deviations estimation for censored and truncated regression models, Institute for Mathematical Studies in the Social Sciences, the Economics Series, Technical Report 356, Stanford University, 1981.

[8] Womereley, R.S., Censord discrete linear L_1 approximation, CAM-R42-83, The Australian National University.