Vietnam Journal of Mathematics **30**:1 (2002) 55–69

Vietnam Journal of MATHEMATICS
© NCST 2002

New Lagrange Multipliers Rules for Constrained Quasidifferentiable Optimization

Yan Gao

School of Management,
University of Shanghai for Science and Technology,
Jungong Road, Shanghai 200093, China

Received August 12, 2000
Revised November 28, 2000

Abstract. New first-order necessary optimality conditions with Lagrange multipliers for quasidifferentiable optimization with equality and inequality constraints are proposed. Kuhn–Tucker sufficient optimality conditions are also developed. Two kinds of differences for two convex compact sets, proposed by Demyanov and by Rubinov and Akhundov, respectively, are used.

1. Introduction

Quasidifferential calculus, developed by Demyanov and Rubinov, plays an important role in nonsmooth analysis and optimization. Since it is closely related to the classical directional derivative, quasidifferential could be used to describe the behavior of extreme point more strongly, and could be easily calculated in some cases. As we know, the early necessary optimality conditions in geometric form for quasidifferentiable optimization were proposed by Polyakova [15] and Shapiro [18]. The versions with Lagrange multipliers for the inequality constrained problem were initially developed by Eppler and Luderer [7] and further investigated by Gao [8] and Luderer [13]. For the equality and inequality constrained case, optimality conditions with Lagrange multipliers were studied by Gao [9] and Yin and Zhang [19]. The optimality conditions in [9] are presented by means of the Demyanov difference of subdifferential and minus superdifferential, where the subdifferential and the superdifferential is in the sense of quasidifferential calculus. It is well known that the optimality conditions with Lagrange multipliers have many advantages over those of geometric form.

The aim of this paper is to explore optimality conditions with Lagrange multipliers for quasidifferentiable optimization with equality and inequality constraints. Three necessary optimality conditions, which improve, in some sense, those given in [9], and corresponding sufficient optimality conditions are proposed. Two kinds of differences of two convex compact sets, proposed by Demyanov [3] and by Rubinov and Akhundov [17], respectively, are used. The remainder of this paper is organized as follows: In Sec. 2, preliminaries on quasidifferential calculus are recalled. In Sec. 3, three necessary optimality conditions are proposed. In Sec. 4, Kuhn–Tucker sufficient optimality conditions are developed.

2. Preliminaries

According to the definition in [5], $f : \mathbb{R}^n \to \mathbb{R}$ is called quasidifferentiable at a point $x \in \mathbb{R}^n$ (in the sense of Demyanov and Rubinov), if it is directionally differentiable at x, i.e., the directional derivative

$$f'(x; d) = \lim_{t \to 0^+} \frac{1}{t}[f(x + td) - f(x)], \quad d \in \mathbb{R}^n \tag{2.1}$$

is well-defined; moreover, the function $f'(x; \cdot)$ is representable as the difference of two sublinear functions. In other words, there exists a pair of convex compact sets $\underline{\partial} f(x), \overline{\partial} f(x) \subset \mathbb{R}^n$ such that

$$f'(x; d) = \max_{v \in \underline{\partial} f(x)} v^T d + \min_{w \in \overline{\partial} f(x)} w^T d, \quad d \in \mathbb{R}^n. \tag{2.2}$$

The pair of sets $Df(x) = [\underline{\partial} f(x), \overline{\partial} f(x)]$ is called a quasidifferential of f at x, $\underline{\partial} f(x)$ and $\overline{\partial} f(x)$ are called a subdifferential and a superdifferential, respectively.

f is said to be uniformly directionally differentiable at x, if the convergence in (2.1) holds uniformly with respect to any unit vector d. It was shown that the directional differentiability is equivalent to the uniformly directional differentiability for a locally Lipschitzian function [5]. Obviously, the quasidifferential is not uniquely defined. Actually, suppose that $[U, V]$ is a quasidifferential of f at x, then for any convex compact set $S \subset \mathbb{R}^n$, the pair of sets $[U + S, V - S]$ is also a quasidifferential of f at x. The class of quasidifferentiable functions contains convex, concave and differentiable functions, but also convex-concave, maximum and other functions. It even contains some functions, which are not locally Lipschitzian.

Let $f : \mathbb{R}^n \to \mathbb{R}$ be locally Lipschitzian in a neighborhood of x. By the definition in [1], the (Clarke) generalized gradient of f at x, denoted by $\partial_{\text{Cl}} f(x)$, is of the form:

$$\partial_{\text{Cl}} f(x) = \text{co}\{u \in \mathbb{R}^n \mid u = \lim_{x_n \to x} \nabla f(x_n), \nabla f(x_n) \text{ exists}, \ x_n \to x\},$$

where "co" denotes convex hull.

In what follows, we review some of related concepts from [5, 11, 16, 17].

Let $S \subset \mathbb{R}^n$ be a convex compact set.

$$P_S(x) = \max_{u \in S} u^T x, \quad x \in \mathbb{R}^n$$

is called the support function of the set S. It is true that P_S is a convex function on $\mathbb{R}^n$ with

$$\partial P_S(x) = \{u \in S \mid u^T x = P_S(x)\}, \tag{2.3}$$

particularly $\partial P_S(0) = S$, where "∂" denotes subdifferential in the sense of convex analysis. From (2.3), it follows that P_S is differentiable at x if and only if the right hand side of (2.3) is a singleton. Given a point $x \in \mathbb{R}^n$, if the set $\{u \in S \mid u^T x = P_S(x)\}$ is a singleton, denoted by $\{u_0\}$, where $u_0 \in S$, then $\nabla P_S(x) = u_0$.

A set $T \subset \mathbb{R}^n$ is called of full measure (with respect to $\mathbb{R}^n$), if $\mathbb{R}^n \setminus T$ is a set of measure zero. Let $U, V \subset \mathbb{R}^n$ be two convex compact sets and $T \subset \mathbb{R}^n$ be a full measure set such that their support functions P_U and P_V are differentiable at every point $x \in T$. The set $U \dot{-} V$, called the Demyanov difference of U and V, is defined as the following:

$$U \dot{-} V = \text{clco}\{\nabla P_U(x) - \nabla P_V(x) \mid x \in T\}, \tag{2.4}$$

where "cl" and "co" denote closure and convex hull, respectively. It has been shown that $U \dot{-} V$ does not depend on the specific choice of the set T, so it is well-defined.

The Demyanov difference was implicitly introduced by Demyanov [3] in order to establish a relation between quasidifferential and the generalized gradient. According to [5, 17], the following relation holds:

$$\partial_{\text{Cl}} f'(x;y)|_{y=0} = \underline{\partial} f(x) \dot{-} (-\overline{\partial} f(x)). \tag{2.5}$$

Therefore, if both $[U_1, V_1]$ and $[U_2, V_2]$ are quasidifferentials of f at x, then $U_1 \dot{-} (-V_1) = U_2 \dot{-} (-V_2)$. That is to say the set $\underline{\partial} f(x) \dot{-} (-\overline{\partial} f(x))$ is independent from the specific choice of the quasidifferential.

Let S be a set in $\mathbb{R}^n$. Given a point $x \in \mathbb{R}^n$, put

$$G_x(S) = \{u \in S \mid u^T x = P_S(x)\},$$

$$\widetilde{G}_x(S) = \{u \in S \mid u^T x = \min_{u \in S} u^T x\}.$$

The set $G_x(S)$ and $\widetilde{G}_x(S)$ are called the max-face and min-face of the set S generated by x, respectively.

Let U and V be convex compact sets, the operation $\ddot{-}$ of U and V, proposed by Rubinov and Akhundov [17], is defined by

$$U \ddot{-} V = \text{clco} \bigcup_{x \neq 0} [G_x(U) - G_x(V)]. \tag{2.6}$$

We call $U \ddot{-} V$ the Rubinov difference of U and V.

Let $T \subset \mathbb{R}^n$, we say the set T has the property $(\mathcal{E})$ with respect to the pair of sets $[U, V]$, if the measure of the set $\mathbb{R}^n \setminus T$ is zero, and both $G_x(U)$ and $\widetilde{G}_x(V)$ for any $x \in T$ are singletons.

Let $M(x)$ denote a family of functions f defined on an open set $X \subset \mathbb{R}^n$ with $x \in X$ satisfying:

(a) f is locally Lipschitzian in a neighborhood of x, denoted by $N(x, \delta)$.

(b) f is quasidifferentiable at x.

(c) There exist a subset $Q \subset V_f$ of full measure (with respect to $N(x, \delta)$), where V_f is the set of the point $y \in N(x, \delta)$ where the gradient of f exists at y, and a set T possessing the property $(\mathcal{E})$ with respect to the pair of sets $[\underline{\partial} f(x), \overline{\partial} f(x)]$ such that the relation

$$g_k \to g, \ t_k \to 0^+, \ x_k = x + t_k g_k \in Q, \ g \in T,$$

implies

$$\nabla f(x_k) \to \operatorname{Arg} \max_{u \in \underline{\partial} f(x)} u^T g + \operatorname{Arg} \min_{w \in \overline{\partial} f(x)} w^T g.$$

The following relation between quasidifferential and the generalized gradient was obtained by Rubinov and Akhundov [17], see also [5]:

$$\underline{\partial} f(x) \dot{-} (-\overline{\partial} f(x)) \subset \partial_{\mathrm{Cl}} f(x), \ \forall f \in M(x). \tag{2.7}$$

Along with $M(x)$, let us consider the function family $\widetilde{M}(x)$ consisting of functions having the properties (a), (b) and the following property (d):

(d) One can find: (1) a subset $Q \subset N(x, \delta)$ of full measure (with respect to $N(x, \delta)$) at every point y of which the gradient $\nabla f(y)$ exists; (2) a quasidifferential $[\underline{\partial} f(x), \overline{\partial} f(x)]$ of the function f at the point x such that the relations

$$g_k \to g, \ t_k \to 0^+, \ x_k = x + t_k g_k \in Q, \ \nabla f(x_k) \to v$$

imply $v \in G_g(U) + \widetilde{G}_g(V)$.

It is true that $\widetilde{M}(x) \subset M(x)$. Any one of families $\widetilde{M}(x)$ and $M(x)$ contains convex, concave functions and maximum of smooth functions. If $f_1, \dots, f_m$ belong to $M(x)$(or $\widetilde{M}(x)$), then $\max_{1 \le i \le m} f_i$ and $g(f_1(\cdot), \dots, f_m(\cdot))$ still belong to $M(x)$ (or $\widetilde{M}(x)$), where g is continuously differentiable on $\mathbb{R}^m$. The relation below holds:

$$\partial_{\mathrm{Cl}} f(x) \subset \underline{\partial} f(x) \ddot{-} (-\overline{\partial} f(x)), \ \forall f \in \widetilde{M}(x); \tag{2.8}$$

moreover,

$$\underline{\partial} f(x) \dot{-} (-\overline{\partial} f(x)) \subset \partial_{\mathrm{Cl}} f(x) \subset \underline{\partial} f(x) \ddot{-} (-\overline{\partial} f(x)), \ \forall f \in \widetilde{M}(x).$$

Now, let us consider the following problem:

$$\begin{aligned} (\mathrm{P}_1) \qquad & \text{minimize} \ \ f_0(x), \\ & \text{subject to} \ \ f_i(x) \le 0, \ i = 1, \dots, m, \\ & \qquad\qquad\quad h_j(x) = 0, \ j = 1, \dots, p, \end{aligned}$$

where $f_i, h_j : \mathbb{R}^n \to \mathbb{R}$ for $i = 0, 1, \dots, m$, $j = 1, \dots, p$ are quasidifferentiable.

Given a point $x \in \mathbb{R}^n$, denote

$$\Gamma(x) = \Big\{ y \in \mathbb{R}^n \mid \exists \lambda \geq 0,\, x_k \to x,\, x_k \neq x,\\ h_j(x_k) = 0,\, 1 \leq j \leq p,\, \frac{x_k - x}{\| x_k - x \|} \to y_0,\, y = \lambda y_0 \Big\},$$

$$\gamma(x) = \{ y \in \mathbb{R}^n \mid h_j'(x; y) = 0,\, j = 1, \ldots, p \}.$$

The following two hypotheses for h_j $(j = 1, \ldots, p)$ at a point will be used to establish necessary optimality conditions for the problem (P_1).

Hypothesis 1. $\Gamma(x) = \gamma(x)$.

Hypothesis 2. For every $y \in \gamma(x)$, both $G_y(\underline{\partial} h_j(x))$ and $\widetilde{G}_y(\overline{\partial} h_j(x))$ are singletons, denoted by $G_y(\underline{\partial} h_j(x)) = \{a_j(x, y)\}$ and $\widetilde{G}_y(\overline{\partial} h_j(x)) = \{b_j(x, y)\}$, respectively. Moreover, vectors $a_1(x, y) + b_1(x, y), \ldots, a_p(x, y) + b_p(x, y)$ are linearly independent.

It is easy to see that both $G_y(\underline{\partial} h_j(x)) = \{a_j(x, y)\}$ and $\widetilde{G}_y(\overline{\partial} h_j(x)) = \{b_j(x, y)\}$ imply that functions $\max_{v \in \underline{\partial} h_j(x)} v^T y$ and $\min_{w \in \overline{\partial} h_j(x)} w^T y$ are differentiable at y with

$$\nabla_y \max_{v \in \underline{\partial} h_j(x)} v^T y = a_j(x, y), \;\; j = 1, \ldots, p$$

and

$$\nabla_y \min_{w \in \overline{\partial} h_j(x)} w^T y = b_j(x, y), \;\; j = 1, \ldots, p.$$

Thus, the function $h_j'(x; \cdot)$ is differentiable at y with

$$\nabla_y h_j'(x; y) = a_j(x, y) + b_j(x, y), \;\; j = 1, \ldots, p.$$

On the other hand, the fact that $a_1(x, y) + b_1(x, y), \ldots, a_p(x, y) + b_p(x, y)$ are linearly independent implies that $p \leq n$ and Jacobian of $(h_1'(x; y), \ldots, h_p'(x; y))^T$ with y as variable is of full rank. Hypotheses 1 and 2 were initially proposed to derive the optimality conditions of geometric form for the problem (P_1) by Polyakova [15] and Shapiro [18], respectively.

We now give some notations, which will be often used later on. Denote

$$I(x) = \{ i \mid f_i(x) = 0, i = 1, \ldots, m \}, \tag{2.9}$$

denote $U \dot{+} V$ for $U \dot{-} (-V)$ and denote $U \ddot{+} V$ for $U \ddot{-} (-V)$.

Proposition 2.1 [9]. *Suppose that $\overline{x}$ is a minimizer to the problem (P_1), f_i for $i = 1, \ldots, m$ are uniformly directionally differentiable at $\overline{x}$ and at least one of Hypothesis 1 and Hypothesis 2 holds. Denote*

$$f(x) = \max\{ f_0(x) - f_0(\overline{x}), f_i(x) \mid i \in I(\overline{x}) \}.$$

Then $y = 0$ is a minimizer to the following problem:

$$(\mathrm{P}_2) \qquad \begin{array}{ll} \text{minimize} & f'(\overline{x}; y), \\ \text{subject to} & h'_j(\overline{x}; y) = 0, \ j = 1, \ldots, p. \end{array}$$

Proposition 2.2 [9]. *Suppose that $\overline{x} \in \mathbb{R}^n$ is a minimizer to the problem (P_1), f_i for $i = 0, 1, \ldots, m$ and h_j for $j = 1, \ldots, p$ are uniformly directionally differentiable at $\overline{x}$ and at least one of Hypothesis 1 and Hypothesis 2 holds at $\overline{x}$. Then, there exist scalars $\lambda_i \geq 0$, $i = 0, 1, \ldots, m$, μ_j, $j = 1, \ldots, p$, not all zero, such that*

$$0 \in \sum_{i=0}^{m} \lambda_i(\underline{\partial} f_i(\overline{x}) \dot{+} \overline{\partial} f_i(\overline{x})) + \sum_{j=1}^{p} \mu_j(\underline{\partial} h_j(\overline{x}) \dot{+} \overline{\partial} h_j(\overline{x})), \tag{2.10}$$

$$\lambda_i f_i(\overline{x}) = 0, \ i = 1, \ldots, m. \tag{2.11}$$

Proposition 2.3 [9]. *Suppose that $\overline{x} \in \mathbb{R}^n$ is a minimizer to the problem (P_1), and h_j for $j = 1, \ldots, p$ are locally Lipschitzian in a neighborhood of $\overline{x}$. Then, for any set of $w_i \in \overline{\partial} f_i(\overline{x})$, $i = 0, 1, \ldots, m$, there exist scalars $\lambda_i(w) \geq 0$, $i = 0, 1, \ldots, m$, $\mu_j(w)$, $j = 1, \ldots, p$, not all zero, such that*

$$0 \in \sum_{i=0}^{m} \lambda_i(w)(\underline{\partial} f_i(\overline{x}) + w_i) + \sum_{j=1}^{p} \mu_j(w) \partial_{\mathrm{Cl}} h_j(\overline{x}), \tag{2.12}$$

$$\lambda_i(w) f_i(\overline{x}) = 0, \ i = 1, \ldots, m, \tag{2.13}$$

where $\lambda_i(w)$, $i = 1, \ldots, m$ and $\mu_j(w)$, $j = 1, \ldots, p$ depend on the specific choice of $w = (w_0, w_1, \ldots, w_m)$.

Proposition 2.4 [19]. *Suppose that $\overline{x} \in \mathbb{R}^n$ is a minimizer to the problem (P_1), and f_i for $i = 0, 1, \ldots, m$ and h_j for $j = 1, \ldots, p$ are uniformly directionally differentiable at $\overline{x}$ and Hypothesis 1 holds at $\overline{x}$. Then, for any set of $w_i \in \overline{\partial} f_i(\overline{x})$, $i = 0, 1, \ldots, m$, there exist scalars $\lambda_i(w) \geq 0$, $i = 0, 1, \ldots, m$, $\mu_j(w)$, $j = 1, \ldots, p$, not all zero, such that*

$$0 \in \sum_{i=0}^{m} \lambda_i(w)(\underline{\partial} f_i(\overline{x}) + w_i) + \sum_{j=1}^{p} \mu_j(w)(\underline{\partial} h_j(\overline{x}) + \overline{\partial} h_j(\overline{x})), \tag{2.14}$$

$$\lambda_i f_i(\overline{x}) = 0, \ i = 1, \ldots, m, \tag{2.15}$$

where $\lambda_i(w)$, $i = 1, \ldots, m$ and $\mu_j(w)$, $j = 1, \ldots, p$ depend on the specific choice of $w = (w_0, w_1, \ldots, w_m)$.

3. Necessary Optimality Conditions

In this section, three necessary optimality conditions with Lagrange multipliers for the problem (P_1) are developed.

Theorem 3.1. *Suppose that $\overline{x} \in \mathbb{R}^n$ is a minimizer to the problem* (P$_1$), *and f_i for $i = 0, 1, \ldots, m$ and h_j for $j = 1, \ldots, p$ are uniformly directionally differentiable at $\overline{x}$ and at least one of Hypotheses 1 and 2 holds at $\overline{x}$. Then, for any set of $w_i \in \overline{\partial} f_i(\overline{x}), i = 0, 1, \ldots, m$, there exist scalars $\lambda_i(w) \geq 0, i = 0, 1, \ldots, m, \mu_j(w), j = 1, \ldots, p$, not all zero, such that*

$$0 \in \sum_{i=0}^{m} \lambda_i(w)(\underline{\partial} f_i(\overline{x}) + w_i) + \sum_{j=1}^{p} \mu_j(w)(\underline{\partial} h_j(\overline{x}) \dot{+} \overline{\partial} h_j(\overline{x})), \tag{3.1}$$

$$\lambda_i f_i(\overline{x}) = 0, \quad i = 1, \ldots, m, \tag{3.2}$$

where $\lambda_i(w)$, $i = 1, \ldots, m$ and $\mu_j(w)$, $j = 1, \ldots, p$ depend on the specific choice of $w = (w_0, w_1, \ldots, w_m)$.

Proof. It follows from Proposition 2.1 that $y = 0$ is a minimizer to the problem (P$_2$). Noting

$$f'(\overline{x}; y) = \max\{f'_i(\overline{x}; y) \mid i \in \{0\} \bigcup I(\overline{x})\},$$

one has that $(y, z) = 0$ is a minimizer to the following problem:

$$(\mathrm{P}_3) \qquad \text{minimize } z,$$
$$\text{subject to } f'_i(\overline{x}; y) - z \leq 0, i \in \{0\} \bigcup I(\overline{x}),$$
$$h'_j(\overline{x}; y) = 0, \; j = 1, \ldots, p,$$

where $y \in \mathbb{R}^n$, $z \in \mathbb{R}$, $(y, z) \in \mathbb{R}^{n+1}$. By virtue of (2.2), for any set of $w_i \in \overline{\partial} f_i(\overline{x})$, $i = 0, 1, \ldots, m$, the following relation holds:

$$\max_{u \in \underline{\partial} f_i(\overline{x})} u^T y + w_i^T y \leq f'_i(\overline{x}; y).$$

Hence, for any set of $w_i \in \overline{\partial} f_i(\overline{x})$, $i = 0, 1, \ldots, m$, (y, z) is a minimizer of the following problem:

$$(\mathrm{P}_4) \qquad \text{minimize } z,$$
$$\text{subject to } \max_{u \in \underline{\partial} f_i(\overline{x})} u^T y + w_i^T y \leq 0, i \in \{0\} \bigcup I(\overline{x})$$
$$h'_j(\overline{x}; y) = 0, \; j = 1, \ldots, p.$$

Evidently, all corresponding functions in (P$_4$) with $(y, z) \in \mathbb{R}^{n+1}$ as variable, are locally Lipschitzian. Hence, from the Fritz John necessary optimality condition in terms of the generalized gradient for (P$_4$) [1, Th. 6.1.1], it follows that there exist scalars $\overline{\lambda}(w) \geq 0, \lambda_i(w) \geq 0, i \in \{0\} \bigcup I(\overline{x}), \mu_j(w), j = 1, \ldots, p$, not all zero, such that

$$0 \in \overline{\lambda}(w) \partial_{\mathrm{Cl}} z \mid_{(y,z)=0} + \sum_{i \in \{0\} \cup I(\overline{x})} \lambda_i(w) \partial_{\mathrm{Cl}} (\max_{u \in \underline{\partial} f_i(\overline{x})} u^T y + w_i^T y) \mid_{(y,z)=0}$$
$$+ \sum_{j=1}^{p} \mu_j(w) \partial_{\mathrm{Cl}} h'_j(\overline{x}; y) \mid_{(y,z)=0}, \tag{3.3}$$

where $0 \in \mathbb{R}^{n+1}$. By virtue of (2.3), (2.5) and (3.3), we have that

$$0 \in \overline{\lambda}(w)(0,1)^T + \sum_{i\in\{0\}\cup I(\overline{x})} \lambda_i(w)(\underline{\partial} f_i(\overline{x}) + w_i, -1) + \sum_{j=1}^{p} \mu_j(w)(\underline{\partial} h_j(\overline{x}) \dot{+} \overline{\partial} h_j(\overline{x}), 0). \tag{3.4}$$

This leads to

$$0 \in \sum_{i\in\{0\}\cup I(\overline{x})} \lambda_i(w)(\underline{\partial} f_i(\overline{x}) + w_i) + \sum_{j=1}^{p} \mu_j(w)(\underline{\partial} h_j(\overline{x}) \dot{+} \overline{\partial} h_j(\overline{x})), \tag{3.5}$$

$$\overline{\lambda}(w) = \sum_{i\in\{0\}\cup I(\overline{x})} \lambda_i(w). \tag{3.6}$$

The relation (3.6) implies that $\lambda_i(w)$ for $i \in \{0\} \bigcup I(\overline{x})$ and $\mu_j(w)$ for $j = 1, \ldots, p$ are not all zero. (Otherwise, it follows from (3.6) that $\overline{\lambda}(w) = 0$, thus $\overline{\lambda}(w)$, $\lambda_i(w)$, $i \in \{0\} \bigcup I(\overline{x})$, $\mu_j(w)$, $j = 1, \ldots, p$ are all zero.) Taking $\lambda_i(w) = 0$ for $i \in \{1, \ldots, m\} \setminus I(\overline{x})$, we obtain (3.1) and (3.2). This completes the proof of the theorem. ■

If $h_j \in M(\overline{x})$ for $j = 1, \ldots, p$, the relation (3.1) yields the relation (2.12) because of $\underline{\partial} h_j(\overline{x}) \dot{+} \overline{\partial} h_j(\overline{x}) \subset \partial_{\mathrm{Cl}} h_j(\overline{x})$. This shows us that the optimality condition (3.1) is sharper than the one given in (2.12) in the case $h_j \in M(\overline{x})$.

In the light of the notion of the first order approximation and a related optimality condition, introduced by Ioffe [12], we can develop another optimality condition for the problem (P_1).

Given $f : \mathbb{R}^n \to \mathbb{R}$, we say $g : \mathbb{R}^n \to \mathbb{R}$ to be a first order approximation for f at $x \in \mathbb{R}^n$, provided that

$$g(ty) = tg(y), \quad \forall t \geq 0, \ y \in \mathbb{R}^n \tag{3.7}$$

and

$$\limsup_{t\to 0^+} \frac{1}{t}[f(x+ty) - f(x) - tg(y)] \leq 0, \quad \forall y \in \mathbb{R}^n. \tag{3.8}$$

Proposition 3.1. [12, Th. 3] *Suppose that $\overline{x}$ is a minimizer to the problem (P_2), and h_j for $j = 1, \cdots, p$ are locally Lipschitzian in a neighborhood of $\overline{x}$ and each g_i is a convex and continuous first order approximation for f_i at $\overline{x}$. Then there exist scalars $\lambda_i \geq 0$, $i = 0, 1, \cdots, m$, μ_j, $j = 1, \cdots, p$, not all zero, such that*

$$0 \in \sum_{i=0}^{m} \lambda_i \partial g_i(0) + \sum_{j=1}^{p} \mu_j \partial_{\mathrm{Cl}} h_j(\overline{x}), \tag{3.9}$$

$$\lambda_i f_i(\overline{x}) = 0, \ i = 1, \ldots, m. \tag{3.10}$$

We now present an optimality condition for the problem (P$_1$) without assumptions except for the Lipschitzian property of h_j.

Theorem 3.2. *Suppose that $\overline{x} \in \mathbb{R}^n$ is a minimizer to the problem* (P$_1$), *h_j for $j = 1, \ldots, m$ are locally Lipschitzian in a neighborhood of $\overline{x}$. Then, there exist scalars $\lambda_i \geq 0$, $i = 0, 1, \ldots, m$, μ_j, $j = 1, \ldots, p$, not all zero, such that*

$$0 \in \sum_{i=0}^{m} \lambda_i(\underline{\partial} f_i(\overline{x}) \dot{+} \overline{\partial} f_i(\overline{x})) + \sum_{j=1}^{p} \mu_j \partial_{\mathrm{Cl}} h_j(\overline{x}), \tag{3.11}$$

$$\lambda_i f_i(\overline{x}) = 0, \quad i = 1, \ldots, m. \tag{3.12}$$

Proof. Let g_i be the support function of the set $\underline{\partial} f_i(\overline{x}) \dot{+} \overline{\partial} f_i(\overline{x})$, i.e.,

$$g_i(y) = \max_{v \in \underline{\partial} f_i(\overline{x}) \dot{+} \overline{\partial} f_i(\overline{x})} v^T y, \quad i = 0, 1, \ldots, m.$$

We next proceed to show $f_i'(x; y) \leq g_i(y)$. Let

$$T_i = \{y \in \mathbb{R}^n \mid G_y(\underline{\partial} f_i(\overline{x})),\ \widetilde{G}_y(\overline{\partial} f_i(\overline{x})) \text{ are singletons}\}, \quad i = 1, \ldots, m. \tag{3.13}$$

Denote

$$G_y(\underline{\partial} f_i(\overline{x})) = \{v_i(y)\}, \quad \forall y \in T_i,$$

$$\widetilde{G}_y(\overline{\partial} f_i(\overline{x})) = \{w_i(y)\}, \quad \forall y \in T_i.$$

Indeed, each T_i is the set where the convex function $\max_{u \in \underline{\partial} f_i(\overline{x})} u^T y$ and the concave function $\min_{u \in \overline{\partial} f_i(\overline{x})}$ are differentiable. Therefore, the set T_i is of full measure. We set $U = \underline{\partial} f_i(\overline{x})$, $V = -\overline{\partial} f_i(\overline{x})$ and $T = T_i$ in (2.4). It is easy to see that

$$v_i(y) + w_i(y) \in \underline{\partial} f_i(\overline{x}) \dot{+} \overline{\partial} f_i(\overline{x}).$$

Hence, one has that

$$\begin{aligned} f_i'(\overline{x}; y) &= (v_i(y) + w_i(y))^T y \\ &\leq g_i(y), \quad \forall y \in T_i. \end{aligned} \tag{3.14}$$

Since T_i is of full measure, as well as $f_i'(\overline{x}; y)$ with y as variable and g_i are continuous, it follows from (3.14) that

$$f_i'(\overline{x}; y) \leq g_i(y), \ \forall y \in \mathbb{R}^n.$$

This implies that

$$\limsup_{t \to 0^+} \frac{1}{t}[f_i(x + ty) - f_i(x) - t g_i(y)] \leq 0, \quad \forall y \in \mathbb{R}^n.$$

On the other hand, it is easy to see

$$g_i(ty) = tg_i(y), \ \ \forall t \geq 0, \ y \in \mathbb{R}^n.$$

By definition, each g_i is a first order approximation for f_i at $\overline{x}$. The convexity and continuity of g_i contribute to the fact that g_i is a support function. According to Proposition 3.1 and the formulation (2.3), we conclude the result of the theorem. ■

Based on Theorem 3.2 and properties of the functions in the family $\widetilde{M}(x)$, we obtain the following theorem immediately.

Theorem 3.3. *Suppose that $\overline{x} \in \mathbb{R}^n$ is a minimizer to the problem* (P_1) *and $h_j \in \widetilde{M}(\overline{x})$ for $j = 1, \ldots, p$. Then, there exist scalars $\lambda_i \geq 0$, $i = 0, 1, \ldots, m$, μ_j, $j = 1, \ldots, p$, not all zero, such that*

$$0 \in \sum_{i=0}^{m} \lambda_i(\underline{\partial} f_i(\overline{x}) \dot{+} \overline{\partial} f_i(\overline{x})) + \sum_{j=0}^{m} \lambda_j(\underline{\partial} h_j(\overline{x}) \ddot{+} \overline{\partial} h_j(\overline{x})), \tag{3.15}$$

$$\lambda_i f_i(\overline{x}) = 0, \ \ i = 1, \ldots, m. \tag{3.16}$$

We have obtained three necessary optimality conditions with Lagrange multipliers. These results improve those given in [9] and [19] in some sense. If $h_j \in M(x)$ then $(\underline{\partial} h_j(\overline{x}) \dot{+} \overline{\partial} h_j(\overline{x})) \subset \partial_{\mathrm{Cl}} h_j(\overline{x})$, therefore the relation (3.1) implies (2.12). That is to say (3.1) is sharper than (2.12) in the case $h_j \in M(x)$ for $j = 1, \ldots, m$. Besides, the relation $\underline{\partial} h_j(\overline{x}) \dot{+} \overline{\partial} h_j(\overline{x}) \subset \underline{\partial} h_j(\overline{x}) + \overline{\partial} h_j(\overline{x})$ always holds. Hence, (3.1) is always sharper than (2.14). Compared with (2.12), the Lagrange multipliers in (3.11) are independent of the specific choice of the set of elements of superdifferentials. In the optimality condition (3.15), the Lagrange multipliers are independent of the set of elements of superdifferentials, and only the quasidifferential is utilized. In [15] and [18], necessary optimality conditions were presented in a different form where Lagrange multipliers are not used.

4. Sufficient Optimality Conditions

This section is devoted to several sufficient optimality conditions for the problem (P_1) under the assumption of the so-called F_η-quasiconvexity and F_η-pseudoconvexity, which were used by Yin and Zhang [19] to establish sufficient optimality conditions corresponding to Proposition 2.4.

Definition 4.1. $F : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ *is called right sublinear, if for any fixed $x \in \mathbb{R}^n$, $F(x, \cdot)$ is sublinear.*

Definition 4.2. $f : \mathbb{R}^n \to \mathbb{R}$ *is called F_η-quasiconvex with respect to the convex compact set $S \subset \mathbb{R}^n$ at $\overline{x}$, if there exist a right sublinear function F and a vector* *mapping* $\eta : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ *such that for any* $x \in \mathbb{R}^n$, *the relation* $f(x) \leq f(\overline{x})$ *implies*

$$F(\eta(x,\overline{x}), w) < 0, \ \ \forall w \in S.$$

Definition 4.3. $f : \mathbb{R}^n \to \mathbb{R}$ *is called* F_η*-pseudoconvex with respect to the convex compact set* $S \subset \mathbb{R}^n$ *at* $\overline{x}$, *if there exist a right sublinear function* F *and a vector mapping* $\eta : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ *such that for any* $x \in \mathbb{R}^n$, *the relation* $f(x) < f(\overline{x})$ *implies*

$$F(\eta(x,\overline{x}), w) < 0, \ \ \forall w \in S.$$

Theorem 4.1. *Let* $\overline{x}$ *be a feasible solution of the problem* (P_1) *and let* $\overline{x}$ *satisfy the Kuhn–Tucker optimality condition corresponding to Theorem 3.2, i.e.,*

$$0 \in (\underline{\partial} f_0(\overline{x}) \dot{+} \overline{\partial} f_0(\overline{x})) + \sum_{i=1}^{m} \lambda_i (\underline{\partial} f_i(\overline{x}) \dot{+} \overline{\partial} f_i(\overline{x})) + \sum_{j=1}^{p} \mu_j \partial_{\mathrm{Cl}} h_j(\overline{x}) \tag{4.1}$$

and (3.12) *hold. Denote*

$$J^+ = \{j \mid \mu_j > 0, \ j = 1, \dots, p\},$$

$$J^- = \{j \mid \mu_j < 0, \ j = 1, \dots, p\}.$$

Suppose that there exist a right sublinear function F *and a vector mapping* $\eta : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$, *such that* f_0 *is* F_η*-pseudoconvex with respect to* $\underline{\partial} f_0(\overline{x}) \dot{+} \overline{\partial} f_0(\overline{x})$ *at* $\overline{x}$, f_i *is* F_η*-quasiconvex with respect to* $\underline{\partial} f_i(\overline{x}) \dot{+} \overline{\partial} f_i(\overline{x})$ *at* $\overline{x}$, h_j $(j \in J^+)$ *is* F_η*-quasiconvex with respect to* $\partial_{\mathrm{Cl}} h_j(\overline{x})$ *at* $\overline{x}$ *and* h_j $(j \in J^-)$ *is* F_η*-quasiconvex with respect to* $-\partial_{\mathrm{Cl}} h_j(\overline{x})$ *at* $\overline{x}$. *Then,* $\overline{x}$ *is a minimizer of the problem* (P_1).

Proof. By virtue of (4.1) and (3.12), there exist $w_i \in \underline{\partial} f_i(\overline{x}) \dot{+} \overline{\partial} f_i(\overline{x})$ and $v_j \in \partial_{\mathrm{Cl}} h_j(\overline{x})$ such that

$$\begin{aligned} 0 &\in w_0 + \sum_{i \in I(\overline{x})} \lambda_i w_i + \sum_{j=1}^{p} \mu_j v_j \\ &= w_0 + \sum_{i \in I(\overline{x})} \lambda_i w_i + \sum_{j \in J^+} \mu_j v_j + \sum_{j \in J^-} (-\mu_j) v_j. \end{aligned} \tag{4.2}$$

Suppose that x is a feasible point of (P_1), then $f_i(x) \leq 0 = f_i(\overline{x})$, $i \in I(\overline{x})$, $h_j(x) = 0 = h_j(\overline{x})$. Since f_0 is F_η-pseudoconvex with respect to $\underline{\partial} f_0(\overline{x}) \dot{+} \overline{\partial} f_0(\overline{x})$ at $\overline{x}$, f_i is F_η-quasiconvex with respect to $\underline{\partial} f_i(\overline{x}) \dot{+} \overline{\partial} f_i(\overline{x})$ at $\overline{x}$, h_j $(j \in J^+)$ is F_η-quasiconvex with respect to $\partial_{\mathrm{Cl}} h_j(\overline{x})$ at $\overline{x}$ and h_j $(j \in J^-)$ is F_η-quasiconvex with respect to $-\partial_{\mathrm{Cl}} h_j(\overline{x})$ at $\overline{x}$, it follows that

$$F(\eta(x,\overline{x}), w_i) \leq 0, \ \ i \in I(\overline{x}),$$

$$F(\eta(x,\overline{x}), v_j) \leq 0, \ \ j \in J^+,$$

$$F(\eta(x,\overline{x}),v_j)\leq 0,\ j\in J^-.$$

Because F is right sublinear, the relation (4.2) implies that

$$\begin{aligned}0&\leq F(\eta(x,\overline{x}),w_0+\sum_{i\in I(\overline{x})}\lambda_i w_i+\sum_{j=1}^{p}\mu_j v_j)\\&\leq F(\eta(x,\overline{x}),w_0)+\sum_{i\in I(\overline{x})}\lambda_i F(\eta(x,\overline{x}),w_i)\\&\quad+\sum_{j\in J^+}\mu_j F(\eta(x,\overline{x}),v_i)+\sum_{j\in J^-}(-\mu_j)F(\eta(x,\overline{x}),v_i).\end{aligned}\tag{4.3}$$

This yields

$$0\leq F(\eta(x,\overline{x}),w_0).$$

Noting that f_0 is F_η-pseudoconvex, we obtain $f_0(\overline{x})\leq f_0(x)$. That is to say $\overline{x}$ is a minimizer to the problem (P_1). ■

Similarly to Theorem 4.1, we obtain a Kuhn–Tucker sufficient optimality condition which corresponds to Theorem 3.1 as follows.

Theorem 4.2. *Let $\overline{x}$ be a feasible solution of the problem* (P_1) *and there be a set of $w_i\in\overline{\partial}f_i(\overline{x})$ for $i=0,1,\ldots,m$ such that the related Kuhn–Tucker optimality condition in the form of Theorem* 3.1, *i.e.,*

$$0\in\underline{\partial}f_0(\overline{x})+w_0+\sum_{i=1}^{m}\lambda_i(w)(\underline{\partial}f_i(\overline{x})+w_i)+\sum_{j=1}^{p}\mu_j(w)(\underline{\partial}h_j(\overline{x})\dot{+}\overline{\partial}h_j(\overline{x}))$$

and (3.2) *hold. Suppose that there exist a right sublinear function F and a vector mapping $\eta:\mathbb{R}^n\times\mathbb{R}^n\to\mathbb{R}^n$, such that f_0 is F_η-pseudoconvex with respect to $\underline{\partial}f_0(\overline{x})+w_0$ at $\overline{x}$, f_i is F_η-quasiconvex with respect to $\underline{\partial}f_i(\overline{x})+w_i$ at $\overline{x}$, h_j $(j\in J^+)$ is F_η-quasiconvex with respect to $\underline{\partial}h_j(\overline{x})\dot{+}\overline{\partial}h_j(\overline{x})$ at $\overline{x}$ and h_j $(j\in J^-)$ is F_η-quasiconvex with respect to $-(\overline{\partial}h_j(\overline{x})\dot{+}\overline{\partial}f(\overline{x}))$ at $\overline{x}$, where J^+ and J^- are defined as in Theorem* 4.1. *Then, $\overline{x}$ is a minimizer of the problem* (P_1).

5. Concluding Remarks

In the Lagrange multipliers rules obtained in the above sections, the Demyanov difference and the Rubinov defference for subdifferential and minus superdifferential of a quasidifferentiable function are utilized. As we know, the definitions of $\dot{-}$ and $\ddot{-}$ do not give, in general, the formulae and calculating methods of $U\dot{-}V$ and $U\ddot{-}V$. Whereas, if both U and V are convex hulls of a finite number of points, the sets $U\dot{-}V$ and $U\ddot{-}V$ could be expressed and calculated.

Let

$$U=\mathrm{co}\{u_i\mid i\in I\}\tag{5.1}$$

and

$$V = \text{co}\{v_j \mid j \in J\}, \tag{5.2}$$

where $u_i, v_j \in \mathbb{R}^n, I, J$ are finite index sets. Without loss of generality, we can suppose that $u_i \neq u_j$, $\forall i, j \in I$, $i \neq j$ and $v_i \neq v_j$, $\forall i, j \in J$, $i \neq j$. Given a pair of indices $i \in I$, $j \in J$, construct two systems of linear inequalities, denoted by (L_{ij}) and $(\overline{\mathrm{L}}_{ij})$, as follows:

$$(\mathrm{L}_{ij}) \qquad \begin{aligned} &(u_s - u_i)^T x < 0, \ \ \forall s \in I \setminus \{i\}, \\ &(v_t - v_j)^T x < 0, \ \ \forall t \in J \setminus \{j\}, \end{aligned}$$

and

$$(\overline{\mathrm{L}}_{ij}) \qquad \begin{aligned} &(u_s - u_i)^T x \leq 0, \ \ \forall s \in I \setminus \{i\}, \\ &(v_t - v_j)^T x \leq 0, \ \ \forall t \in J \setminus \{j\}, \end{aligned}$$

where $x \in \mathbb{R}^n$.

Obviously, each (L_{ij}) is a system with n variables and $\text{card}(I) + \text{card}(J) - 2$ strictly linear inequalities, and each $(\overline{\mathrm{L}}_{ij})$ is a system with n variables and $\text{card}(I) + \text{card}(J) - 2$ linear inequalities, where "card" denotes cardinality, as well as the coefficient matrices of (L_{ij}) and of $(\overline{\mathrm{L}}_{ij})$ coincide with each other.

Let U and V be defined as in (5.1) and (5.2). According to [9] and [10], the sets $U \dot{-} V$ and $U \ddot{-} V$ have the following forms

$$U \dot{-} V = \text{co}\{u_i - v_j \mid (\mathrm{L}_{ij}) \text{ is consistent}\} \tag{5.3}$$

and

$$U \ddot{-} V = \text{co}\{u_i - v_j \mid (\overline{\mathrm{L}}_{ij}) \text{ has non-zero solutions}\}. \tag{5.4}$$

Based on formulations (5.3) and (5.4), we can calculate the sets $U \dot{-} V$ and $U \ddot{-} V$. Thus, the optimality conditions obtained in above sections can be verified for some cases. Indeed, the quasidifferentials of the functions, which are generated from smooth functions by finitely many maximum, minimum and smooth composition operations, are pairs of polyhedron. In what follows, we take a class of quasidifferentiable functions as an example.

Consider a smooth composition of max-type functions of the form:

$$h(x) = g(\max_{j \in J_1} f_{1j}(x), \ldots, \max_{j \in J_m} f_{mj}(x)), \tag{5.5}$$

where $g : \mathbb{R}^m \to \mathbb{R}$ and each $f_{ij} : \mathbb{R}^n \to \mathbb{R}$ are continuously differentiable, J_i for $i = 1, \ldots, m$ are finite index sets. This class of functions is interesting and important, in some sense. Many publications dealt with the problem of minimizing it, see for instance [4, 6]. Given a point $x \in \mathbb{R}^n$, denote

$$f_i(x) = \max_{j \in J_i} f_{ij}(x), \ \ i = 1, \ldots, m,$$

$$J_i(x) = \{j \in J_i \mid f_{ij}(x) = f_i(x)\}, \ \ i = 1, \ldots, m,$$

$$I_+(x) = \{i \in \{1,\dots,m\} \mid \frac{\partial g(f_1,\dots,f_m)}{\partial f_i}|_{(f_1,\dots,f_m)=(f_1(x),\dots,f_m(x))} \geq 0\},$$

$$I_-(x) = \{i \in \{1,\dots,m\} \mid \frac{\partial g(f_1,\dots,f_m)}{\partial f_i}|_{(f_1,\dots,f_m)=(f_1(x),\dots,f_m(x))} < 0\}.$$

According to [5], h is quasidifferentiable on $\mathbb{R}^n$ and its quasidifferential can be expressed as follows:

$$\underline{\partial} h(x) = \mathrm{co}\{u \mid u = \sum_{i\in I_+(x)} \frac{\partial g(f_1,\dots,f_m)}{\partial f_i}|_{(f_1,\dots,f_m)=(f_1(x),\dots,f_m(x))} \nabla f_{ij_i}(x), \quad j_i \in J_i(x)\},$$

$$\overline{\partial} h(x) = \mathrm{co}\{v \mid v = \sum_{i\in I_-(x)} \frac{\partial g(f_1,\dots,f_m)}{\partial f_i}|_{(f_1,\dots,f_m)=(f_1(x),\dots,f_m(x))} \nabla f_{ij_i}(x), \quad j_i \in J_i(x)\}.$$

In the light of (5.3) and (5.4), we can calculate the sets $\underline{\partial} h(x) - (-\overline{\partial} h(x))$ and $\underline{\partial} h(x) - (-\overline{\partial} h(x))$. Besides, it is easy to see that $h \in \widetilde{M}(x)$.

The above discussion is applicable to any function, which is generated from smooth functions by finitely many maximum, minimum and smooth operations, since its quasidifferential is a pair of polyhedra and can be calculated according to quasidifferential calculus. For instance, let

$$H(x) = \max_{k\in K} g_k(\max_{j\in J_{1k}} f_{1jk}(x), \dots, \max_{j\in J_{mk}} f_{mjk}(x)), \tag{5.6}$$

where $g_k : \mathbb{R}^m \to \mathbb{R}$, $f_{ij} : \mathbb{R}^n \to \mathbb{R}$ are continuously differentiable, J_{ik}, K are finite index sets. Unlike h, H is generated by twice maximum operations.

By the definition in [14], $f : \mathbb{R}^n \to \mathbb{R}^m$ is said to be piecewise C^k on an open set $S \subset \mathbb{R}^n$, where k is a positive integer, if there exists a finite family of C^k functions $f_i : S \to \mathbb{R}^m$ for $i = 1,\dots,l$, called the C^k pieces of f, such that f is continuous on S and for every $x \in S$, $f(x) = f_i(x)$ for at least one index $i \in \{1,\dots,l\}$.

According to [2], any piecewise C^1 function f can be formulated as a minimax of finitely many smooth functions, i.e.,

$$f(x) = \min_{i\in I} \max_{j\in J} f_{ij}(x),$$

where $f_{ij} : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable, I and J are finite index sets. This shows that any piecewise C^1 function is contained in $M(x)$ and $\widetilde{M}(x)$, and its quasidifferential is a pair of polyhedra.

Acknowledgements. The author is grateful to the anonymous referee for his valuable suggestions and comments.

References

1. F. H. Clarke, *Optimization and Nonsmooth Analysis*, John Wiley & Sons, New York, 1983.
2. S. Dempe and D. Pallaschke, Quasidifferentiability of optimal solutions in parametric nonlinear optimization, *Optimization* **40** (1997) 1–24.
3. V. F. Demyanov, The relation between the Clarke subdifferential and a quasidifferential, Vestnik Leningrade University, **13** (1981) 183–189.
4. V. F. Demyanov, S. Gamidov, and T. I. Sivelina, An algorithm for minimizing a certain class of quasidifferentiable functions, *Mathematical Programming Study* **29** (1986) 74–84.
5. V. F. Demyanov and A. M. Rubinov, *Constructive Nonsmooth Analysis*, Peterlang, Frankfurt am Main, 1995.
6. G. Di Pillo, L. Grippo, and S. Lucidi, Smooth transformation of the generalized minimax problem, *J. Optimization and Applications* **95** (1997) 1–24.
7. K. Eppler and B. Luderer, The Lagrange principle and quasidifferential calculus, *Wiss. Z. d. TU Karl-Marx-Stadt* **29** (1987) 187–192.
8. Y. Gao, A weak Fritz John condition for quasidifferentiable functions, *Chinese J. Operations Research* **9** (1990) 43–44 (Chinese).
9. Y. Gao, Demyanov difference of two sets and optimality conditions of Lagrange multipliers type for constrained quasidifferentiable optimization, *J. Optimization Theory and Applications* **104** (2000) 377–394.
10. Y. Gao, Differences of two convex polyhedrons and relations among various subdifferentials, *OR Transactions* **4** (2000) 57–66.
11. J. B. Hiriart-Urruty and C. Lemarechal, *Convex Analysis and Minimization Algorithms*, Springer-Verlag, Berlin, 1993.
12. A. D. Ioffe, Necessary and sufficient conditions for local minimum. 1: A reduction theorem and first order conditions, *SIAM J. Control and Optimization* **17** (1979) 245–250.
13. B. Luderer, Directional derivative estimates for the optimal value function of a quasidifferentiable programming problem, *Mathematical Programming* **51** (1991) 333–348.
14. J. S. Pang and D. Ralph, Piecewise smoothness, local invertibility and parametric analysis of normal maps, *Mathematics of Operations Research* **21** (1996) 401–426.
15. L. N. Polyakova, On the minimization of a quasidifferentiable function subject to equality-type constraints, *Mathematical Programming Study* **29** (1986) 44–55.
16. R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, 1972.
17. A. M. Rubinov and I. S. Akhundov, Difference of compact sets in the sense of Demyanov and its application to nonsmooth analysis, *Optimization* **22** (1992) 179–188.
18. A. Shapiro, On optimality conditions in quasidifferentiable optimization, *SIAM J. Control and Optimization* **22** (1984) 610–617.
19. H. Yin and K. Zhang, The Fritz John type conditions in quasidifferentiable programming with equality constraints and their sufficiency, *Applied Mathematics - J. Chinese Universities* **13** (1998) 99–104 (Chinese).