Journal of Mathematical Research and Exposition
Vol.11 No.4. Nov.1991

An Algorithm for Minimizing a Class of Quasidifferentiable Functions*

Zunquan Xia

(Dalian University of Technology)

Abstract

Some results concerning approximating quasidifferentials by means of polyhedrons and an algorithm for minimizing a class of quasidifferentiable functions are given in this paper. The main idea here is following the one in [8].

1. Introduction

For minimizing a general quasidifferentiable function f in the sense of Demyanov and Rubinov [3], [4], [6], the necessary condition for a point x to be a (local) minimizer reads that

$$-\bar{\partial} f(x) \subset \underline{\partial} f(x) \tag{1.1}$$

due to [7], or

$$\operatorname*{man}_{w \in \partial f(x)} \min_{v \in \partial f(x)} \| v + w \| = 0 \tag{1.2}$$

[4]. If the condition (1.1) or (1.2) is not satisfied at x, then the direction

$$d = -(v_o + w_o) / \| v_o + w_o \| \tag{1.3}$$

is a steepest descent direction, where

$$\| v_o + w_o \| = \max_{w \in \partial f(x)} \min_{v \in \partial f(x)} \| v + w \| > 0 \tag{1.4}$$

[4]. For a general quasidifferentiable function it is not easy to construct numerical methods. This is not only because it is very difficult to calculate (1.4) or to check (1.2), i.e., (1.1), but also because the uniqueness in the sense of the equivalence class of quasidifferentials of a quasidifferentiable function at a point can not be used in practice. For this reason we only study a certain class of quasidifferentiable functions in this paper.

Definition 1.1. Let f be quasidifferentiable in $\mathbf{R}^n$. A pair of nonemty convex compact sets $[\partial_* f(x), \partial^* f(x)]$ is said to be the quasidifferential kernel of f at x if the following conditions are satisfied

$$\begin{aligned} f'(x; d) &= \underline{f}'(x; d) - \bar{f}'(x; d) \\ &= \delta^*(d \mid \partial_* f(x)) - \delta^*(d \mid \partial^* f(x)), \quad d \in \mathbf{R}^n \end{aligned} \tag{1.5}$$

and

Received Dec.5, 1990. In memory of: Via Pogliani 4-6, 20037 Paderno Dugnano Italia; Danhausergasse 10/13. A1040 Wien, Österrich.

$$\partial_* f(x) = \bigcap_{[\underline{\partial} f(x), \bar{\partial} f(x)] \in \mathscr{D} f(x)} [\underline{\partial} f(x) + \bar{\partial} f(x)]$$
$$\partial^* f(x) = \bigcap_{[\underline{\partial} f(x), \bar{\partial} f(x)] \in \mathscr{D} f(x)} [\bar{\partial} f(x) - \bar{\partial} f(x)] \quad (1.6)$$

[9].

Let $Q(\mathbf{R}^n)$ denote such a subclass of quasidifferentiable functions that every function in $Q(\mathbf{R}^n)$ satisfies the conditions (1.5) and (1.6). Correspondingly, $\partial_* f(x)$ and $\partial^* f(x)$ are called sub- and super-quasidifferential kernel, respectively. Clearly, convex and concave, D.C. functions in the sense of Hiriart-Urruty belong to $Q(\mathbf{R}^n)$.

The necessary condition for $f \in Q(\mathbf{R}^n)$ to take a minimum at x is

$$\partial^* f(x) \subset \partial_* f(x). \quad (1.7)$$

This condition is very difficult to be verified. We will pay our attention to a special case where

$$\partial^* f(x) = \mathrm{co} S \quad (1.8)$$

and S is a finite set of points.

The ABS methods by Abaffy, Broyden and Spedicato, [1,2], can be used to construct an algorithm of polyhedron approximating subdifferentials of finite convex functions [2]. In the convex case in which

$$S = \{0\},$$

and hence

$$\partial^* f(x) = \{0\}.$$

The algorithm SAPA given in [8] is actually a descent algorithm for minimizing a finite convex function. This approach can be extended to the case in which $\partial^* f(x)$ is of the form (1.8). We confine our goal to minimizing such a function being of the form

$$f(x) := f_0(x) - \max_{i \in I} f_i(x) \quad (1.9)$$

where f_i is convex, $f_i \in C^1$, $i \in I$, and I is a finite index set.

In Sec.2 a general representation for polyhedral approximations to quasidifferentials is given and the principle of an algorithm for minimizing a function being of the form (1.9) is given in Sec.3.

2. Polyhedral approximations

2.1 General formulation

Consider the following problem

$$\min f(x) = f_0(x) - \max_{i \in I} f_i(x) \qquad x \in \mathbf{R}^n \quad (2.1)$$

where $f_0: \mathbf{R}^n \to \mathbf{R}^1$ is convex, not necessarily continuously differentiable, $f_i: \mathbf{R}^n \to \mathbf{R}^1 \in C^1$, not necessarily convex, $I := \{1, 2, \cdots, m\}$ is a finite index set.

The function in (2.1) belongs to $Q(\mathbf{R}^n)$.

Let $\varphi_d(t):=f(x+td)$, $d\in\mathbf{R}^n$, $t\in\mathbf{R}^1_+$. Since

$$\varphi_d'(t;k)=\lim_{\lambda\downarrow 0}(\varphi_d(t+\lambda k)-\varphi_d(t))/\lambda$$
$$=\lim_{\lambda\downarrow 0} f(x+(t+\lambda k)d-f(x+td))/\lambda$$
$$=\delta^*(kd|\partial f(x+td))-\delta^*(kd|-\partial f(x+td)),$$

one has that

$$\varphi_d'(0;k)=f'(x;kd)$$
$$=\max_{m\in\langle\partial f(x),d\rangle} km+\min_{n\in\langle\partial f(x),d\rangle} kn.$$

Hence, one obtains the general expression of a quasidifferential of φ_d at the origin,

$$\underline{\partial}\varphi_d(0)=\langle\underline{\partial}f(x),d\rangle,\ \bar{\partial}\varphi_d(0)=\langle\bar{\partial}f(x),\ d\rangle.$$

It follows that

$$\underline{\partial}\varphi_d(0)+\bar{\partial}\varphi_d(0)=\langle\underline{\partial}f(x)+\bar{\partial}f(x),d\rangle,$$
$$\bar{\partial}\varphi_d(0)-\bar{\partial}\varphi_d(0)=\langle\bar{\partial}f(x)-\bar{\partial}f(x),d\rangle.$$

According to the results due to [5], one has that

$$\partial_*\varphi_d(0)=\bigcap_{\mathscr{D}f(x)}\langle\underline{\partial}f(x)+\bar{\partial}f(x),d\rangle$$
$$\partial^*\varphi_d(0)=\bigcap_{\mathscr{D}f(x)}\langle\bar{\partial}f(x)-\bar{\partial}f(x),d\rangle. \tag{2.2}$$

Now for any vector u such that

$$u\in\partial_*f(x)=\bigcap_{\mathscr{D}f(x)}(\underline{\partial}f(x)+\bar{\partial}f(x))$$

and any vector w such that

$$w\in\partial^*f(x)=\bigcap_{\mathscr{D}f(x)}(\bar{\partial}f(x)-\bar{\partial}f(x)),$$

we obtain

$$\langle u,d\rangle\in\partial_*\varphi_d(o),\ \ \langle w,d\rangle\in\partial^*\varphi_d(0)$$

and

$$\langle\partial_*f(x),d\rangle\subset\partial_*\varphi_d(0),\ \ \langle\partial^*f(x),d\rangle\subset\partial^*\varphi_d(0).$$

Let

$$[\underline{a},\underline{\beta}]:=\bigcap_{\mathscr{D}f(x)}\langle\underline{\partial}f(x)+\bar{\partial}f(x),d\rangle, \tag{2.3}$$
$$[\bar{a},\bar{\beta}]:=\bigcap_{\mathscr{D}f(x)}\langle\bar{\partial}f(x)-\bar{\partial}f(x),d\rangle. \tag{2.4}$$

Then we have the following

Lemma 2.1 The scalars $\underline{a},\underline{\beta},\bar{a},\bar{\beta}$ defined above can be evaluated by the following formulae

$$\underline{a}=\min\{f'(x;d),-f'(x;-d)\}=\min\partial_*\varphi_d(0), \tag{2.5}$$
$$\underline{\beta}=\max\{f'(x;d),-f'(x;-d)\}=\max\partial^*\varphi_d(0), \tag{2.6}$$
$$\bar{a}=\min\{o,f'(x;d)+f'(x;-d)\}=\min\partial^*\varphi_d(0), \tag{2.7}$$

$$\bar{\beta} = -\underline{\alpha} = \max \partial^* \varphi_d(0).$$

Proof Since $\partial_* \varphi_d(o)$ and $\partial^* \varphi_d(0)$ are nonempty and finite intervals, one has that

$$\underline{\alpha} = \min\{v | v \in \partial_* \varphi_d(0)\} = \max_{\mathscr{D}f(x)} \min_{u \in \underline{\partial} f(x) + \bar{\partial} f(x)} \langle u, d \rangle$$

$$\underline{\beta} = \max\{v | v \in \partial_* \varphi_d(0)\} = \min_{\mathscr{D}f(x)} \max_{u \in \underline{\partial} f(x) + \bar{\partial} f(x)} \langle u, d \rangle$$

$$\bar{\alpha} = \min\{v | v \in \partial^* \varphi_d(0)\} = \max_{\mathscr{D}f(x)} \min_{u \in \underline{\partial} f(x) + \bar{\partial} f(x)} \langle u, d \rangle$$

$$\bar{\beta} = -\bar{\alpha}.$$

According to the definitions in Sec.1 and the results due to [5], the preceding expressions on the right hand side of equalities given above are equl to $\quad -\underline{f}'(x; -d), \underline{f}'(x; d), -\bar{f}'(x; -d), \bar{f}(x; d)$, respectively, therefore, one has that

$$\underline{\alpha} = -\underline{f}'(x; -d) = \min\{f'(x; d), -f'(x; -d)\} = \min \partial_* \varphi_d(0)$$

$$\underline{\beta} = \underline{f}'(x; d) = \max\{f'(x; d), -f'(x; -d)\} = \max \partial_* \varphi_d(0)$$

$$\bar{\alpha} = -\bar{f}'(x; d) = -\max(0, -f'(x; d) - f'(x; -d)\} = \min \partial^* \varphi_d(0)$$

$$\bar{\beta} = -\bar{\alpha} = \max \partial^* \varphi_d(0).$$

The lemma has been proved. ■

Let $\mathrm{bd} B_1(0) := \{u \in \mathbf{R}^n \mid \|u\| = 1\}$. Then $\partial_* \varphi_d(0)$ and $\partial^* \varphi_d(0)$ can be regarded as the projections of $\partial_* f(x)$ and $\partial^* f(x)$ onto the direction d, respectively. Taking $\{d_i\}_1^m \subset \mathrm{bd} B_1(0)$, we have that

$$\partial_* f(x) \subset \Delta_i := [\underline{\alpha}_i, \underline{\beta}_i] d_i + \{v | v^T d = 0\}$$

and

$$\partial^* f(x) \subset \Delta^i := [\bar{\alpha}_i, \bar{\beta}_i] d_i + \{v | v^T d = 0\},$$

in other words, one has that

$$\partial_* f(x) \subset \bigcap_{i=1}^m \Delta_i, \quad \partial^* f(x) \subset \bigcap_{i=1}^m \Delta^i.$$

Since $\partial_* f(x)$ and $\partial^* f(x)$ are compact convex, it follows that when d runs over $\mathrm{bd} B_1(0)$ the intersections above comprise just the kernel of quasidifferential of f at x, i.e.,

$$\partial_* f(x) = \bigcap_{d \in \mathrm{bd} B_1(0)} \Delta_d \tag{2.9}$$

$$\partial^* f(x) = \bigcap_{d \in \mathrm{bd} B_1(0)} \Delta^d. \tag{2.10}$$

The equalities (2.9) and (2.10) imply that the $\partial_* f(x)$ and $\partial^* f(x)$ can be approximated by means of polyhedrons determined by a set of unit vectoss.

2.2 Determination of descent directions

We denote by $P_i(x_k)$ the ith polyhedral approximations to $\partial_* f(x_k)$ and $\partial^* f(x_k)$, respectively, at the kth iteration.

Lemma 2.2. If $P^i(x_k) \subset P_i(x_k)$, then $\partial^* f(x_k) \subset \partial_* f(x_k)$.

Proof By contradiction, assume that

$$\partial^* f(x_k) \subset \partial_* f(x_k).$$

Since

$$\mathrm{Proj}_{P_i(x_k)} d_l = \mathrm{Proj}_{\partial^* f(x_k)} d_l = \mathrm{Proj}_{\Delta_l} d_l$$

and

$$\mathrm{Proj}_{P^i(x_k)} d_l = \mathrm{Proj}_{\partial^* f(x_k)} d_l = \mathrm{Proj}_{\Delta^l} d_l,$$

one has that for any $l \in \{1, \cdots, m\}$,

$$\delta^*(d_l | P_i(x_k)) = \delta^*(d_l | \partial_* f(x_k)) = \delta^*(d_l | \Delta_l), \quad (2.11)$$

$$\delta^*(-d_l | P_i(x_k)) = \delta^*(-d_l | \partial_* f(x_k)) = \delta^*(-d_l | \Delta_l). \quad (2.12)$$

For $\delta^*(d_l | P^i(x_k))$ and $\delta^*(-d_l | P^i(x_k))$ we have the equalities similar to (2.11) and (2.12). By the assumption of contradiction we have

$$\delta^*(d_l | \partial_* f(x_k)) \geqq \delta^*(d_l | \partial^* f(x_k)), \quad (2.13)$$

$$\delta^*(-d_l | \partial_* f(x_k)) \geqq \delta^*(-d_l | \partial^* f(x_k)). \quad (2.14)$$

It follows from (2.11) to (2.14) that

$$\delta^*(d_l | P_i(x_k)) \geqq \delta^*(d_l | P^i(x_k)), \quad (2.15)$$

$$\delta^*(-d_l | P_i(x_k)) \geqq \delta^*(-d_l | P^i(x_k)), \quad (2.16)$$

$$\forall l \in \{1. \cdots, m\}.$$

According to the structures of $P_i(x_k)$ and $P^i(x_k)$ and to (2.16) − (2.16) and to the assumption of contradiction, we have

$$P_i(x_k) = \bigcap_1^m \Delta_l = P^i(x_k). \quad (2.17)$$

The inclusion relationship (2.17) contradicts the hypothesis of the lemma

$$P^i(x_k) \subset P_i(x_k).$$

Hence $\partial^* f(x_k) \subset \partial_* f(x_k)$. The demonstration is completed. ∎

Let

$$G_i := \{u \in \mathbf{R}^n | u = d_l \text{ or } -d_l, \|d_l\| = 1, l \in \{1, \cdots, m_i\}\},$$

where G_i is used to construct the ith approximation.

Theorem 2.3. If $P^{i-1}(x_k) \subset P_{i-1}(x_k)$, but

$$P^i(x_k) \not\subset P_i(x_k),$$

then there exists at least one $d \in G_i$ such that

$$f'(x_k; d) < 0, \quad (2.19)$$

i.e., d is a descent direction of f at x_k. If x_k is a minimizer to (2.1), then one has

$$P^i(x_k) \subset P_i(x_k).$$

Proof If the condition (2.18) holds, then according to Lem. 2.2 one has

$$\partial^* f(x_k) \not\subset \partial_* f(x_k).$$

This implies that x_k is not a (local) minimizer. The second assertion can be issued immediately from this.

According to (2.18) and the structures of $P_i(x_k)$ and $P^0(x_k)$, there exists at least one $d \in G_i$ such that

$$\delta^*(d\,|P^i(x_k)) > \delta^*(d\,|P_i(x_k)). \qquad (2.20)$$

Since

$$\delta^*(d\,|P^i(x_k)) = \delta^*(d\,|\partial^* f(x_k)),\ \delta^*(d\,|P_i(x_k)) = \delta^*(d\,|\partial_* f(x_k))$$

due to (2.11) and (2.12), one has

$$f'(x_k; d) = \delta^*(d\,|\partial_* f(x_k)) - \delta^*(d\,|\partial^* f(x_k))$$
$$= \delta^*(d\,|P_i(x_k)) - \delta^*(d\,|P^i(x_k)).$$

It follows from (2.20) that

$$f'(x_k; d) < 0. \qquad (2.21)$$

The inequality (2.21) implies that d is a descent direction of f at x_k. The demonstration is completed. ■

Corollary If $\partial^* f(x_k) = \{0\}$, $m_i = l$, and (2.18) holds, then $d = d_l$ or $-d_l$ is a descent direction. ■

Noticing that it can not be conjectured that

$$\partial^* f(x_k) \subset \partial_* f(x_k).$$

We will concentrate our attention in the next section to the case where

$$P^i(x_k) \subset P_i(x_k). \qquad (2.22)$$

3. An algorithm for the problem (2.1)

An algorithm, named by SAPA, for minimizing a finite convex function was given in [8]. To begin with, we go back to this algorithm.

3.1 The algorithm SAPA for minimizing a finite convex function.

Algorithm SAPA

Let the initial data x_k, $\varepsilon_D (>0)$, $i := 0$ be given.

Step 1. Execute Algorithm IPA.

The following steps are carried out in multiprocessors.

Step 2. Compute

$\mathscr{K}(v_i)$ by Algorithm NVS.

In practical computation, we could define

$$s := \sum_{v \in \mathscr{K}(v_i)} \|v - v_i\|$$

If $s \leq \varepsilon_D$, then set $g := v_i$ and stop.

Step 3. Execute Algorithm CTO.

Step 4. Execute Algorithm DDD.

Step 5. $i := i + 1$ and go to Step 2.

//SAPA//

3.2 An algorithm for the problem (2.1)

For simplicity the notation $[\underline{\partial} f(x_k), \bar{\partial} f(x_k)]$ is used instead of $[\partial_* f(x_k), \partial^* f(x_k)]$.

Lemma 3.1. If $f \in Q(\mathbf{R}^n)$ and $w \in \bar{\partial} f(x_k)$, but $w \notin \underline{\partial} f(x_k)$, then the direction

$$d = w - \mathrm{proj}_{\partial f(x_k)} w = w - Nr_w \partial f(x_k) = -Nr(\partial f(x_k) - w)$$

is a descent direction.

Proof Since $O \in \bar{\partial} f(x_k)$, one has

$$\bar{f}'(x_k; h) \geqslant 0, \quad \forall h \in \mathbf{R}^n.$$

Since

$$\underline{f}'(x_k; d_0) = \mathrm{Proj}_{\underline{\partial} f(x_k)} d_o = \mathrm{Proj}_{d_o} u < 0,$$

where $d_0 = d / \|d\|$, $u = \mathrm{Proj}_{\underline{\partial} f(x_k)} w \in \underline{\partial} f(x_k)$, it follows that

$$f'(x_k; d) = \underline{f}'(x_k; d) - \bar{f}'(x_k; d) < 0$$

due to [9].

Let $R(x_k) := \{ i \in I \mid f_i(x_k) = \max_{i \in I} f_i(x_k) \}$. Then we have that

$$\underline{\partial} f(x_k) = \partial f_0(x_k) - B, \quad \bar{\partial} f(x_k) = B - B,$$

where

$$B := \mathrm{co}\{ \nabla f_i(x_k) \mid i \in \mathrm{R}(x_k) \},$$

[10]. Clearly, $\bar{\partial} f(x_k)$ is a polyhedron formed by taking the convex hull of a finite number of extreme points. Let $\mathrm{Ex}\bar{\partial} f(x_k)$ denote the set of all extreme points of $\partial f(x_k)$. ■

Algorithm AQD

$x_1, I := \{1, \cdots, N\}$, $i := 1, k := 1$.

Step 1. Compute

$$R(x_k)$$

$$\mathrm{Ex}\bar{\partial} f(x_k).$$

Step 2. For $w_i \in \mathrm{Ex}\bar{\partial} f(x_k\ (x_k)$ execute

SAPA.

If a descent direction d_k is found, determine the stepsize a_k by a line search and go to Step 4.

Step 3. If $i = N$ then set $x^* := x_k$ and stop,

otherwise set $i := i + 1$ and go to Step 2.

Step 4. Compute

$x_{k+1} := x_k + a_k d_k$, $k := k + 1$.

Go to Step 1.

//AQD//

A convergence theorem similar to the one given in [8, Th. 6.2] and a corollary corresponding to the one given in [8, Corol. of Th. 6.2] can be established by following the similar way given there

Acknowledgement

The author is grateful to Prof. E. Spedicato for providing the opportunity of visiting the University of Bergamo, where this work was done.

Thanks are due to the Centro Calcolo of the University of Bergamo for financially supporting this wor .

References

〔1〕Abaffy J., Broyden C.G. and Spedicato E., A class of direct methods for linear equations, Numerische Mathematik 45, 361－376, 1984

〔2〕Abaffy J. and Spedicato E., ABS projection algorithms: Mathematical techniques for linear and nonlinear equations, Ellis Horwood, Chichester, 1989.

〔3〕Demyanov V.F. and Rubinov A.M., On quasidifferentiable functions Soviet Math. Dokl., 21:1, 14－17, 1980.

〔4〕Demyanov V.F., Nonsmooth analysis and directional derivatives, 3.163(456), Dipt. di Matem., Univ. di Pisa, Italia, 1989.

〔5〕Gao Y., The star-kernel for quasidifferentiable functions in one dimensional space, J. Math. Res. & Exposition, 6:1, 1988.

〔6〕Pschenichnyj B.N., Necessary conditions for extremal problems, Marcel Dekker, New York, 1971.

〔7〕Polyakova L.N., Necessary conditions for an extremum of quasidifferentiable functions, Vestnik Leningrad Univ. Math. (English Transl.), 13, 241－147, 1981.

〔8〕Spedicato E. and Xia Z.-Q., An approach for polyhedral approximations to subdifferentials of finite convex functions, to appear, Bergamo.

〔9〕Xia Z.-Q., Quasidifferential kernel, to appear, Bergamo.

〔10〕Xia Z.-Q., A note on star－kernel for quasidifferentiable functions, WP－87－66, SDS/IIASA, Laxenburg, Austria, 1987.

一类拟可微函数的极小化算法

夏 尊 铨

（大连理工大学应用数学系）

本文给出一类拟可微函数的极小化问题

$$\min\ f(x)=f_0(x)-\max f_i(x),\ x\in\mathbf{R}^n$$

的算法，其中f_0是凸函数，f_i是连续可微函数，I是一个有限的指标集．算法的核心是对次微分作外接多面体近似．该算法属于下降算法．有关算法的理论作了详细的论述．