Some Methods for Minimizing of d.c. Functions

L. Polyakova(✉), V. Karelin, S. Myshkov, and E. Stankova

St. Petersburg State University, 7/9, Universitetskaya nab.,
St. Petersburg 199034, Russia
lnpol07@mail.ru, vkarelin@mail.ru, skmyshkov@mail.ru, e.stankova@spbu.ru
http://www.springer.com/gp/computer-science/lncs

Abstract. In the paper, we present algorithms for minimization of d.c. functions (difference of two convex functions) on the whole space R^n. Many nonconvex optimization problems can be described using these functions. D.c. functions are used in various applications especially in optimization, but the problem to characterize them is not trivial, due to the fact that these functions are not differentiable and certainly are not convex. The class of these functions is contained in the class of quasidifferentiable functions. Proposed algorithms are based on known necessary optimality conditions and d.c. duality. Convergence to inf-stationary points is established under fairly general natural assumptions.

Keywords: Convex function · Difference of convex functions (d.c. functions) · Subdifferential · Quasidifferentiable functions

1 Introduction

Nonsmooth optimization problems often arise in the applied fields of science. Nonconvex problems with nondifferentiable objective functions are interesting for researching. Therefore the development of theory and construction of numerical methods for solving nonsmooth problems optimization is a necessary object of research As it is well-known, functions represented as the difference of two convex functions (d.c. functions) form a linear space dense in the space of continuous functions. Hence the problem of d.c. programming forms an important class of optimization. It should be noted that algorithms of convex optimization are ineffective for the decision of nonsmooth optimization problems. But in the investigation of optimization properties of d.c. functions, of course, a well-developed apparatus of convex analysis is used. At present there exists a lot of papers devoted to the study of properties of the difference of convex functions which are the difference of two continuous convex functions (see, for example, [1–11]). The class of d.c. functions is a subclass of locally Lipschitz functions that is of interest both in analysis and optimization. An investigation of properties of such functions is, certainly, based on the fundamental knowledge and theorems

© Springer Nature Switzerland AG 2019
S. Misra et al. (Eds.): ICCSA 2019, LNCS 11622, pp. 667–677, 2019.
https://doi.org/10.1007/978-3-030-24305-0_49

of convex analysis as the class of convex functions is one of the most investigated among nonsmooth functions. A lot of real processes can be modeled with using the nondifferentiable optimization.

We will consider the following unconstrained nonsmooth optimization problem:

$$\min\{f_1(x) - f_2(x)\}, \quad x \in R^n,$$

where functions f_1, f_2 are convex on R^n.

1.1 Elementary Properties of Convex Sets and Convex Functions

The concepts of convex sets and convex functions are fundamental in Convex Analysis. Convex sets and convex functions are the main tool in theoretical studies in many subjects of nondifferentiable optimization. Formulate some definitions and statements [12,13].

A set $X \subset R^n$ is called convex, if for all $x_1 \in X$ and $x_2 \in X$ the following formula

$$\lambda x_1 + (1 - \lambda)x_2 \in X \quad \forall \lambda \in [0, 1] \subset R,$$

holds. A convex set of vectors is a set that contains all the points of any line segment joining two points of the set (see the next figure). Assume that the empty set $\emptyset$ is convex by definition. The sum of two convex sets $X_1, X_2 \subset R^n$ is called the set of

$$X = X_1 + X_2 = \{x_1 + x_2 \mid x_1 \in X_1, \quad x_2 \in X_2\}.$$

Sometimes the set $X = X_1 + X_2$ is called the algebraic sum of two convex sets X_1 and X_2 or the Minkowski sum. By writing $X_1 - X_2$ we will understand the set $X_1 + (-X_2)$. If a set $X \subset R^n$ is convex, then for $\alpha \in R$

$$\alpha X = \{y \in R^n \mid y = \alpha x, \ x \in X\}.$$

A set αX for $\alpha > 0$ is called the image set X with stretching or squeezing the space R^n in α times relative to the origin. Note some elementary properties of convex sets.

1. Let I be an arbitrary set of indices i, a set $X_i \subset R^n$ is convex for each index $i \in I$. Then $X = \bigcap_{i \in I} X_i$ is convex as well.
2. Let sets $X_1 \subset R^n$ and $X_2 \subset R^n$ be convex, then the algebraic sum $X_1 + X_2$ is also convex.

Let $f : R^n \to R \bigcup \{+\infty\} \bigcup \{-\infty\}$. The set dom $f = \{x \in R^n \mid f(x) < +\infty\}$ is called an effective domain of a convex function f. The set

$$\text{epi } f = \left\{ \begin{pmatrix} x \\ \mu \end{pmatrix} \in R^n \times R \mid f(x) \leq \mu \right\}$$

is called an epigraph of f. A function f is called convex if its epigraph is a convex set. A convex function f is said to be proper if its epigraph is non-empty and

contains no vertical lines, i.e. if $f(x) < +\infty$ for at least one x and $f(x) > -\infty$ for all x. Further we will consider only proper convex functions. For proper convex functions it is possible to give another definition which equivalent to the above. A function $f : R^n \to R \cup \{+\infty\}$ is called convex if the following relations

$$f(\lambda_1 x_1 + \lambda_2 x_2) \le \lambda_1 f(x_1) + \lambda_2 f(x_2),\ \lambda_1, \lambda_2 \ge 0,\ \lambda_1 + \lambda_2 = 1,\ \forall x_1, x_2 \in R^n,$$

hold. The conjugate of a function f is

$$f^*(v) = \sup_{x \in R^n} \{\langle x, v\rangle - f(x)\},\ v \in R^n,$$

where $\langle x, v\rangle$ is the scalar product of two vectors x and v. A conjugate function f^* is closed and convex (even when f is not). If f is a closed proper convex function then f^* is also a closed proper convex function and the equality $f(x) = f^{**}(x)$ is true. A set

$$\partial f(x) = \{v \in R^n \mid f(z) - f(x) \le \langle v, z - x\rangle\ \forall z \in R^n\}$$

is called a subdifferential of the convex function f at x. Any $v \in \partial f(x)$ is called a subgradient of f. The subdifferential $\partial f(x)$ is a convex set. If a convex function is finite on R^n then its subdifferential is nonempty compact convex set. Let f be a convex function on R^n. For any $x \in R^n$ and any $g \in R^n$ we have

$$f'(x, g) = \sup_{v \in \partial f(x)} \langle v, g\rangle,$$

where $f'(x, g)$ is the directional derivative of f at x in the direction g.

1.2 Quasidifferentiable Functions

Let a function f be defined on R^n and be directionally differentiable at a point $x \in R^n$ and its directional derivative $f'(x, g)$ can be represented in the form

$$f'(x, g) = \lim_{\lambda \downarrow 0} \frac{f(x + \lambda g) - f(x)}{\lambda} = \max_{v \in \underline{\partial} f(x)} \langle v, g\rangle + \min_{w \in \overline{\partial} f(x)} \langle w, g\rangle,$$

where $\underline{\partial} f(x) \subset R^n$, $\overline{\partial} f(x) \subset R^n$ are convex compact sets in R^n. The function f is called a quasidifferentiable at a point $x \in R^n$. A pair of sets $\mathcal{D}f(x) = [\underline{\partial} f(x), \overline{\partial} f(x)]$ is called a quasidifferential of a quasidifferentiable function f at x. The set $\underline{\partial} f(x) \subset R^n$ is called a subdifferential of f at x, the set $\overline{\partial} f(x) \subset R^n$ is called a superdifferential of f at x.

The class of quasidifferentiable functions has been introduced by V.F. Demyanov and A.M. Rubinov [14,15]. Differentiable, convex, concave functions, the maximum functions are quasidifferentiable functions.

Let f_1, f_2 be finite convex functions on R^n and

$$f(x) = f_1(x) - f_2(x), \quad x \in R^n.$$

The function f is quasidifferentiable on R^n and $\mathcal{D}f(x) = [\partial f_1(x), -\partial f_2(x)]$ is its quasidifferential at a point $x \in R^n$, where $\partial f_i(x)$ are the subdifferentials of convex functions $f_i(x)$, $i = 1, 2$, at the point $x \in R^n$ in the sense of Convex Analysis.

Necessary Conditions for a Minimum of the Difference of Convex Functions. Consider an optimization problem: find

$$\inf_{x \in R^n} f(x).$$

The following necessary optimality conditions for the function f on R^n have been proven in [3].

Theorem 1. *For a point $x^* \in R^n$ to be a minimizer of the function f on R^n, it is necessary that*

$$\partial f_2(x^*) \subset \partial f_1(x^*). \tag{1}$$

For a point $x^* \in R^n$ to be a maximizer of the function f on R^n, it is necessary that

$$\partial f_1(x^*) \subset \partial f_2(x^*).$$

If the inclusion

$$\partial f_2(x^*) \subset int\, \partial f_1(x^*)$$

holds at the point $x^* \in R^n$ then this point is a strict local minimizer of the function f on R^n. If the inclusion

$$\partial f_1(x^*) \subset int\, \partial f_2(x^*)$$

is satisfied at the point $x^* \in R^n$ then this point is a strict local maximizer of the function f on R^n.

A point $x^* \in R^n$ satisfying inclusion Eq. (1) is called an inf-stationary point of the function f.

2 Minimization of d.c. Functions

Let f_1 and f_2 be finite convex functions on R^n. Consider the difference of these functions $f(x) = f_1(x) - f_2(x)$. Suppose that functions f_1, f_2 are strongly convex on R^n with strong convexity constant $m > 0$. Note that this assumption is not burdensome, since, if they are not strongly convex, then adding to each function $f_i (i = 1, 2)$, arbitrary strongly convex function on R^n , for example, $||x||^2$, we get the property we need.

Consider the optimization problem: find

$$\inf_{x \in R^n} f(x).$$

Since both functions f_1 and f_2 are strongly convex due to the properties of strongly convex functions, the function

$$f^o(v) = f_2^*(v) - f_1^*(v), \quad v \in R^n,$$

is continuously differentiable in R^n and

$$\inf_{x \in R^n} f(x) = \inf_{v \in R^n} f^o(v).$$

Thus, the problem of finding a global minimizer of the function f on R^n is equivalent to the problem of finding the global minimizer of continuously differentiable functions f^o on R^n.

Remind that if a convex function φ is defined and convex on R^n, then the decomposition

$$\varphi(x) = \sup_{z \in R^n} \{\varphi(z) + \langle w(z), x - z\rangle\} \tag{2}$$

holds, where $w(z)$ is a subgradient of φ at z. Therefore

$$f_2(x) = \sup_{z \in R^n} \{f_2(z) + \langle w(z), x - z\rangle\}, \quad w(z) \in \partial f_2(z),$$

We have

$$f(x) = f_1(x) - f_2(x) = f_1(x) - \sup_{z \in R^n} \{f_2(z) + \langle w(z), x - z\rangle\}$$

$$= \inf_{z \in R^n} \{f_1(x) - f_2(z) - \langle w(z), x - z\rangle\}.$$

Fix the point $z \in R^n$ and arbitrary subgradient $w(z) \in \partial f_2(z)$. Let

$$\psi(x, z) = f_1(x) - f_2(z) - \langle w(z), x - z\rangle.$$

Obviously, the function $\psi(x, z)$, as the function of the argument x is strongly convex on R^n with strong convexity constant $m > 0$ for each fixed $z \in R^n$ and $w(z) \in \partial f_2(z)$. Therefore, we have

$$\inf_{x \in R^n} f(x) = \inf_{x \in R^n} \inf_{z \in R^n} \psi(x, z) = \inf_{z \in R^n} \inf_{x \in R^n} \psi(x, z),$$

and for every $z \in R^n$ and $w(z) \in \partial f_2(z)$, by using Eq. (2), we have

$$f(x) \le \psi(x, z) = f_1(x) - f_2(z) - \langle w(z), x - z\rangle.$$

Note that

$$\inf_{x \in R^n} \psi(x, z) = \min_{x \in R^n} \psi(x, z) = \psi(\bar{x}(z), z),$$

moreover, the point $\bar{x}(z)$ is unique.

Choose an arbitrary point $x_0 \in R^n$ and an arbitrary subgradient $w(x_0) \in \partial f_2(x_0)$. Put

$$\varphi_0(x) = \psi(x, x_0) = f_1(x) - f_2(x_0) - \langle w(x_0), x - x_0\rangle.$$

It is obvious that $f(x_0) = \varphi_0(x_0)$. Find

$$\min_{x \in R^n} \varphi_0(x) = \varphi_0(x_1) = \varphi_0(x_1(w_0)).$$

Due to the strong convexity of the function φ_0 and the uniqueness of the point x_1, we have

$$0 \in \partial\varphi_0(x_1) = \partial f_1(x_1) - w(x_0),$$

$$\psi(x, x_0) - \psi(x_1, x_0) \geq m||x - x_1||^2 \quad \forall x \in R^n.$$

Thus, $w(x_0) \in \partial f_1(x_1) \cap \partial f_2(x_0)$.

Consider two cases:

(a) If $w(x_0) \in \partial f_1(x_0)$, then $x_1 = x_0$.

(b) If $w(x_0) \notin \partial f_1(x_0)$, then $x_1 \neq x_0$.

As $w_0 \in \partial f_2(x_0), \quad w_0 \in \partial f_1(x_1), \quad$ then $x_0 = (f_2^*)'(w_0), \quad x_1 = (f_1^*)'(w_0)$. Hence the equality

$$x_0 - x_1 = (f_2^*)'(w_0) - (f_1^*)'(w_0) = (f^o)'(w_0)$$

holds. As

$$f_1(x_1) + f_1^*(w_0) = \langle x_1, w_0 \rangle, \quad f_2(x_0) + f_2^*(w_0) = \langle x_0, w_0 \rangle,$$

then

$$f^o(w_0) = f_2^*(w_0) - f_1^*(w_0) = f_1(x_1) - f_2(x_0) - \langle w_0, x_1 - x_0 \rangle = \psi(x_1, x_0). \quad (3)$$

But since

$$f_2(x_1) \geq f_2(x_0) + \langle w_0, x_1 - x_0 \rangle + m||x_1 - x_0||^2,$$

then

$$f^o(w_0) = f_2^*(w_0) - f_1^*(w_0) \geq f_1(x_1) - f_2(x_1) + m||x_1 - x_0||^2 = f(x_1) + m||x_1 - x_0||^2.$$

Therefore

$$f(x_1) \leq f^o(w_0) - m||x_1 - x_0||^2 \leq f(x_0) - 2m||x_1 - x_0||^2.$$

Remark 1. Note the fact that formula Eq. (3) is true for any continuous convex functions f_1 and f_2 and for each point x_1, which minimizes $\psi(x, x_0)$ on R^n when x_0 and $w_0 \in \partial f_2(x_0)$ are also fixed.

Thus, we can calculate the value of the difference of conjugate functions $f^o(v) = f_2^*(v) - f_1^*(v)$ at any point $v \in R^n$, although you need to know the point x_0, in which the vector is a subgradient of the function f_2.

Thus, if at k-th step the point x_k is not an inf - the stationary point of the function f on R^n, that is, $\partial f_2(x_k) \not\subset \partial f_1(x_k)$, then we can find the point x_{k+1} for which an inequality

$$f(x_{k+1}) \leq f^o(w_k) - m||x_{k+1} - x_k||^2 \leq f(x_k) - 2m||x_{k+1} - x_k||^2$$
$$= f(x_k) - 2m||(f^o)'(w_k)||^2,$$

holds, where $w_k \in \partial f_2(x_k), \quad w_k \notin \partial f_1(x_k)$. From these inequalities the next relations

$$f(x_{k+1}) \leq f(x_k) - 2m||x_{k+1} - x_k||^2, \quad f(x_{k+1}) \leq f(x_0) - 2m\sum_{i=0}^{k} ||x_{i+1} - x_i||^2,$$

$$f^o(w_k) \le f(x_k) - m||x_{k+1} - x_k||^2, \quad f^o(w_k) \le f(x_0) - m\sum_{i=0}^{k} ||x_{i+1} - x_i||^2,$$

$$f^o(w_k) \le f^o(w_{k-1}) - m(||x_{k+1} - x_k||^2 + ||x_k - x_{k-1}||^2)$$
$$= f^o(w_{k-1}) - m(||(f^o)'(w_k)||^2 + ||(f^o)'(w_{k-1}||^2).$$

follow. Describe the "principal" scheme of this relaxation algorithm for minimizing the function f on R^n.

2.1 Algorithm 1.

Step. 1. Take the arbitrary point $x_0 \in R^n$. If

$$\partial f_2(x_0) \subset \partial f_1(x_0),$$

then the point x_0 is inf - stationary point of the function f on R^n and the process is finished.

Step 2. Find a point $x_k \in R^n$, which is not an inf - stationary point of f on R^n. Then choose an arbitrary subgradient $w_k = w(x_k) \in \partial f_2(x_k)$. Put

$$\varphi_k(x) = \psi(x, x_k) = f_1(x) - f_2(x_k) - \langle w_k, x - x_k \rangle.$$

Find $\min\limits_{x \in R^n} \varphi_k(x) = \varphi_k(x_{k+1})$.

Step 3. If $\partial f_2(x_{k+1}) \subset \partial f_1(x_{k+1})$, then the process is finished, otherwise go to Step 2.

If a sequence $\{x_k\}$ is finite, then, by construction, the last received point is an inf - stationary point of the function f on R^n. Consider the case when the sequence $\{x_k\}$ is infinite.

Theorem 2. *If the series*

$$\sum_{i=0}^{\infty} ||x_{i+1} - x_i||^2 \tag{4}$$

diverges, then the function f is unbounded from below. If this series converges and the set

$$\mathcal{L} = \mathcal{L}(x_0) = \{x \in R^n \mid f(x) \le f(x_0)\}$$

is limited then

$$(f^o)'(w_k) \to 0.$$

Proof. Consider the first case when series Eq. (4) diverges. Then we have

$$f(x_{k+1}) \le f(x_k) - 2m||x_{k+1} - x_k||^2 \le f(x_0) - 2m\sum_{i=0}^{k} ||x_{i+1} - x_i||^2.$$

From this inequality it is easy to notice that, under our assumption on the divergence of series Eq. (4), the function f decreases infinitely on the sequence $\{x_k\}$. Let series Eq. (4) converge, then

$$||(f^o)'(w_k)|| = ||x_{k+1} - x_k|| \to 0 \quad \text{if} \quad k \to +\infty.$$

As the set $\mathcal{L}(x_0)$ is compact then the relations

$$f(x_k) \to \bar{f} > -\infty, \quad f^o(w_k) > \bar{f}$$

are true. Also by virtue of the construction the sequence $\{x_k\}$ is limited because $x_k \in \mathcal{L}(x_0)$. Therefore, the sequence of points $\{w_k\}$ is limited. Since the function f^o is continuously differentiable on R^n, then any limit point w^* of the sequence $\{w_k\}$ is a stationary point for the function f^o, that is, $(f^o)'(w^*) = 0$. The theorem is proved.

Remark 2. It is easy to see that for the convergence of the sequence $\{x_k\}$ series it is sufficient the convergence of a series $\sum\limits_{i=0}^{\infty} ||x_{i+1} - x_i|| < +\infty$.

Here is an example illustrating this algorithm for the difference of two quadratic forms with positive definite matrices.

Example 1. Consider the functions

$$f_1(x) = \frac{1}{2}\langle Ax, x\rangle + \langle c, x\rangle \quad f_2(x) = \frac{1}{2}\langle Bx, x\rangle, \quad x, c \in R^n,$$

where A and B are positive definite matrices.

It is obvious that the function f is smooth by R^n. Let the zero point $x_0 = 0$ be a starting point. If $c = 0$ then x_0 is the stationary point of the function f, otherwise we have $x_1 = \arg\min\limits_{x \in R^n} f_1(x) = -A^{-1}c$. Further put

$$\varphi_1(x) = f_1(x) - f_2(x_1) - \langle Bx_1, x - x_1\rangle, \; x_2 = -A^{-1}(c - Bx_1) = x_1 + A^{-1}Bx_1.$$

Similar to the previous cases we obtain

$$x_3 = -A^{-1}(c - Bx_2) = x_1 + A^{-1}B\,x_1 + \left(A^{-1}B\right)^2 x_1.$$

Let we find a point x_k, then

$$x_{k+1} = -A^{-1}(c - Bx_k) = x_1 + A^{-1}B\,x_1 + \cdots + \left(A^{-1}B\right)^k x_1 =$$

$$= \left[E + A^{-1}B + \cdots + \left(A^{-1}B\right)^k\right] x_1. \tag{5}$$

It is not difficult to notice that $||x_{k+1} - x_k|| = ||\left(A^{-1}B\right)^k x_1||$. Assume that all eigenvalues of the matrices $a^{-1}B$ are less than one, then the ratio is true

$$\sum_{i=0}^{\infty} \left(A^{-1}B\right)^i = \left(E - A^{-1}B\right)^{-1} =$$

$$= \left(A^{-1}A - A^{-1}B\right)^{-1} = \left(A^{-1}[A - B]\right)^{-1} = (A - B)^{-1}A.$$

If the sequence $\{x_k\}$ is infinite, then from Eq. (5) we have

$$x_k \to x^* = \sum_{i=0}^{\infty} \left(A^{-1}B\right)^i x_1 = (A-B)^{-1} A x_1.$$

Let

$$f(x) = \bar{f}_1(x) - \bar{f}_2(x) = 1/2x_1^2 - 1/2x_2^2, \quad x = (x_1, x_2) \in R^2.$$

It is not difficult to see that the function f is unlimited from below.

Since functions $\bar{f}_1, \bar{f}_2$ are not strongly convex functions, then add to each of them a function $1/2\langle x, x\rangle$. Then $f(x) = f_1(x) - f_2(x)$, where

$$f_1(x) = 1/2\langle Ax, x\rangle,\ f_2(x) = 1/2\langle Bx, x\rangle,\ A = \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} > 0, \quad B = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} > 0.$$

Case 1. Let $x_0 = (0,0) \in R^2$. Then we will not get out of the zero point, because $f_1'(x_0) = f_2'(x_0) = (0,0)$.

Case 2. Let $x_0 = (1,1) \in R^2$. Then

$$\varphi_1(x) = f_1(x) - f_2(x_0) - \langle Bx_0, x - x_0\rangle,$$

$$\varphi_1'(x) = Ax - Bx_0, \quad x_1 = A^{-1}Bx_0 = \begin{pmatrix} 0.5 & 0 \\ 0 & 2 \end{pmatrix} x_0.$$

And so on. On the k - step we find a point $x_k = \left(A^{-1}B\right)^k x_0$. Then

$$x_{k+1} - x_k = \left(A^{-1}B\right)^k \left(A^{-1}B - E\right) x_0 =$$

$$= \begin{pmatrix} (0.5)^k & 0 \\ 0 & 2^k \end{pmatrix} \begin{pmatrix} -0.5 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} (0.5)^k & 0 \\ 0 & 2^k \end{pmatrix} \begin{pmatrix} -0.5 \\ 1 \end{pmatrix} = \begin{pmatrix} -(0.5)^{k+1} \\ 2^k \end{pmatrix}.$$

Therefore, from a given starting point as a result of applying this of the algorithm, we obtain a sequence of points on which the function f is unbounded from below. Note that if we take a starting point $x_0 = (1,0) \in R^2$, then the resulting sequence of points $\{x_k\}$ tends to zero.

Example 2. Consider

$$f(x) = 5x^2 - |x| = 6x^2 - (|x| + x^2) = f_1(x) - f_2(x),\ x \in R.$$

Functions f_1, f_2 are strongly convex. The function f is unbounded from above, it has two global minimizers $x_1^* = -0.1$, $x_2^* = 0.1$, $f(x_1^*) = f(x_2^*) = -0.05$, and a local minimizer $x_0^* = 0$.

For every point $x_0 \in R$

$$f(x_0) = 6x^2 - (|x_0| + x^2) - \langle w(x_0), x - x_0\rangle,$$

where $w(x_0) \in \partial f_2(x_0)$, $f_2(x) = |x| + x^2$, $x \in R$. Consider some variants of choosing subgradients at the point $x_0 = 0$.

1. Let $w(x_0) = 1$. At this case

$$\varphi_0(x) = 5x^2 - x, \ \min_{x \in R} f_0(x) = f_0(x_1), \ x_1 = 0.1,$$

Form function $\varphi_1(x) = 5x^2 - x = \varphi_0(x)$. Then $x_2 = x_1$

$$\partial f_1(x_1) = f_{1'}^1(x_1) = 0.2, \quad \partial f_2(x_1) = f_{1'}^2(x_1) = 0.2.$$

Therefore the point x_1 is an inf - stationary point for the function f.
2. Let $w(x_0) = -1$, then we give a point $\overline{x}_2 = -0.1$, which is an inf-stationary point for the function f.
3. Let $w(x_0) = 0$, then $\varphi_0(x) = 5x^2$ and $x_1 = 0$. Therefore the point x_0 is a q-stationary point for the function f.

If the sequence $\{x_k\}$ is finite, then, by construction, the last received point is an inf - stationary point of the function f on R^n. If the sequence $\{x_k\}$ is infinite, then the theorem similar to Theorem 2 holds. Unfortunately, under this choice of a subgradient w_k we also cannot claim that any limit point of the sequence $\{x_k\}$ is an inf-stationary point of the function f on R^n.

2.2 Algorithm 2

Step 1. Take the arbitrary point $x_0 \in R^n$. If the condition (1.2) holds then x_0 is an inf - stationary point of the function f and the process is finished.
Step 2. Let we find a point $x_k \in R^n$. If at the point x_k necessary condition (1.2) holds, then the point x_k - inf is a stationary point for the function f and the process is finished. Otherwise, choose an arbitrary subgradient $w(x_k)$ from the condition $w_k = w(x_k) \in \partial f_2(x_k) \backslash \partial f_1(x_k)$. Let

$$f_k(x) = \psi(x, x_k) = f_1(x) - f_2(x_k) - \langle w_k, x - x_k \rangle.$$

Find $\min\limits_{x \in R^n} f_k(x) = f_k(x_{k+1})$. Herewith

$$f(x_{k+1}) \leq f(x_k) - m||x_{k+1} - x_k||^2.$$

If the sequence $\{x_k\}$ is finite, then by construction the last received point is an inf - stationary point of the function f on R^n. If the sequence $\{x_k\}$ is infinite, then the theorem similar to Theorem 2 holds. Unfortunately, under this choice of a subgradient w_k we also cannot claim that any limit point of the sequence $\{x_k\}$ is an inf - the stationary point of the function f on R^n.

3 Conclusion

The class of d.c. functions covers a broad family of nonconvex and often nondifferentiable functions. D.c. functions are widely used in many areas such as optimization, machine learning and statistics [16,17]. In the computational aspect the considered algorithm is significantly simplified in the case of minimizing of the difference of polyhedral functions, since in this case the quasidifferential of d.c. function consists of two polyhedra.

References

1. Hiriart-Urruty, J.-B.: From convex minimization to nonconvex minimization: necessary and sufficient conditions for global optimality. In: Clarke, F.N., Demyanov, V.F., Giannessi, F. (eds.) Nonsmooth Optimization and Related Topics. Ettore Majorana International Science Series, vol. 43, pp. 219–239. Springer, Boston (1989). https://doi.org/10.1007/978-1-4757-6019-4_13
2. Martínez-Legaz, J.-E., Volle, M.: Duality in D.C. programming: the case of several D.C. constraints. J. Math. Anal. Appl. **237**(2), 657–671 (1999). https://doi.org/10.1006/jmaa.1999.6496
3. Polyakova, L.N.: Necessary conditions for an extremum of quasidifferentiable functions. Vestnik St. Petersburg Univ. Math. **13**, 57–62 (1980). (In Russian)
4. Polyakova, L.: On global unconstrained minimization of the difference of polyhedral functions. J. Glob. Optim. **50**(2), 179–195 (2011). https://doi.org/10.1007/s10898-010-9589-6
5. Polyakova, L.N., Karelin, V.V.: On a continuous method for minimizing of nonsmooth functions. In: Petrosyan, L.A., Zhabko, A.P. (eds.) International Conference "Stability and Control Processes" in Memory of V.I. Zubov SCP 2015, pp. 338–341. IEEE (2015). https://doi.org/10.1109/SCP.2015.7342133
6. Stavroulakis, G.E., Polyakova, L.N.: Difference convex opitimization techniques in nonsmooth computational mechanics. Optim. Meth. Softw. **2**, 57–81 (1996). https://doi.org/10.1080/10556789608805644
7. Strekalovskij, A.S.: To a problem of a global extremum. Dokl. USSR Acad. Sci. **292**(5), 1062–1066 (1989). (In Russian)
8. Tao, P.D., Hoai An, L.T.: Convex analysis approaches to DC programming: theory, algorithms and applications. Acta Math. Vietnam. **22**(1), 289–355 (1997)
9. Thoai, N.V.: A modified version of Tuy's method for solving D.C. programming problems. Optimization **19**(5), 665–674 (1988)
10. Törn, A., Žilinskas, A. (eds.): Global Optimization. LNCS, vol. 350. Springer, Heidelberg (1989). https://doi.org/10.1007/3-540-50871-6
11. Tuy, H.: Global minimization of a difference of two convex functions. Math. Program. Study **30**, 150–182 (1983)
12. Leichtweiß, K.: Konvexe Mengen. Hochschultext, 1st edn. Springer-Verlag, Berlin (1980)
13. Rockafellar, R.T.: Convex Analysis. Princeton University Press, Princeton (1970)
14. Demyanov, V.F., Rubinov, A.M.: Quasidifferential Calculus. Optimization Software Inc. Publications Division, New-York (1986)
15. Demyanov, V.F., Stavroulakis, G.E., Polyakova, L.N., Panagiotopoulos, P.D.: Quasidifferentiability and Nonsmooth Modelling in Mechanics Engineering and Economics. Kluwer Academic, Dordrecht (1996)
16. Stankova, E.N., Balakshiy, A.V., Petrov, D.A., Shorov, A.V., Korkhov, V.V.: Using technologies of OLAP and machine learning for validation of the numerical models of convective clouds. In: Gervasi, O., et al. (eds.) ICCSA 2016. LNCS, vol. 9788, pp. 463–472. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-42111-7_36
17. Stankova, E.N., Ismailova, E.T., Grechko, I.A.: Algorithm for processing the results of cloud convection simulation using the methods of machine learning. In: Gervasi, O., et al. (eds.) ICCSA 2018. LNCS, vol. 10963, pp. 149–159. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-95171-3_13