International Journal of Pure and Applied Mathematics
Volume 20 No. 1 2005, 69-79

SECOND-ORDER OPTIMALITY CONDITIONS FOR A CONSTRAINED OPTIMIZATION

Wenhui Zhao[1], Yan Gao[2 §]
[1,2]School of Management
University of Shanghai for Science and Technology
516 Jungong Road, Shanghai, 200093, P.R. CHINA
[1]e-mail: wenhui_zhao@163.com
[2]e-mail: gaoyan1962@yahoo.com.cn

Abstract: In this paper, we study an optimization problem with equality and inequality constraints, where the corresponding functions have quasidifferentiable gradients. Two second-order necessary optimality conditions for this problem are proposed. The differences for convex compact sets proposed by Demyanov, and by Rubinov and Akhundov are used, respectively.

AMS Subject Classification: 49K52, 49J52, 90C26
Key Words: second-order optimality conditions, optimization, nonsmooth analysis, quasidifferential calculus

1. Introduction

Quasidifferential calculus, developed by Demyanov and Rubinov, is related closely to the classical directional derivative and the quasidifferentials can be computed easily for a large class of functions. So the calculus is viewed as an important tool in nonsmooth analysis and optimization and it is of main concern in the recent research in optimization. The first-order necessary optimality conditions for quasidifferentiable problem were proposed by Polyakova and Shapiro and were further developed by many authors, see for instance, [4], [5], [9], [10], [13], [14]. Problems which do not involve C^2 data also play impor-

Received: March 4, 2005
§Correspondence author

tant role in the theory of mathematical programming problems, too. Until now, most works done in this area were for the class of $C^{1,1}$ functions, that is, for the class of functions which are continuously differentiable and whose gradient mappings are locally Lipschitzian, see [7], [8]. Nevertheless, the problem with quasidifferentiable gradient has been studied less extensively. To our knowledge, only [15] dealt with quasidifferentiable gradient problem. A second-order necessary condition for unconstrained problem was proposed.

In the present paper, we intend to investigate the second-order optimality conditions for a constrained optimization problem with quasidifferentiable gradient via quasidifferential.

2. Preliminaries

In this section, we recall some related concepts that will be used later on.

As in [5], $F : \Re^n \to \Re^m$, a vector-value function, is called quasidifferentiable at $x \in \Re^n$ in the sense of Demyanov and Rubinov, if it is directionally differentiable at x and there exists a pair of convex compact sets $\underline{\partial} F(x), \overline{\partial} F(x) \subset \Re^{m\times n}$ such that

$$F'(x;d) = \max_{v\in\underline{\partial} F(x)} vd + \min_{w\in\overline{\partial} F(x)} wd, \quad \forall d \in \Re^n.$$

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. Let f_i be the i-th component of F with a quasidifferential $[\underline{\partial} f_i(x), \overline{\partial} f_i(x)]$. Evidently, F is quasidifferentiable if and only if all of its components are quasidifferentiable. The quasidifferential of F can be expressed by

$$[\underline{\partial} F(x), \overline{\partial} F(x)] = [\underline{\partial} f_1(x), \overline{\partial} f_1(x)] \times \cdots \times [\underline{\partial} f_m(x), \overline{\partial} f_m(x)].$$

The class of quasidifferentiable function contains convex, concave and differentiable functions, but also convex-concave, maximum and other functions.

Suppose f is a continuously differentiable function defined in $\Re^n$, and its gradient is a quasidifferentiable mapping. A quasidifferential of its gradient is called a second-order quasidifferential, which is denoted by

$$[\underline{\partial}^2 f(x), \overline{\partial}^2 f(x))] = [\underline{\partial} \nabla f(x), \overline{\partial} \nabla f(x)].$$

Assume that f is a continuously differentiable function defined in $\Re^n$, and its gradient is quasidifferentiable. As in [15], a generalized second-order directional derivative of f at x in the directions (d_1, d_2) is of the form:

$$f''(x; d_1, d_2) = \lim_{\lambda \downarrow 0} \frac{1}{\lambda}(\nabla f(x + \lambda d_2)^T d_1 - \nabla f(x)^T d_1).$$

The following propositions given by Xia and Zhao in [15] will be used to discuss necessary optimality conditions for the problem (P) that will be given in Section 3.

Proposition 2.1. *The function f defined in $\Re^n$ is continuously differentiable and its gradient is quasidifferentiable. Then for any $x \in \Re^n, d \neq 0, d \in \Re^n$ and $\lambda > 0$, one has the following second-order expansion (Taylor's expansion of second-order):*

$$f(x + \lambda d) = f(x) + \lambda f'(x; d) + \frac{\lambda^2}{2} d^T (V + W) d,$$

for some V and W, where

$$V \in \underline{\partial}^2 f(x + \zeta d)$$

and

$$W \in \overline{\partial}^2 f(x + \zeta d), \ \zeta \in (0, \lambda).$$

Proposition 2.2. *Under the assumptions given in Proposition 2.1, suppose x_0 is a local minimum point of f. Then for every $d \in \Re^n$ there exist $V \in \underline{\partial}^2 f(x_0)$ and $W \in \overline{\partial}^2 f(x_0)$ such that*

$$d^T (V + W) d \geq 0,$$

or

$$d^T H d \geq 0,$$

where $H \in \underline{\partial}^2 f(x_0) + \overline{\partial}^2 f(x_0)$.

In what follows, we review some of related concepts from [3], [1], [6], [11], [12].

Let S be a set in $\Re^n$. Given a point $x \in \Re^n$, put

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

and

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

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 $S \subset \Re^n$ be a convex compact set; the support function of the set S, denoted by $P_S(x)$, is defined by

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

It is true that P_S is a convex function on $\Re^n$ with

$$\partial P_S(x) = G_x(S). \tag{2.1}$$

Particularly, $\partial P_S(0) = S$, where "∂" denotes subdifferential in the sense of convex analysis. From (2.1), it follows that P_S is differentiable at x if and only if the right hand side of (2.1) is a singleton.

A set $T \subset \Re^n$ is called of full measure (with respect to $\Re^n$), if $\Re^n \setminus T$ is a set of measure zero. Let $U, V \subset \Re^n$ be convex compact sets and $T \subset \Re^n$ 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

$$U \dot{-} V = \operatorname{cl}\operatorname{co}\{\nabla P_U(x) - \nabla P_V(x) \mid x \in T\}, \tag{2.2}$$

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

Let U and V be convex compact sets, the difference $U \ddot{-} V$ of U and V, proposed by Rubinov and Akhundov [12], is defined by

$$U \ddot{-} V = \operatorname{cl}\operatorname{co} \bigcup_{x \neq 0} [G_x(U) - G_x(V)]. \tag{2.3}$$

The set $U \ddot{-} V$ is called the Rubinov difference of U and V.

3. Second-Order Necessary Optimality Conditions

We now consider the following constrained problems:

$$\begin{array}{lll} \text{(P)} & \text{minimize} & f(x) \\ & \text{subject to} & g_i(x) \le 0, i = 1, \dots, m, \\ & & h_j(x) = 0, j = 1, \dots, n, \end{array}$$

where $f, g_i, h_j, i = 1, \ldots, m, j = 1, \ldots, n$ are continuously differentiable functions in $\Re^n$, and their gradients are quasidifferentiable. Let x_0 be a local minimum point for (P). Moreover, assume the following constraint qualification is satisfied:

(H) $\nabla g_i(x_0), i \in I(x_0), \nabla h_j(x_0), j = 1, \ldots, n$ are linearly independent,

where $I(x_0) = \{i \mid g_i(x_0) = 0\}$. Then there exists a vector $(\lambda_1, \ldots, \lambda_m, \mu_1, \ldots, \mu_n) \in \Re^{m+n}$ such that the Karsh-Kuhn-Tucker optimality conditions are satisfied

$$\text{(KKT)} \quad \nabla f(x_0) + \sum_{i=1}^{m} \lambda_i \nabla g_i(x_0) + \sum_{j=1}^{n} \mu_j \nabla h_j(x_0) = 0,$$

$$\lambda_i \geq 0, \quad \lambda_i g_i(x_0) = 0, \quad i = 1, \ldots, m.$$

To get the second-order conditions, we associate with each multiplier $\lambda = (\lambda_1, \ldots, \lambda_m)$, a set $G(\lambda)$ is defined as

$$G(\lambda) = \{x \mid g_i(x) = 0 \;\text{ when }\; \lambda_i > 0,\; g_i \leq 0 \;\text{ when }\; \lambda_i = 0,\; h_j(x) = 0,\; j = 1, \ldots, n\},$$

and denote the cone of feasible directions to $G(\lambda)$ at x_0 by

$$F(G(\lambda), x_0) = \{d \mid \exists \delta > 0, \text{s.t.} \forall \theta \in (0, \delta], x = x_0 + \theta d \in G(\lambda)\}.$$

If we assume that (H) is satisfied, then there exists at least one Karsh-Kuhn-Tucker multiplier (λ, μ). The usual Lagrangian function of f at x, denoted by $L(x, \lambda, \mu)$, is of the form:

$$L(x, \lambda, \mu) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) + \sum_{j=1}^{n} \mu_j h_j(x), \tag{3.1}$$

and denoted by $L''(x, \lambda, \mu; d)$ the second-order directional derivative of $L(\cdot; \lambda, \mu)$ at x. Then the second-order necessary conditions can be expressed as the following theorem.

Theorem 3.1. *Let x_0 be a local minimum point of* (P) *and let* (H) *hold. Then for each Karsh-Kuhn-Tucker multiplier vector (λ, μ) satisfying (KKT) at x_0, for each $d \in F(G(\lambda), x_0)$, we have $L''(x_0, \lambda, \mu; d) \geq 0$.*

Proof. Let λ, μ and d be fixed. By definition of $F(G(\lambda), x_0)$, there exists a $\delta > 0$ for any $\forall\theta \in (0, \delta]$, such that $x_0 + \theta d \in G(\lambda)$. From (3.1), one has that

$$L(x_0 + \theta d, \lambda, \mu) = f(x_0 + \theta d) + \sum_{i=1}^{m} \lambda_i g_i(x_0 + \theta d) + \sum_{j=1}^{n} \mu_j h_j(x_0 + \theta d).$$

By virtue of the assumptions and the chain rules on quasidifferential calculus [2], [5], $L(x, \lambda, \mu)$ is quasidifferentiable. On the other hand, by Proposition 2.1, it follows that

$$L(x_0 + \theta d, \lambda, \mu) = L(x_0, \lambda, \mu) + \theta L'(x_0, \lambda, \mu; d) + \frac{\theta^2}{2} d^T (V + W)d,\ d \in F(G(\lambda), x_0),$$

for some V and W, where

$$V \in \underline{\partial}^2 L(x_0 + \zeta d, \lambda, \mu) \text{ and } W \in \overline{\partial}^2 L(x_0 + \zeta d, \lambda, \mu), \quad \zeta \in (0, \theta).$$

Since $\theta \neq 0$, we have

$$\frac{L(x_0 + \theta d, \lambda, \mu) - L(x_0, \lambda, \mu)}{\theta} = L'(x_0, \lambda, \mu; d) + \frac{\theta}{2} d^T (V + W)d. \tag{3.2}$$

From the Mean Value Theorem and (3.2), we may conclude that there exists a $t = t(\theta)$ such that

$$L'(x_0 + td, \lambda, \mu; d) - L'(x_0, \lambda, \mu; d) = \frac{\theta}{2} d^T (V + W)d, \quad d \in F(G(\lambda), x_0), \tag{3.3}$$

where

$$V \in \underline{\partial}^2 L(x_0 + \zeta d, \lambda, \mu) \text{ and } W \in \overline{\partial}^2 L(x_0 + \zeta d, \lambda, \mu), \quad \zeta \in (0, \theta).$$

Dividing (3.3) by t and then taking limits as $\theta \downarrow 0$, we obtain

$$\lim_{\theta \downarrow 0} \frac{L'(x + td, \lambda, \mu; d) - L'(x_0, \lambda, \mu; d)}{t} = \lim_{\theta \downarrow 0} d^T (\frac{\theta}{2t}(V + W))d,\ d \in F(G(\lambda), x_0).$$

The fact that $L''(x_0, \lambda, \mu; d, d)$ exists and $t \to 0^+$ when $\theta \downarrow 0$ leads to the following equalities hold:

$$\begin{aligned}\lim_{\theta\downarrow 0} \frac{L'(x_0 + td, \lambda, \mu; d) - L'(x_0, \lambda, \mu; d)}{t} \\ &= \lim_{\theta\downarrow 0} \frac{L'(x_0 + \theta d, \lambda, \mu; d) - L'(x_0, \lambda, \mu; d)}{\theta} \\ &= L''(x_0, \lambda, \mu; d, d) = L''(x_0, \lambda, \mu; d), \quad d \in F(G(\lambda), x_0).\end{aligned}$$

Hence,

$$\frac{\theta^2}{2} d^T (V + W) d = \theta t L''(x_0, \lambda, \mu; d) + \theta t(\epsilon), \quad d \in F(G(\lambda), x_0),$$

where $\epsilon \to 0$, as $\theta \to 0$ and $t = t(\theta) \to 0$. Thus, for any $\theta > 0$ small enough, there exists a $t = t(\theta)$ such that

$$\begin{aligned}L(x_0 + \theta d, \lambda, \mu) = L(x_0, \lambda, \mu) + \theta \nabla L(x_0, \lambda, \mu; d) \\ + \theta t L''(x_0, \lambda, \mu; d) + 0(\theta t), \quad d \in F(G(\lambda), x_0). \quad (3.4)\end{aligned}$$

Since x_0 is a minimum point to (P), the following relations are satisfied

$$\begin{cases} \nabla L(x_0, \lambda, \mu) = 0, \\ L(x_0, \lambda, \mu) = 0, \\ f(x_0 + \theta d) \geq f(x_0). \end{cases} \quad (3.5)$$

Then, gathering (3.4), (3.5) and the definition of $L(x, \lambda, \mu)$ gives:

$$L''(x_0, \lambda, \mu; d) \geq 0 \text{ for any } d \in F(G(\lambda), x_0).$$

This completes the proof of the theorem. □

To get another optimality condition of (P), we give the following lemma.

Lemma 3.1. *Suppose f is a continuously differentiable function defined in $\Re^n$ and its gradient is quasidifferentiable. Then*

$$f''(x; d, d) = \Big(\max_{v \in \underline{\partial}^2 f(x)} v^T d + \min_{w \in \overline{\partial}^2 f(x)} w^T d\Big) d.$$

Proof. By deducing, we have

$$\begin{aligned} f''(x;d,d) &= \lim_{\lambda\downarrow 0}\frac{1}{\lambda}(\nabla f(x+\lambda d)^T d - \nabla f(x)^T d) \\ &= \lim_{\lambda\downarrow 0}\frac{1}{\lambda}(\nabla f(x+\lambda d) - \nabla f(x))^T d = \lim_{\lambda\downarrow 0}\sum_{i=1}^{n}\frac{1}{\lambda}(\nabla f(x+\lambda d) - \nabla f(x))_i d_i \\ &= \sum_{i=1}^{n} d_i(\max_{v_i\in\underline{\partial}(\nabla f(x))_i} v_i^T d + \min_{w_i\in\overline{\partial}(\nabla f(x))_i} w_i^T d) = \max_{v\in\underline{\partial}^2 f(x)}(v^T d)^T d \\ &\qquad + \min_{w\in\overline{\partial}^2 f(x)}(w^T d)^T d. \end{aligned}$$

Thus, the conclusion is obtained. □

We proceed now to present another necessary optimality condition of (P) .

Theorem 3.2. *Let* x_0 *be a local minimum point of* (P) *and let* (H) *hold. For each Karsh-Kuhn-Tucker multiplier vector* (λ,μ) *satisfying* (KKT) *at* x_0 *and each* $d \in F(G(\lambda), x_0)$, *there exists at least one* $\zeta \in \underline{\partial}^2 L(x_0) \,\ddot{-}\, (-\overline{\partial}^2 L(x_0))$ *such that* $d^T\zeta d \geq 0$.

Proof. Given a $d \in F(G(\lambda), x_0)$, denote

$$G_d(\underline{\partial}^2 L(x_0)) = \{v \in \underline{\partial}^2 L(x_0) \mid vd = \max_{v\in\underline{\partial}^2 L(x_0)} vd\}$$

and

$$\widetilde{G}_d(\overline{\partial}^2 L(x_0)) = \{w \in \overline{\partial}^2 L(x_0) \mid wd = \min_{w\in\overline{\partial}^2 L(x_0)} wd\}.$$

We set $U = \underline{\partial}^2 L(x_0)$ and $V = -\overline{\partial}^2 L(x_0)$ in (2.3). Then,

$$\underline{\partial}^2 L(x_0) \,\ddot{-}\, (-\overline{\partial}^2 L(x_0)) = \operatorname{cl}\operatorname{co}\bigcup_{x\neq 0}[G_d(\underline{\partial}^2 L(x_0)) + \widetilde{G}_d(\overline{\partial}^2 L(x_0))].$$

Hence, there exist $v \in \underline{\partial}^2 L(x_0)$ and $w \in \overline{\partial}^2 L(x_0)$ such that

$$(\nabla L(x_0;d))' = \max_{v\in\underline{\partial}^2 L(x_0)} vd + \min_{w\in\overline{\partial}^2 L(x_0)} wd = (v+w)d,$$

$$v + w \in \underline{\partial}^2 L(x_0) \,\ddot{-}\, (-\overline{\partial}^2 L(x_0)). \tag{3.6}$$

From Lemma 3.1 and (3.6), given a $d \in F(G(\lambda), x_0)$ we have

$$(\nabla L(x_0; d))'d = L''(x_0; d, d) = d^T(v + w)d \geq 0,$$

$$v + w \in \underline{\partial}^2 L(x_0) \mathbin{\ddot{-}} (-\overline{\partial}^2 L(x_0)).$$

It is easy to see that

$$\max_{v \in \underline{\partial}^2 L(x_0),\, w \in \overline{\partial}^2 L(x_0)} d^T(v + w)d \geq 0.$$

In other words, for each $d \in F(G(\lambda), x_0)$, there exists at least one $\zeta \in \underline{\partial}^2 L(x_0) \mathbin{\ddot{-}} (-\overline{\partial}^2 L(x_0))$ such that

$$L''(x_0; d, d) = d^T \zeta d \geq 0.$$

We thus complete the proof of the theorem. □

Now, we have obtained two second-order necessary optimality conditions for (P). Theorem 3.2 improves the necessary condition given in [5]. If f is a continuously differentiable function defined in $\Re^n$ and its gradient is quasidifferentiable, then the relation $\underline{\partial}^2 f(x) \mathbin{\ddot{-}} (-\overline{\partial}^2 f(x)) \subset \underline{\partial}^2 f(x) + \overline{\partial}^2 f(x)$ always holds. Therefore, Theorem 3.2 implies Proposition 2.2. That is to say the former is sharper than the latter.

Proposition 3.1. *Let x_0 be a local minimum point of* (P) *and let* (H) *hold. Suppose there exists*

$$T' = \{d \in \Re^n \mid G_d(\underline{\partial}^2 L(x_0)) \textit{ and } \widetilde{G}_d(\overline{\partial}^2 L(x_0)) \textit{ are singletons}\},$$

then there exists at least one $\zeta \in \underline{\partial}^2 L(x_0) \mathbin{\dot{-}} (-\overline{\partial}^2 L(x_0))$ such that $d^T \zeta d \geq 0, d \in T' \cap F(G(\lambda), x_0)$.

Proof. For any $d \in T'$, we denote

$$G_d(\underline{\partial}^2 L(x_0)) = \{v(d)\} \text{ and } \widetilde{G}_d(\overline{\partial}^2 L(x_0)) = \{w(d)\}.$$

In fact, if $G_d(\underline{\partial}^2 L(x_0))$ and $\widetilde{G}_d(\overline{\partial}^2 L(x_0))$ are singletons, T' is the set where the convex function $\max_{v \in \underline{\partial}^2 L(x_0)} vd$ and the concave function $\min_{w \in \overline{\partial}^2 L(x_0)} wd$ are differentiable. We set $U = \underline{\partial}^2 L(x_0)$, $V = -\overline{\partial}^2 L(x_0)$ and $T = T'$ in (2.2). It is easy to see that

$$v(d) + w(d) \in \underline{\partial}^2 L(x_0) \mathbin{\dot{-}} (-\overline{\partial}^2 L(x_0)).$$

This yields

$$(\nabla L(x_0;d))' = (v(d)+w(d))d, \quad d \in T' \cap F(G(\lambda), x_0). \tag{3.7}$$

According to Lemma 3.1 and (3.7), for each $d \in T' \cap F(G(\lambda), x_0)$, there exists at least one $\zeta \in \underline{\partial}^2 L(x_0) \dot{-} (\overline{\partial}^2 L(x_0))$ such that

$$L''(x_0;d,d) = (\nabla L(x_0;d))'d = d^T \zeta d \geq 0. \qquad \square$$

Acknowledgements

This work was supported by Shanghai Education Committee for key project (04EA01).

References

[1] V.F. Demyanov, The relation between the Clarke subdifferential and a quasidifferential, *Vestnik Leningrade University*, **13** (1981), 183-189.

[2] V.F Demyanov, A.M. Rubinov, *Quasidifferential Calculus, Optimization Software*, Inc., New York (1986).

[3] V.F. Demyanov, A.M. Rubinov, *Constructive Nonsmooth Analysis*, Peterlang, Frankfurt am Main (1995).

[4] Y. Gao, Demyanov difference of two sets and optimality conditions of Lagrange multipliers type for constrained quasidifferentiable optimization, *Journal of Optimization Theory and Applications*, **104** (2000), 377-394.

[5] Y. Gao, Representation of the Clarke generalized Jacobian via the quasidifferential, *Journal of Optimization Theory and Applications*, **123** (2004), 519-532.

[6] J.B. Hiriart-Urruty, C. Lemarechal, *Convex Analysis and Minimization Algorithms*, Springer-Verlag, Berlin (1993).

[7] J.B. Hiriart-Urruty, J.J. Strodiot, V. Hien Nguyen, Generalized Hessian matrix and second order optimality conditions for problems with $C^{1,1}$ data, *Applied Mathematics and Optimization*, **11** (1984), 43-56.

[8] D. Klatte, Upper Lipschitz behavior of solutions to perturbed $C^{1,1}$ optimization problems, *Optimization*, **19** (1988), 169-179.

[9] B. Luderer, Directional derivative estimates for the optimal value function of a quasidifferentiable programming problem, *Mathematical Programming*, **51** (1991), 333-348.

[10] L.N. Polyakova, On the minimization of a quasidifferentiable function subject to equality-type constraints, *Mathematical Programming Study*, **29** (1986), 44-55.

[11] R.T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton (1972).

[12] A.M. Rubinov, I.S. Akhundov, Difference of compact sets in the sense of Demyanov and its application to nonsmooth analysis, *Optimization*, **22** (1992), 179-188.

[13] A. Shapiro, On optimality conditions in quasidifferentiable optimization, *SIAM J. Control and Optimization*, **22** (1984), 610-617.

[14] A. Uderzo, Quasi-multipliers rules for quasidifferentiable extremum problems, *Optimization*, **51** (2002), 761-795.

[15] Z.Q. Xia, C.J. Zhao, Second-order expansions for a class of quasidifferentiable functions, *Demonstratio Mathematica*, **12** (1989), 365-389.