Computational Optimization and Applications, 24, 83–93, 2003

A Solution Method for a Special Class of Nondifferentiable Unconstrained Optimization Problems

B. LUDERER
J. WEIGELT
Chemnitz University of Technology, Faculty of Mathematics, 09107 Chemnitz, Germany

Abstract. We consider quasidifferentiable functions in the sense of Demyanov and Rubinov, i. e. functions, which are directionally differentiable and whose directional derivative can be expressed as a difference of two sublinear functions, so that its "subdifferential", called the quasidifferential, consists of a pair of sets. For these functions a generalized gradient algorithm is proposed. Its behaviour is studied in detail for the special class of continuously subdifferentiable functions. Numerical test results are given. Finally, the general quasidifferentiable case is simulated by means of "perturbed" subdifferentials, where we make use of the non-uniqueness in the quasidifferential representation.

Keywords: nonsmooth optimization, quasidifferential calculus, unconstrained minimization, subdifferentiable functions, steepest descent method

1. Introduction

The study of quasidifferentiable functions is of considerable interest in nonsmooth optimization. After creating the mathematical apparatus and developing necessary optimality conditions, the interest focusses more and more on minimization methods for such problems. So far, only a few methods have been proposed, most of them being of the steepest-descent type. A good survey of numerical methods for solving unconstrained quasidifferentiable problems is given in [1], while problems involving equality constraints are discussed in [11].

The aim of the present paper consists in describing a modified gradient method for unconstrained minimization of quasidifferentiable functions, where the case of subdifferentiable functions is studied in more detail. While the convergence analysis of the algorithm proposed can be find in [8], we report here, above all, on some numerical results.

2. Theoretical background

We recall some notions needed in the following: Let Ω be an open subset of $\mathbb{R}^n$. A function f is called *quasidifferentiable* at the point x if it is directionally differentiable at this point and there exist two compact convex subsets $\underline{\partial} f(x)$ and $\bar{\partial} f(x)$ such that

$$f'(x;r) = \max_{v \in \underline{\partial} f(x)} \langle v, r\rangle + \min_{w \in \bar{\partial} f(x)} \langle w, r\rangle$$

$\forall r \in \mathbb{R}^n$. The pair of sets $Df(x) = [\underline{\partial} f(x), \bar{\partial} f(x)]$ is said to be the *quasidifferential* of the function f at the point x (consisting of the *sub-* and the *superdifferential*). It is not uniquely defined, since for an arbitrary convex compact set A, $[\underline{\partial} f(x) - A, \bar{\partial} f(x) + A]$ is also a quasidifferential. A function f is referred to as *subdifferentiable* if there exists a quasidifferential of the form $[\underline{\partial} f(x), \mathbf{0}]$.

The quasidifferential of a quasidifferentiable function considered as a point-to-set mapping need not be Hausdorff continuous and not even upper semicontinuous. However, upper semicontinuity of the generalized gradient is a desirable property e.g. to ensure convergence of minimization algorithms. Thus we focus our investigations on the class of functions for which an upper semicontinuous quasidifferential mapping Df exists. Due to a statement from [6], this class of functions coincides with the class of so-called *continuously codifferentiable* functions.

A well-known necessary criterion (see [4]) for a point x to be a local minimizer of a quasidifferentiable function f is the validity of the inclusion

$$-\bar{\partial} f(x) \subset \underline{\partial} f(x). \tag{1}$$

Points satisfying (1) are called *inf-stationary*. We aim for finding points that fulfil (1) or some weaker stationarity condition.

Let x be not an inf-stationary point. Then there are directions of descent, i.e. such directions $r \in \mathbb{R}^n$ that $f'(x;r) < 0$. Directions $\bar{r}$ of steepest descent can be determined via

$$\bar{r} = -\frac{\bar{v} + \bar{w}}{\|\bar{v} + \bar{w}\|}, \quad \|\bar{v} + \bar{w}\| = \max_{w \in \bar{\partial} f(x)} \min_{v \in \underline{\partial} f(x)} \|v + w\| \,. \tag{2}$$

They are not unique in general (cf. [4], Ch. 17).

3. General scheme of descent

In order to unify several approaches and to develop new versions of minimization methods for quasidifferentiable functions, we consider a general model of a monotonuous method of descent. To this aim, we introduce some point-to-set mapping P serving as an outer approximation of the subdifferential, where the quality of approximation depends on the concrete choice of P.

Let P be a point-to-set mapping from $\mathbb{R}^n$ into the space of convex compact subsets of $\mathbb{R}^n$ that fulfils the following assumptions:

(A1) P is upper semicontinuous on $\mathbb{R}^n$

(A2) $\underline{\partial} f(x) \subset P(x) \;\; \forall x \in \mathbb{R}^n$

(A3) If some sequence $\{x_k\}_{k=1}^{\infty}$ converges to x^*, then $\forall \delta > 0 \; \exists k_0$:

$$\underline{\partial} f(x^*) \subset P(x_k) + \delta B \quad \forall k \geq k_0 \,.$$

Following an approach of [2], the sequence $\{r_k\}$ of directions used in the algorithm described below is assumed to satisfy the following assumptions:

(B1) $\varphi(x_k, r_k) \stackrel{\text{def}}{=} \max_{v \in P(x_k)} \langle v, r_k \rangle + \min_{w \in \bar{\partial} f(x_k)} \langle w, r_k \rangle \ < \ 0 \quad \forall k$

(B2) If the sequence $\{x_k\}_{k=1}^{\infty}$ of iteration points converges to a point x^* having the property $-\bar{\partial} f(x^*) \not\subset P(x^*)$, then $\liminf_{k \to \infty} |\varphi(x_k, r_k)| \stackrel{\text{def}}{=} \gamma(x^*) > 0$.

(B3) $\|r_k\| = 1 \quad \forall k$.

Since $\varphi(x_k, r_k)$ is an upper approximation of $f'(x_k; r_k)$ it is clear from (B1) that r_k provides a direction of descent.

Algorithm

0. Given x_0, $k := 0$.
1. Determine a quasidifferential $Df(x_k)$ as well as a set $P(x_k)$ that fulfils (A1)–(A3).
2. If $-\bar{\partial} f(x_k) \subset P(x_k)$, then STOP.
3. Choose a direction r_k satisfying (B1)–(B3).
4. Evaluate $x_{k+1} = x_k + \alpha_k r_k$, where the step size α_k is chosen to be optimal (Cauchy step size).
5. $k := k + 1$, go to 1.

Remarks.

1. Due to Assumption (A2), the inclusion in Step 2 is necessary for x to be inf-stationary. Points satisfying this relation will be called *P-stationary*.
2. The requirement of determining the Cauchy step size in Step 4 is needed for the proof of convergence. In concrete evaluations it will be weakened by applying line search.

Theorem 1. *Let f be quasidifferentiable with an upper semicontinuous quasidifferential mapping. Then every accumulation point x^* of the sequence $\{x_k\}$ generated by the algorithm decribed above is P-stationary.*

The proof can be found in [8].

Note that various choices of ε-subdifferentials known from the literature meet the assumptions (A1)–(A3), e.g. the ε-subdifferential due to Demy'anov, Gamidov and Sivelina [3] as well as the Pallaschke-Recht-ε-subdifferential borrowed from [9] (for details see [8]).

Remark. In order to apply the algorithm to the minimization of a quasidifferentiable function f, apart from P one has to determine directions of descent satisfying (B1)–(B3). In the general quasidifferentiable case, however, it seems to be complicated to ensure (B2), even if r_k is chosen as an ε-steepest direction. Another approach is that of [7], where at every step a descent of at least τ is guaranteed. But this algorithm ("common descent") terminates after a finite number of steps at a point satisfying the so-called (ε, τ)-*stationarity condition*, which is only necessary for a point to be ε-inf-stationary. Reducing then τ, it is again not very clear how to guarantee (B2). Moreover, the determination of $\bar{\partial} f(x)$ is not stable in general, because the mapping $\bar{\partial} f$ is only upper semicontinuous.

For these reasons, we focus our further considerations on the restricted class of continuously subdifferentiable functions.

4. Continuously subdifferentiable case

Let $f : \mathbb{R}^n \to \mathbb{R}$ be a subdifferentiable function having an upper semicontinuous quasidifferential representation of the form $Df(x) = [\underline{\partial} f(x), \mathbf{0}]$, i.e. $\forall\, x \in \mathbb{R}^n$ the superdifferential of f consists only of zero. This leads to certain simplifications of the algorithm presented above which allow for specifying concrete directions r_k of descent satisfying (B1)–(B3). The condition of P-stationarity reduces to

$$\mathbf{0} \in P(x),$$

whereas the conditions (B1) and (B2) turn into

$$\max_{v \in P(x_k)} \langle v, r_k \rangle < 0 \tag{3}$$

and, for $x_k \to x^*$ with $\mathbf{0} \notin P(x^*)$,

$$\liminf_{k \to \infty} \left| \max_{v \in P(x_k)} \langle v, r_k \rangle \right| = \gamma(x^*) > 0\,. \tag{4}$$

We describe now two possible choices of directions r_k which fulfil (3) and (4).

(1) P-steepest direction
In this case a direction r_k is chosen in such a way that $\varphi(x_k, r) = \max\limits_{v \in P(x_k)} \langle v, r \rangle$ achieves its minimal value over all $\|r\| = 1$, i.e.

$$\varphi(x_k, r_k) = \min_{\|r\|=1} \varphi(x_k, r)\,,$$

which in case of $P = \underline{\partial}_\varepsilon f$ leads to ε-steepest directions. We get $r_k = -\frac{v_k}{\|v_k\|}$, where $\|v_k\| = \min_{v \in P(x_k)} \|v\|$ provided that $\|v_k\| \neq 0$, i.e. $\mathbf{0} \notin P(x_k)$. Otherwise we had found a P-stationary point x_k. It is not hard to see that r_k satisfies conditions (B1) and (B2) or, equivalently, (3) and (4), since

$$\max_{v \in P(x_k)} \langle v, r_k \rangle = \left\langle v_k, -\frac{v_k}{\|v_k\|} \right\rangle = -\|v_k\| < 0, \tag{5}$$

i.e. (3) is fulfilled. Furthermore, if $x_k \to x^*$, $\mathbf{0} \notin P(x^*)$, then $\min_{v \in P(x^*)} \|v\| = \|v^*\| \stackrel{\text{def}}{=} a > 0$, and from the upper semicontinuity of the mapping P it follows that there is an index k_0 such that $\|v_k\| \geq \frac{a}{2} > 0 \;\; \forall\, k \geq k_0$. This as well as (5) immediately imply (4) for P-steepest directions.

Note that this choice of directions is associated with that proposed in [3], [7], when setting P equal to the corresponding ε-subdifferentials.

(2) Modified direction
Let P fulfil (A1)–(A3). Given an outer approximation $P(x)$ of the subdifferential $\underline{\partial} f(x)$, a new set $\hat{P}(x)$ will be constructed as follows. Let $P(x) = \text{co}\,\{p_\lambda \mid \lambda \in \Lambda(x)\}$, where p_λ, $\lambda \in \Lambda(x)$, are the extreme points of $P(x)$. Set $\hat{P}(x) = \text{co}\{\frac{p_\lambda}{\|p_\lambda\|},\ \lambda \in \Lambda(x)\}$ if $\mathbf{0} \notin P(x)$ and $\hat{P}(x) = \text{co}\{\mathbf{0}, \{\frac{p_\lambda}{\|p_\lambda\|},\ \lambda \in \Lambda(x),\ \ p_\lambda \neq 0\}\}$ if $\mathbf{0} \in P(x)$, where "co" denotes the convex hull.

Let $\hat{v}_k \in \hat{P}(x_k)$ satisfy the relation $\|\hat{v}_k\| = \min_{\hat{v} \in \hat{P}(x_k)} \|\hat{v}\|$. If $\|\hat{v}_k\| > 0$, we set $\hat{r}_k = -\hat{v}_k / \|\hat{v}_k\|$ and call it *modified direction*.

It is not very hard to see that the stationarity conditions associated with P-steepest and modified directions are equivalent, i.e. $\mathbf{0} \in P(x_k) \Leftrightarrow \mathbf{0} \in \hat{P}(x_k)$. Moreover, if $\mathbf{0} \notin \hat{P}(x_k)$, then the modified direction $\hat{r}_k$ is a direction of descent. Finally, if the sequence $\{x_k\}_{k=1}^{\infty}$ converges to some $x^* \in \mathbb{R}^n$ with $\mathbf{0} \notin \hat{P}(x^*)$, then $\liminf_{k\to\infty} |\max_{v \in P(x_k)} \langle v, \hat{r}_k \rangle| > 0$.

Thus, the constructed direction $\hat{r}_k$ also fulfils the assumptions needed for the convergence theorem.

Remarks.

1. Modified directions have been introduced to overcome some difficulties when using P-steepest directions, especially, if the latter ones are located near the boundary of the cone of descent directions. In this way, sometimes better numerical results are obtained. Note, furthermore, that polyhedrality of the subdifferential (and the set $P(x)$, respectively) is very important for the construction of effective algorithms.
2. When using modified directions, the distance $d(\mathbf{0}, P(x_k))$ will not be calculated explicitly, but estimates are available:

$$d(\mathbf{0}, \hat{P}_{\min}(x_k) \leq d(\mathbf{0}, P(x_k)) \leq d(\mathbf{0}, \tilde{P}(x_k)),$$

where $\tilde{P}(x)$=co $\{p_\mu \mid \mu \in \hat{\Lambda}(x)\}$, and $\hat{\Lambda}(x) \subset \Lambda(x)$ contains a minimal number of elements such that $\min_{v \in \hat{P}(x)} \|v\| = \min\{\|v\| \mid v \in \text{co}\{\frac{p_\mu}{\|p_\mu\|} \mid \mu \in \hat{\Lambda}(x)\}\}$, while $\hat{P}_{\min}(x) \overset{\text{def}}{=} \|p_{\min}\| \hat{P}(x)$ with $\|p_{\min}\| = \min_{\lambda \in \Lambda(x)} \|p_\lambda\|$.

5. Numerical results in the subdifferentiable case

We want to describe some experience with small test examples. The algorithm proposed for continuously subdifferentiable functions has been implemented by the program "QD–optimization" written in Turbo Pascal.

The aim consists in finding a P-stationary point, where the mapping P is chosen as $P = \underline{\partial}_\varepsilon$ (in the sense of Demyanov, Gamidov, Sivelina) and the directions are the P-steepest or the modified ones. Starting with a relatively large value ε_0, during the iteration process the parameter ε is decreased until it reaches some $\varepsilon_{\min}$. For the evaluation of the distance from a point (of the superdifferential) to a set (ε-subdifferential), an algorithm due to [5] is used. The line search is realized by a combination of Fibonacci and Armijo search, including a special rule for reducing the number of function calls.

The iteration process showed the expected behaviour: In most cases the sequence of iteration points achieved some neighbourhood of the minimizer after a few number of iterations, but a stopping criterion was not yet fulfilled. This effect was mainly due to the discontinuity of the subdifferential mapping at the optimal point.

Especially in those cases, when the direction of steepest descent turns out to be unsuitable, the modified direction may be much better (cf. e.g. MI 1).
Among others, the following examples borrowed from [10] have been tested:

MI 1: (Mifflin) $f(x) = \max\{-x_1, -x_1 + 20(x_1^2 + x_2^2 - 1)\}$, $x^* = (1, 0)^\top$, $f^* = -1$, $x^{(0)} = (0.8, 0.8)^\top$
MI 2: (Mifflin) $f(x) = -x_1 + 2(x_1^2 + x_2^2 - 1) + 1.75|x_1^2 + x_2^2 - 1|$, $x^* = (1, 0)^\top$, $f^* = -1$, $x^{(0)} = (-1, 1)^\top$
LQ: $f(x) = \max\{-x_1 - x_2, -x_1 - x_2 + x_1^2 + x_2^2 - 1\}$, $x^* = (\frac{1}{2}\sqrt{2}, \frac{1}{2}\sqrt{2})^\top$, $f^* = -\sqrt{2}$, $x^{(0)} = (-\frac{1}{2}, \frac{1}{2})^\top$
QL: $f(x) = \max\{x_1^2 + x_2^2, x_1^2 + x_2^2 + 10(-4x_1 - x_2 + 4), x_1^2 + x_2^2 + 10(-x_1 - 2x_2 + 6)\}$, $x^* = (1.2, 2.4)^\top$, $f^* = 7.2$, $x^{(0)} = (-1, 5)^\top$
CB 1: (Charalambous, Bandler) $f(x) = \max\{f_1(x), f_2(x), f_3(x)\}$, $f_1(x) = x_1^2 + x_2^4$, $f_2(x) = (2 - x_1)^2 + (2 - x_2)^2$, $f_3(x) = 2e^{-x_1^2 + x_2^2}$, $x^* = (1.1390\ldots, 0.8996\ldots)^\top$, $f^* = 1.952\ldots$, $x^{(0)} = (1, 0.1)^\top$
CB 2 (Charalambous, Bandler) $f(x) = \max\{f_1(x), f_2(x), f_3(x)\}$, $f_1(x) = x_1^4 + x_2^2$, $f_2(x) = (2 - x_1)^2 + (2 - x_2)^2$, $f_3(x) = 2e^{-x_1^2 + x_2^2}$, $x^* = (1, 1)^\top$, $f^* = 2$, $x^{(0)} = (2, 2)^\top$

The following table shows the convergence behaviour of the examples mentioned above, where $\bar{x}$ denotes the last iteration point, *it*—the number of iterations and *nf*—the number of function calls. The parameters ε_0 and $\varepsilon_{\min}$ had been chosen as 0.5 and 10^{-4}, respectively.

	Steepest direction				Modified direction			
Example	$\bar{x}$	$f(\bar{x})$	*it*	*nf*	$\bar{x}$	$f(\bar{x})$	*it*	*nf*
MI 1	1.00308	−0.86008	75	426	0.99961	−0.99961	8	19
	0.03116				0.01595			
MI 2	1.00000	−0.99999	10	93	0.99016	−0.99501	9	28
	−0.00142				0.01366			
LQ	0.70665	−1.41393	12	94	0.70688	−1.41404	16	95
	0.70728				0.70716			
QL	1.20389	7.21993	16	70	1.19841	7.20573	12	73
	2.40220				2.40199			
CB 1	1.13886	1.95350	9	80	1.13966	1.95272	13	116
	0.89911				0.89925			
CB 2	1.00037	2.00087	16	106	1.00021	2.00118	22	124
	0.99969				1.00016			

6. Simulation of the quasidifferentiable case

It has been mentioned above that, although a proof of convergence was given for the general quasidifferentiable case, we did not succeed in finding a rule ensuring that condition (B2) is always fulfilled in this case. Nevertheless, we tried to examine some kind of quasidifferentiable problems as well. This is important, because also in the subdifferentiable case we can meet with such a situation, since the quasidifferential of a subdifferentiable function need not be given in a "minimal" form, i.e. in the form $Df(x) = [\underline{\partial} f(x), \mathbf{0}]$ and a transformation of a minimal form is a very hard task. Moreover, even if a function is subdifferentiable, one sometimes does not recognize this. For these reasons and for the sake of comparison, we simulated the quasidifferentiable case by transforming the ε-quasidifferential of a subdifferentiable function $[\underline{\partial}_\varepsilon, \mathbf{0}]$ into the equivalent form $[\underline{\partial}_\varepsilon - C, C]$, where C is some randomly generated polyhedron.
How to choose a direction of descent at a non-P-stationary point?

Case 1. Analogously to the "minimal" representation, take the ε-steepest direction. Due to

$$d(-C(x_k), \underline{\partial}_\varepsilon f(x_k) - C(x_k)) = d(\mathbf{0}, \underline{\partial}_\varepsilon f(x_k)),$$

where $d(\cdot, \cdot)$ means the usual Hausdorff distance, the concrete choice of ε-quasidifferentials does not have any influence on the (ε)-steepest direction. Thus the generated sequence of directions meets (B1)–(B3) and, consequently, fulfils the assumptions of the convergence theorem. However, this approach requires for all extreme points $w \in C(x_k)$ the solution of a distance problem

$$d(-w, \underline{\partial}_\varepsilon f(x_k) - C(x_k)) = \min_{v \in \underline{\partial}_\varepsilon f(x_k) - C(x_k)} \|v\| \tag{6}$$

in order to determine a direction r_k according to

$$r_k = -\frac{\bar{v} + \bar{w}}{\|\bar{v} + \bar{w}\|}, \quad \|\bar{v} + \bar{w}\| = \max_{w \in C(x_k)} \min_{v \in \underline{\partial}_\varepsilon f(x_k) - C(x_k)} \|v + w\|.$$

Case 2. In order to reduce the amount of calculation, one can make use of the following statement.

Lemma. *If* $\bar{w} \in \bar{\partial} f(x)$ *and* $d(-\bar{w}, \underline{\partial}_\varepsilon f(x) - C(x)) > 0$, *then*

$$r = -\frac{\bar{v} + \bar{w}}{\|\bar{v} + \bar{w}\|}, \quad \|\bar{v} + \bar{w}\| = \min_{v \in \underline{\partial}_\varepsilon f(x_k) - C(x_k)} \|v + \bar{w}\|$$

is a direction of descent.

Proof: We have

$$\begin{aligned} f'(x; r) &= \max_{v \in \underline{\partial} f(x)} \langle v, r \rangle + \min_{w \in \bar{\partial} f(x)} \langle w, r \rangle \\ &\le \max_{v \in \underline{\partial}_\varepsilon f(x) - C(x)} \langle v, r \rangle + \langle \bar{w}, r \rangle = \max_{v \in \underline{\partial}_\varepsilon f(x) - C(x)} \langle v + \bar{w}, r \rangle \\ &= -\left\langle \bar{v} + \bar{w}, \frac{\bar{v} + \bar{w}}{\|\bar{v} + \bar{w}\|} \right\rangle = -\|\bar{v} + \bar{w}\| < 0. \end{aligned}$$

□

Since $C(x_k)$ is a polyhedron and due to the lemma, one has to choose extreme points w of $C(x_k)$ and to solve distance problems of the kind (6) as long as a point w_0 had been found (if it exists) such that

$$d(-w_0, \underline{\partial}_\varepsilon f(x_k) - C(x_k)) \geq 0.1.$$

Then for the corresponding direction

$$r_k^0 = -\frac{\bar{v} + w_0}{\|\bar{v} + w_0\|}, \quad \|\bar{v} + w_0\| = \min_{v \in \underline{\partial}_\varepsilon f(x_k) - C(x_k)} \|v + w_0\|$$

we have $f'(x_k; r_k^0) \leq -0.1$, and we proceed in direction r_k^0. If such an element w_0 does not exist, the minimization is carried out in ε-steepest direction.

Again the assumptions of the convergence theorem are fulfilled. Note that analogous considerations can be made for modified directions, too.
The numerical results obtained are not unexpected:

- the convergence behaviour in the "minimal" and the "perturbed" cases differs in the sense that different sequences are generated, but the process converges in every case,
- for "perturbed" problems, the number of iterations and function calls increases, while the accuracy decreases,
- the advantage of modified directions (in some examples, especially MI 1) is "damped".

The approach described here for "perturbed" subdifferentiable functions can be taken as a starting point to heuristic algorithms for the minimization of general quasidifferentiable functions.

7. Further numerical experiments in the quasidifferentiable case

As mentioned above, in the general quasidifferentiable case we cannot ensure the validitiy of condition (B2) and hence the convergence of the presented algorithms. Nevertheless, it is of great interest to verify the practical efficiency and to compare the algorithms proposed above with other existing ones. So, we tested a number of problems described in the recent paper [1] (see Problems 9.1–9.6). The test problems and the obtained results are given below.

Problem 9.1.

$$f(x) = \max_{i=1,2,3} f_i(x) + \min_{i=4,5,6} f_i(x)$$
$$f_1(x) = x_1^4 + x_2^2,\ f_2(x) = (2 - x_1)^2 + (2 - x_2)^2,\ f_3(x) = 2e^{-x_1+x_2},$$
$$f_4(x) = x_1^2 - 2x_1 + x_2^2 - 4x_2 + 4,\ f_5(x) = 2x_1^2 - 5x_1 + x_2^2 - 2x_2 + 4,$$
$$f_6(x) = x_1^2 + 2x_2^2 - 4x_2 + 1;$$
$$x \in \mathbb{R}^2, x_0 = (2, 2),\ x^* = (1, 1),\ f(x^*) = 2$$

Problem 9.3.

$$f(x) = \max_{j=1,\ldots,m} \sum_{i=1}^{n} \frac{(ix_i - 1)^2}{i + j - 1} + \min_{j=1,\ldots,m} \sum_{i=1}^{n} \frac{(ix_i - 1)^2}{i + j - 1}$$

$$x \in \mathbb{R}^n, \quad x_0 = (5, \ldots, 5), \quad x^* = \left(1, \frac{1}{2}, \ldots, \frac{1}{n}\right), \; f(x^*) = 0$$

Problem 9.4.

$$f(x) = \max_{j=1,\ldots,25} \left(\sum_{i=1}^{n} a_{ij}(x_i - x_i^*) \right)^2 + \tau \cdot \min_{j=1,\ldots,25} \left(\sum_{i=1}^{n} a_{ij}(x_i - x_i^*) \right)^2$$

$$a_{ij} = (1,02 - 0,04j)^{i-1}, \; x \in \mathbb{R}^n, x_0 = (5, \ldots, 5), \; x^* = (1, \ldots, 1), \; f(x^*) = 0$$

Problem 9.5.

$$f(x) = \max_{j=1,\ldots,50} \left(\sum_{i=1}^{n} a_{ij}(x_i - x_i^*) \right)^2 + \tau \cdot \min_{j=1,\ldots,50} \left(\sum_{i=1}^{n} a_{ij}(x_i - x_i^*) \right)^2$$

$$a_{ij} = (1,01 - 0,02j)^{i-1}, \; x \in \mathbb{R}^n, \; x_0 = (5, \ldots, 5), \; x^* = (1, \ldots, 1), \; f(x^*) = 0$$

Problem 9.6.

$$f(x) = \max_{j=1,\ldots,60} \exp\left(\sum_{i=1}^{n} a_{ij} x_j (x_j + 1) \right) + \min_{j=1,\ldots,60} \exp\left(\sum_{i=1}^{n} b_{ij} x_j (x_j + 1) \right)$$

$$a_{ij} = \begin{cases} 1/2(i + j - 1), & i = 1, \ldots, 30 \\ -1/2(i + j - 1), & i = 31, \ldots, 60 \end{cases}, \quad j = 1, \ldots, n; \; b_{ij} = 1/2a_{ij}, \; i = 1, \ldots, 60,$$

$$j = 1, \ldots, n, \; x \in \mathbb{R}^n, \; x_0 = (1, \ldots, 1), \; x^* = (0, \ldots, 0), \; f(x^*) = 2$$

In the table below the results obtained by the algorithms described above with those given in [1] are compared. In doing so, we use the following notations:
it—the number of iterations, f—the number of function calls, t—the computation time (in seconds), n—the number of variables, m—the parameter occuring in the formulation of the problems, P—the above described algorithm with P-steepest direction, M—the above described algorithm with modified direction, B4, B5, B7, B9—the algorithms given in [1].

The codes have been written in Delphi 5 and numerical experiments have been carried out on an Intel 200 MHz PC under Windows 98. As above in the subdifferentiable case, we set $\varepsilon_{\min} = 10^{-4}$.

To improve the rate of convergence, in the implementation of algorithms P and M we used the following modification. Whenever some found direction ensured a given minimal decrease, the search for the best descent direction was terminated and the found direction was chosen. In particular, in Problems 9.4 and 9.5 this approach led to a significant improvement with respect to time, number of iterations and number of function calls.

Problem		P	M	B4	B5	B7	B9
9.1	*it*	47	47	33	4	52	15
	f	542	542	81	193	542	167
	t	0,03	0,03	0,05	0,01	0,05	0,01
9.3	*it*	78	80	96	39	36	23
$n = 10$	f	342	364	857	1873	174	377
$m = 25$	t	0,76	0,81	0,82	1,04	0,17	0,17
9.3	*it*	78	77	138	34	31	25
$n = 10$	f	420	421	293	1633	143	420
$m = 75$	t	2,58	2,64	2,04	2,74	0,39	0,66
9.4	*it*	66	101	38	16	37	105
$n = 10$	f	741	962	124	5281	780	2085
$\tau = 1$	t	1,38	2,43	0,50	2,25	0,39	0,82
9.5	*it*	74	74	35	8	20	52
$n = 10$	f	601	633	115	4141	606	1011
$\tau = 1$	t	3,09	2,96	0,88	3,02	0,55	0,71
9.6	*it*	13	13	70	59	282	1354
$n = 10$	f	234	234	84	3943	11259	1354
	t	1,06	1,06	1,64	13,29	55,09	4,67

For Problem 9.2 from [1] our algorithm did not succeed in finding the minimum. It remains to say that in order to develope more effective algorithms for subdifferentiable and quasidifferentiable functions, much work has still to be done.

Acknowledgment

We would like to thank Mrs. Susanne Traegner for implementing the algorithms in the quasidifferentiable case and carrying out a number of numerical experiments with great care. Furthermore, we want to thank an anonymous referee who drew our attention on the recent paper [1].

References

1. A.M. Bagirov, “Numerical methods for minimizing quasidifferentiable functions: A survey and comparison,” in Quasidifferentiability and Related Topics, Ser.: Nonconvex Optimization and its Applications, Demyanov V.F. and Rubinov A.M. (Eds.), vol. 43, pp. 33–71, Kluwer Academic Publishers: Dordrecht, 2000.
2. D.P. Bertsekas, Constrained Optimization and Lagrange Multiplier Methods, Academic Press: New York, 1982.
3. V.F. Demyanov, S. Gamidov, and T.I. Sivelina, “An algorithm for minimizing a certain class of quasidifferentiable functions,” Math. Programming Study, vol. 29, pp. 74–84, 1986.
4. V.F. Demyanov and A.M. Rubinov, Quasidifferential Calculus, Optimization Software, Inc., Publications Division: New York, 1986.

5. H. Jandl, On the Implementation of Quasidifferentiable Optimization Problems Under Constraints, Fortschritt-Berichte VDI, 16/44, VDI-Verlag, Duesseldorf, 1988.
6. L. Kuntz, "A characterization of continuously codifferentiable functions and some consequences," Optimization, vol. 22, pp. 539–547, 1991.
7. L. Kuntz and A. Pielczyk, "The method of common descent for a certain class of quasidifferentiable functions," Optimization, vol. 22, pp. 669–679, 1991.
8. B. Luderer and J. Weigelt, "A generalized steepest descent method for continuously subdifferentiable functions," Optimization, vol. 30, pp. 119–135, 1994.
9. D. Pallaschke and P. Recht, "On the steepest-descent method for a class of quasidifferentiable optimization problems," in Nondifferentiable Optimization: Motivations and Applications, Proc. IIASA Workshop (Sopron, 1984), Lect. Notes Econ. Math. Syst., vol. 255, Springer-Verlag: Berlin 1985, pp. 252–263.
10. H. Schramm, Eine Kombination von Bundle- und Trust-Region-Verfahren zur Loesung nichtdifferenzierbarer Optimierungsprobleme, Bayreuth. Math. Schr., vol. 30, 1989.
11. L.W. Zhang and Z.Q. Xia, "Newton-type methods for quasidifferential equations," J. Optimization Theory Appl., vol. 108, no. 2, pp. 439–456, 2001.