Applied Analysis and Optimization
Volume 2, Number 1, 2018, 159–170

Yokohama Publishers
ISSN 2189-1664 Online Journal

GENERALIZED CUTTING PLANE METHOD BY MEANS OF MINIMUM TYPE SUBDIFFERENTIALS

ALI REZA DOAGOOEI AND CHRISTIANE TAMMER

Abstract. Applying generalized convexity techniques and using minimum type subdifferentials, we design a cutting-plane type algorithm to minimize the degree-one calm and continuous objective functions. We establish conditions to prove the convergence of the algorithm. In addition to proposing some packages from the literature, we will illustrate our method through a geometric example.

1. Introduction

Optimization algorithms constitute a wide branch of research in the theory of applied mathematics. Besides the numerous convergent algorithms which are tailored to the particular types of optimization problems, there are many heuristic algorithms which are fast but not convergent for any kind of objective functions. In other words, the main part to make an algorithm reliable is its convergence. Being an efficient convergent algorithm to solve convex optimization problems, the cutting-plane method iteratively refines a feasible set of a convex objective function through the linear inequalities. Such procedures are commonly used both to solve (not necessarily differentiable) convex optimization problems and to find integer solutions of mixed integer linear programming problems.

In a general situation, the cutting plane method has been developed in the setting of abstract convexity [14]. Then, many papers have been studying this method for several classes of objective functions, see for example [3, 4, 5, 6, 9, 11, 19]. In particular, the cutting angle method, which is a special case of cutting plane method, applies a max-min function of the form

$$h(x) = \max_{i=1,\dots,k} \min_{j=1,\dots,n} l_{ij} x_j,$$

where $x = (x_1, \dots, x_n)$ and $(l_{i1}, \dots, l_{in})$ belong to $\mathbb{R}^n$, for $i = \overline{1,k}$ and for some $k \in \mathbb{N}$. The cutting angle method for classes of increasing and convex along rays functions, increasing and co-radiant functions and for the class of Lipschitz continuous functions are examined in [1, 2, 18].

The class of degree-one calm functions is very large. It contains all locally Lipschitz functions. In this paper, we propose a cutting angle method (generalized cutting plane method) to minimize a degree-one calm function over a compact subset of $\mathbb{R}^n$.

2010 *Mathematics Subject Classification.* 90C26, 52B55, 68Q25.

Key words and phrases. Optimization algorithm, cutting plane method, degree-one calm function, Minimum-type subgradient, abstract convexity.

This work has been supported by the Center for International Scientific Studies and Collaboration (CISSC).

The layout of this paper is as follows: In Section 2, we gather definitions, notations and preliminaries which are used in the paper. In Section 3, we present a method to obtain a min-type subgradient of a degree-one calm function and we apply it to design a generalized cutting plane method. In order to solve a subproblem of aforementioned algorithm, we derive a method to minimize a max-min function in Section 4. We also present another version of cutting angle method which uses approximate solutions and prove its convergence. In addition to proposing some packages from the literature, we will finally illustrate our method through a geometric example.

2. PRELIMINARIES

Let $\mathcal{H}$ be a set of finite-valued functions defined on a set X. Recall that a function $f : X \to \mathbb{R}_{+\infty} := \mathbb{R} \cup \{+\infty\}$ is called abstract convex with respect to $\mathcal{H}$ or $\mathcal{H}$-convex on X [17], if there exists a set $U \subset \mathcal{H}$ such that

$$f(x) = \sup\{l(x) \mid l \in U\} \quad (x \in X)$$

with the convention $\sup \emptyset = -\infty$. We assume that the function $-\infty$, where $-\infty(x) = -\infty$ for all $x \in X$, is also abstract convex. Note also that $\mathcal{H}$ is called the set of all elementary functions.

For a positive integer number k, we denote the set of all functions defined as a minimum of at most k affine functions by $\mathcal{H}_k$. In other words,

$$\mathcal{H}_k := \{h : \mathbb{R}^n \to \mathbb{R} \mid h(\cdot) := \min_{1 \le i \le j} (\langle l_i, \cdot \rangle + \alpha_i),\ l_i \in \mathbb{R}^n,\ \alpha_i \in \mathbb{R},\ j \le k\}. \tag{2.1}$$

The following observation characterizes $\mathcal{H}_{n+1}$-convex functions, which also indicates how large this class of functions is.

Theorem 2.1 ([20], Theorem 5.1)**.** *Suppose $f : \mathbb{R}^n \longrightarrow \mathbb{R}_{+\infty}$ is a proper, lower semicontinuous function. Then, the following statements are equivalent:*

(i) *f is $\mathcal{H}_{n+1}$-convex.*
(ii) *There exists $\ell \in \mathcal{H}_{n+1}$ such that $f \ge \ell$.*

It is worthy saying that every function with a finite global infimum is $\mathcal{H}_{n+1}$-convex. Moreover, every star-shaped function (function with a star-shaped epigraph) is $\mathcal{H}_{n+1}$-convex as well.

Let $\mathrm{dom}(f) := \{x \in X \mid f(x) < +\infty\}$. Recall that the $\mathcal{H}$-subdifferential of the function f at a point $x_0 \in \mathrm{dom}(f)$ is defined by

$$\partial_{\mathcal{H}} f(x_0) := \{h \in \mathcal{H} \mid \forall x \in X : \ f(x) - f(x_0) \ge h(x) - h(x_0)\}.$$

Every element of $\partial_{\mathcal{H}} f(x_0)$ is called an abstract subgradient of f at the point x_0. Notice that $\partial^*_{\mathcal{H}} f(x) \subset \partial_{\mathcal{H}} f(x)$ for all $x \in X$, where

$$\partial^*_{\mathcal{H}} f(x) := \{h \in \mathrm{supp}(f, \mathcal{H}) \mid h(x) = f(x)\}.$$

Recall that a function $f : X \to \mathbb{R}_{+\infty}$ is called degree-one calm at the point $x \in \mathrm{dom}(f)$, if

$$\liminf_{y \to x} \frac{f(y) - f(x)}{\|y - x\|} > -\infty.$$

Equivalently, f is degree-one calm at $x \in \text{dom}(f)$ if and only if there exist $\epsilon > 0$ and $K > 0$ such that

$$\forall y \in B(x, \epsilon) : \quad f(y) - f(x) \geq -K\|y - x\|, \tag{2.2}$$

where $B(x, \epsilon) := \{z \in X \mid \|z - x\| \leq \epsilon\}$. The following observation shows where the $\mathcal{H}_{n+1}$-subdifferential of a function is nonempty.

Theorem 2.2 ([20], Theorem 6.3)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}_{+\infty}$ be a lower semicontinuous non-negative function and $x \in \mathbb{R}^n$ be such that $f(x) \neq +\infty$. Assume f is degree-one calm at the point x. Then $\partial^*_{\mathcal{H}_{n+1}} f(x) \neq \emptyset$.*

Recall that a family of function $\mathcal{G} := \{g \mid \mathbb{R}^n \to \mathbb{R}\}$ is called equicontinuous at a point $x_0 \in \mathbb{R}^n$ if for every $\epsilon > 0$ there exists a neighborhood V_{x_0} such that

$$|g(x) - g(x_0)| < \epsilon,$$

for all $x \in V_{x_0}$ and for all $g \in \mathcal{G}$. $\mathcal{G}$ is called equicontinuous on $C \subseteq \mathbb{R}^n$ if $\mathcal{G}$ is equicontinuous at every point of C.

Our main result will be established upon the following theorem.

Theorem 2.3 ([14], Theorem 9.1.1')**.** *Let (X, d) be a metric space and $\mathcal{H}$ be a equicontinuous family of real-valued functions defined on X. Let f be a continuous function such that*

$$f(x) = \sup\{h(x) + c \mid h \in \mathcal{H},\ c \in \mathbb{R},\ h + c \leq f\}, \tag{2.3}$$

for all $x \in X$. Let $\{\epsilon_k\}$ be a nonincreasing sequence of positive numbers tending to 0. Let the sequences $(x_k)_{k\in\mathbb{N}}$ and $(h_k)_{k\in\mathbb{N}}$ be defined by the following recursive procedure:

For the initial data $x_0, x_1, \ldots, x_{k-1} \in X$ and $h_0, \ldots, h_{k-1}$, the point $x_k \in X$ is chosen so that $\psi_k(x_k) - \inf_{x\in X} \psi_k(x) < \epsilon_k$; where $\psi_k(x) := \max\{h_0(x), h_1(x), \ldots, h_{k-1}(x)\}$. Let $h_k \in \mathcal{H}$ be chosen so that $h_k(x) \leq f(x)$ for all $x \in X$ and

$$\psi_k(x_k) \leq h_k(x_k) \leq f(x_k) \leq \psi_k(x_k) + \epsilon_k.$$

If the sequence (x_k) has a limit point x^*, then x^* is a minimizer of the following minimization problem:

$$f(x) \longrightarrow \inf, \quad x \in X.$$

Remark 2.4. Suppose that the family of functions $\mathcal{H}$ is closed under addition, i.e., $\mathcal{H} + c = \mathcal{H}$ for all $c \in \mathbb{R}$. Then, (2.3) is fulfilled if and only if f is $\mathcal{H}$-convex.

3. Generalized cutting plane method

In this section, we apply generalized cutting plane method [14] by means of minimum type subdifferentials. We first present the generalized cutting plane method and then discuss the details of the algorithm.

Let X be a normed space and $\mathcal{H}$ be a set of functions defined on an open set containing the compact convex set $C \subseteq X$. Assume that f is an abstract convex function with respect to the set $\mathcal{H}$. Consider the following minimization problem.

$$f(x) \to \min,\ x \in C. \tag{3.1}$$

To solve this problem, the generalized cutting-plane method is applied as follows:

Algorithm 1. Generalized cutting plane method (GCPM)
Step 0. Let $k := 0$ and arbitrarily choose an $x_0 \in C$.
Step 1. Find $h_k \in \partial^*_{\mathcal{H}} f(x_k)$.
Step 2. Find a global optimum for the problem

$$\max_{0 \le i \le k} h_i(x) \to \min, \ x \in C.$$

Let y^* be a solution of this subproblem.
Step 3. Set $k := k+1$, $x_k := y^*$ and return to Step 1.

A convergence of this algorithm has been proved in [14] (compare also with [17], Theorem 9.1) under weak assumptions.

Theorem 3.1 ([14], Theorem 9.1.1)**.** *Assume that the set $\mathcal{H}$ consists of continuous functions and let f be a continuous $\mathcal{H}$-convex function. Then, each limit point of the sequence (x_k) produced by GCPM is a global minimizer of the function f over the set C.*

Our main goal is to solve the minimization problem defined by (3.1) by means of a developed version of GCPM, whenever the objective function is degree-one calm and $\mathcal{H}$-convex function. As seen from (2.1), the elements of $\mathcal{H}_{n+1}$ are continuous functions. This fulfills the first assumption of Theorem 3.1. Therefore, the generalized cutting plane method by means of $\mathcal{H}_{n+1}$-subdifferentials is convergent, provided f is a continuous $\mathcal{H}_{n+1}$-convex function.

On the other hand, Theorem 2.2 ensures that $\partial_{\mathcal{H}_{n+1}} f(x)$ is nonempty for every $x \in C$ under the degree-one calm assumption of the function f over the set C. However, we need a formula to calculate a minimum-type subgradient of function f at each point. In [20] Section 6, a method has been presented in order to construct a subgradient of a non-negative, lower semicontinuous and degree-one calm function with respect to the class of min-type functions $\mathcal{H}_{n+1}$. We develop this method and formulate it through the following algorithm.

Algorithm 2. (Constructing a minimum-type subgradient)
Input: A continuous function f, $x \in C$ such that f is degree-one calm at the point x with respect to the constant K and the set $B(x, \epsilon)$, where $0 < \epsilon \le 1$ satisfying (2.2). Let $M \ge 0$ such that $f + M \ge 0$ over C.

Step 1.
If $f(x) = 0$
  $h := 0$ and exit.
Else
  $f := f(x) + M$.

Step 2. ([20], Theorem 6.2)
If $\|x\| \le 1$

$\hat{K} := \sqrt{2}K + f$; $\hat{\epsilon} := \frac{\epsilon}{3}$.
Else
$\hat{K} := \sqrt{2}K\|x\| + f$; $\hat{\epsilon} := \frac{\epsilon}{3\|x\|}$.

Step 3. ([20], Theorem 4.2)
$m = \max\{n, \frac{1}{1-\hat{\epsilon}}\} + \frac{1}{1000}$; $\epsilon_1 := \frac{n}{(1+n\hat{\epsilon})m^3\hat{K}^2}$.

Step 4. ([20], Lemma 2.1)
$\lambda := \frac{\epsilon_1 f^2}{\sqrt{n}\|(x,1)\|}$,
$d_i := \frac{1}{f}(x_1, x_2, \ldots, x_i + \lambda, \ldots, x_n, 1 - \lambda x_i)$ for $i = 1, 2, \ldots, n$.
$d_{n+1} := \frac{1}{f}(x_1 - \lambda, x_2 - \lambda, \ldots, x_n - \lambda, 1 + \lambda(x_1 + \ldots + x_n))$.

Step 5. ([20], Theorem 6.1)
$B := (\frac{d_1}{n+1}\ \frac{d_2}{n+1} \cdots \frac{d_{n+1}}{n+1})$, that is an $(n+1) \times (n+1)$ squared matrix. Here $\frac{d_i}{n+1}$ is the i^{th} column of B for $i = 1, 2, \ldots, n+1$.

Step 6.
Calculate B^{-1}. Let $a_i \in \mathbb{R}^{n+1}$ be the i^{th} row of B^{-1} and $a_i = (l_i, \alpha_i) \in \mathbb{R}^n \times \mathbb{R}$ for $i = 1, \ldots, n+1$.

Output: ([20], Theorem 6.3)
$h := \min_{1 \le i \le n+1}\langle l_i, .\rangle + \alpha_i - M$. Return h as a minimum type subgradient of f at the point x.

Remark 3.2. In Algorithm 2, if f is nonnegative over C, then clearly one could assume $M = 0$. Notice also that the minimum-type subgradient h obtained by Algorithm 2 belongs to $\partial^*_{H_{n+1}} f(x)$. See [20] (Section 6) for the details.

So far, we have proved the convergence of GCPM and presented an algorithm to find a minimum-type subgradient of the function f at each point of C. Thus, the subproblem defined in Step 2 of GCPM is the last challenging task. In the next section, we discuss how to solve this subproblem. In fact, we replace the subproblem 2 of GCPM by an inexact counterpart and then we will prove the convergence of the new algorithm.

4. Inexact generalized cutting plane method and convergence analysis

In this section, first we study some numerical methods to solve the following problem

$$(P1) \qquad \min_{x \in C} \max_{1 \le i \le k} \big(\min_{1 \le j \le n+1} \langle a_{ij}, x\rangle + \alpha_{ij}\big),$$

where C is a subset of $\mathbb{R}^n$, $k \in \mathbb{N}$, $a_{ij} \in \mathbb{R}^n$ and $\alpha_{ij} \in \mathbb{R}$, for $j = \overline{1, n+1}$ and $i = \overline{1, k}$.

There exist only a few methods in the literature to find a stationary point of a max-min function, see for examples [15, 16, 21]. Using Armijo condition [13, Section 2], these methods have been established based on a line search method. It is worthy saying that in our case, $(P1)$ should be solved in each iteration of the algorithm GCPM (Step 2). Hence, we are not able to use such iterative algorithm, which obviously increases the complexity of the algorithm. Therefore, we will use a smooth function as a surrogate function defined in [21] and will replace the objective function in $(P1)$ by it. Then, an inexact solution is computed. More precisely, let

$$\Phi(x) := \max_{1\leq i\leq k} \min_{1\leq j\leq n+1} \langle a_{ij}, x\rangle + \alpha_{ij} \quad (x \in \mathbb{R}^n).$$

Consider the following smooth approximation function for $p > 0$

$$\Phi_p(x) := \frac{1}{p} \ln\Big(\sum_{i=1}^{k} \frac{1}{\sum_{j=1}^{n+1} \exp(-p\langle a_{ij}, x\rangle - p\alpha_{ij})}\Big) + \frac{\ln(n+1)}{p}. \tag{4.1}$$

Applying Proposition 2.1 of [21], one has

$$\Phi(x) \leq \Phi_p(x) \leq \Phi(x) + \frac{\ln k(n+1)}{p} \tag{4.2}$$

for all $x \in \mathbb{R}^n$.

We will show that replacing Φ_p with sufficiently large p in Step 2 of (GCPM) necessitates a convergence to a global minimum of the function f, even if the function Φ_p is minimized approximately. To do this, first we present a new version of GCPM using the approximate under-estimator Φ_p.

Algorithm 3. Generalized cutting plane method with approximate under-estimators (GCPMwAU)
Step 0. Let $k := 0$, $\sigma > 0$, $\epsilon_0 \geq 0$ and $x_0 \in C$ be arbitrary chosen.
Step 1. Find $h_k := \min_{1\leq j\leq n+1}\langle a_{ij}, .\rangle + \alpha_{ij} \in \partial^*_{H_{n+1}} f(x_k)$ (Algorithm 2 may be used here).
Step 2. For $\epsilon_k \geq 0$, choose $p_k > 0$ sufficiently large such that

$$\frac{\ln(k+1)(n+1)}{p_k} < \min\Big\{\frac{1}{k+1}, \epsilon_k\Big\},$$

and let

$$\Phi_k(x) := \frac{1}{p_k} \ln\Big(\sum_{i=0}^{k} \frac{1}{\sum_{j=1}^{n+1} \exp(-p_k\langle a_{ij}, x\rangle - p_k\alpha_{ij})}\Big) + \frac{\ln(n+1)}{p_k}.$$

Step 3. For the following minimization problem

$$\Phi_k(x) \to \min,\ x \in C,$$

find an approximate solution $y^* \in C$ such that

$$\Phi_k(y^*) \leq \inf_{x\in C} \Phi_k(x) + \epsilon_k. \tag{4.3}$$

Step 4. If $|\Phi_k(y^*) - \Phi_{k-1}(x_k)| \leq \sigma - \frac{1}{k}$, then stop with the x_k as a $(\sigma - \epsilon_k)$-solution. That is $|\Phi_{k-1}(x_k) - \min_{x\in C} f(x)| < \sigma - \epsilon_k$.
Otherwise, choose ϵ_{k+1} such that $0 \leq \epsilon_{k+1} \leq \epsilon_k$; set $k := k+1$ and $x_k = y^*$. Go to

Step 1.

Algorithm 3 benefits from two superiorities rather than Algorithm 1. First, in Step 3 of GCPMwAU we have a smooth function to minimize, while in Step 2 of GCPM the function is nonsmooth. Secondly, in Step 3 of GCPMwAU we need to obtain an approximate solution satisfying (4.3), while in Step 2 of GCPM one needs to reach an exact solution. In the end of this section, we will show the convergence of GCPMwAU.

Now we discuss how to find a y^* satisfying (4.3) in the k^{th} iteration. For this end, we discuss two cases: either C is a box or C is defined as a system of equalities in $\mathbb{R}^n$. For the sake of simplicity, we denote Φ_k by ψ.

Case 1: Let $l, u \in \mathbb{R}^n$ and $C := [l, u]$, i.e.,

$$C = \{x \in \mathbb{R}^n \mid \ l_i \leq x_i \leq u_i, \ 1 \leq i \leq n\}, \tag{4.4}$$

where $l = (l_1, l_2, \dots, l_n), \ u = (u_1, u_2, \dots, u_n)$. Since ψ is a smooth function, one could apply the nonlinear gradient projection method [13]. This method is a sequential quadratic type method and uses either line search approaches or trust-region methods. We briefly explain this method using the line search approaches.

At the iteration k, one may replace ψ by a quadratic model at the a point z as follows:

$$q(x) := \psi(z) + \nabla\psi(z)^T(x - z) + \frac{1}{2}(x - z)^T B(x - z),$$

where B is a positive definite approximation to the Hessian $\nabla^2\psi(z)$. Therefore, a line search method is applied. Indeed, at the iterate z_i of the line search method, define the quadratic model of ψ by

$$q_i(x) := \psi(z_i) + \nabla\psi(z_i)^T(x - z_i) + \frac{1}{2}(x - z_i)^T B_i(x - z_i).$$

Then, use the gradient projection method for quadratic programming in order to find an approximate solution of the problem

$$q_i(x) \to \min, \ l \leq x \leq u. \tag{4.5}$$

Let $\hat{z}$ be an approximate solution. The descent direction for the objective function ψ is $\hat{z} - z_i$. Thus, $z_{i+1} := z_i + \alpha_i(\hat{z} - z_i)$ is the new iteration, where α_i is chosen to satisfy Armijo condition. To implement this method, one may apply the software of [22] (see also [7]) in which the Hessian approximation is defined by limited-memory BFGS updating.

Case 2: Let

$$C = \{x \in \mathbb{R}^n | \ l \leq x \leq u, \ g_i(x) = 0, i = 1, 2, \dots, m\}, \tag{4.6}$$

where, $g_i : \mathbb{R}^n \to \mathbb{R}$ is continuously differentiable functions for $i = 1, 2, \dots, m$ and $l, u \in \mathbb{R}^n$. Thus, we can write the subproblem of GCPMwAU by

$$\psi(x) \to \min, \ g_i(x) = 0, \ i = 1, 2, \dots m, \ l \leq x \leq u.$$

The augmented Lagrangian method could be applied in this case (see [8], Chapter 1-3 for the details). For $\lambda_i \in \mathbb{R},\ i = 1, \dots, m$ and $\mu > 0$, let

$$\mathcal{L}(x, \lambda; \mu) = f(x) - \Sigma_{i=1}^m \lambda_i g_i(x) + \frac{\mu}{2} \Sigma_{i=1}^m g_i^2(x),$$

be the augmented Lagrangian function. Consider the following subproblem

$$\mathcal{L}(x, \lambda; \mu) \to \min,\ l \le x \le u,$$

where $\lambda = (\lambda_1, \dots, \lambda_m)$. With the same argument to solve the problem of (4.5), the nonlinear gradient projection is applied to approximately solved this subproblem at each iteration. For the best implementation of augmented Lagrangian method, we refer the reader to MINOS [12] and LANCELOT packages [10].

Remark 4.1. In the iteration k of GCPMwAU, whenever C is determined by one of the aforementioned cases, the subproblem has an approximate solution $y_k \in C$ such that $\Phi_k(y_k) \le \inf_{x \in C} \Phi_k(x) + \epsilon_k$, for a positive tolerance ϵ_k.

Next, we discuss the convergence of the GCPMwAU. For this end, the following lemmas are needed.

Lemma 4.2. *Let $a_{ij} \in \mathbb{R}^n$ and $\alpha_{ij} \in \mathbb{R}$, for $i = 0, 1, 2, \dots$ and $j = \overline{1, n+1}$. Set*

$$\mathcal{A} := \{\xi_i(\cdot) := \min_{1 \le j \le n+1} \langle a_{ij}, \cdot \rangle + \alpha_{ij} \mid i = 0, 1, 2, \dots\}.$$

If $\{a_{ij} \mid j = \overline{1, n+1},\ i = 0, 1, 2, \dots\}$ is bounded, then $\mathcal{A}$ is equicontinuous over $\mathbb{R}^n$.

Proof. Let $x_0 \in \mathbb{R}^n$, $\xi_k \in \mathcal{A}$ and $\epsilon > 0$ be given. Since $\{a_{ij} \mid\ j = \overline{1, n+1},\ i = 0, 1, 2, \dots\}$ is bounded, there exists $M > 0$ such that $\|a_{ij}\| \le M$ for all $j = \overline{1, n+1},\ i = 0, 1, 2, \dots$. Let $\delta := \frac{\epsilon}{M}$. Assume that $x \in \mathbb{R}^n$ satisfying $\|x - x_0\| < \delta$. Then,

$$|\langle a_{ij}, x - x_0 \rangle| < \epsilon, \tag{4.7}$$

for $j = \overline{1, n+1},\ i = 0, 1, 2, \dots$. On the other hand, there exist $s, t \in \{\overline{1, n+1}\}$ such that $\langle a_{ks}, x_0 \rangle + \alpha_{ks} = \xi_k(x_0)$ and $\langle a_{kt}, x \rangle + \alpha_{kt} = \xi_k(x)$. Applying (4.7), we have

$$\xi_k(x) = \langle a_{kt}, x \rangle + \alpha_{kt} > \langle a_{kt}, x_0 \rangle + \alpha_{kt} - \epsilon \ge \xi_k(x_0) - \epsilon$$

and

$$\xi_k(x_0) = \langle a_{ks}, x_0 \rangle + \alpha_{ks} > \langle a_{ks}, x \rangle + \alpha_{ks} - \epsilon \ge \xi_k(x) - \epsilon.$$

Therefore, $|\xi_k(x) - \xi_k(x_0)| < \epsilon$. This completes the proof. □

Lemma 4.3. *Let Φ_k, h_k, x_k and ϵ_k be obtained in the iteration k of GCPMwAU. Then,*

(i) $h_k(x_k) \le \Phi_k(x_k) \le h_k(x_k) + \frac{1}{k+1}$.
(ii) $\Phi_k(x) - \frac{1}{k+1} \le f(x)$, *for all* $x \in C$.
(iii) *For* $k \ge 1$, $\Phi_k(x_k) \ge \min_{x \in C} f(x) \ge \Phi_{k-1}(x_k) - \frac{1}{k} - \epsilon_k$.

Proof. (i) follows from (4.2). (ii) follows from (i) and the fact that $h_k(x) \le f(x)$ for all $x \in C$. Finally, (iii) follows from the following inequality.

$$\Phi_k(x_k) \ge h_k(x_k) = f(x_k) \ge \min_{x \in C} f(x) \ge \min_{x \in C} \Phi_{k-1}(x) - \frac{1}{k} \ge \Phi_{k-1}(x_k) - \frac{1}{k} - \epsilon_k,$$

where the last inequality follows from (4.3). □

By the following Theorem, we show that GCPMwAU is convergent. We denote all the positive integer numbers by $\mathbb{N}$.

Theorem 4.4. *Let $0 < \rho < 1$ and $0 \leq \epsilon_{k+1} \leq \rho\epsilon_k$. Assume that there exists $M > 0$ such that $\|a_{ij}\| \leq M$, where $a_{ij} \in \mathbb{R}^n$ is obtained from Step 1 of GCPMwAU for $j = \overline{1, n+1}$ and $i \in \mathbb{N} \cup \{0\}$. Then, GCPMwAU will terminate after finitely many iterations.*

Proof. Assume, by the way of contradiction, that the number of vectors x_k generated by Step 4 of GCPMwAU is infinite. Let

$$\mathcal{A} := \{\min_{1\leq j\leq n+1} \langle a_{ij}, \cdot\rangle + \alpha_{ij} \mid i = 0, 1, 2, \ldots\}$$

and

$$\mathcal{B} := \{\Phi_k(\cdot) = \frac{1}{p_k} \ln\Big(\sum_{i=0}^{k} \frac{1}{\sum_{j=1}^{n+1} \exp(-p_k\langle a_{ij}, \cdot\rangle - p_k\alpha_{ij})}\Big) + \frac{\ln(n+1)}{p_k} \mid k \geq 0\}.$$

Applying Lemma 4.2, $\mathcal{A}$ is equicontinuous. Now, we are going to show that the family of functions $\mathcal{B}$ is equicontinuous as well. Let $\eta > 0$ be arbitrary and $x_0 \in C$. Using (4.2) and equicontinuity of $\mathcal{A}$, there exists a neighborhood V_{x_0} of x_0 such that

$$|\Phi(x) - \Phi(x_0)| < \eta$$

and

$$\Phi(x) \leq \Phi_k(x) < \Phi(x) + \eta$$

for all $\Phi \in \mathcal{A}$ and $x \in V_{x_0}$. This implies that $|\Phi_k(x) - \Phi_k(x_0)| < \eta$. Thus, $\mathcal{B}$ is equicontinuous.

It follows now from Theorem 2.3 that every limit point of (x_k) is a global minimizer of f over C. Let (x_{k_j}) be a convergent subsequence of (x_k) and $x_{k_j} \to x^* \in C$, where $\min_{x\in C} f(x) = f(x^*)$. Since $\epsilon_k \downarrow 0$, equicontinuity of $\mathcal{B}$ and Lemma 4.3 part (iii) imply that

$$\lim_{j\to+\infty} \Phi_{k_j-1}(x_{k_j}) = \lim_{j\to+\infty} \Phi_{k_j}(x_{k_j}) = f(x^*),$$

which means $|\Phi_{k_j}(x_{k_j}) - \Phi_{k_j-1}(x_{k_j})| \to 0$. Therefore, the stopping criterion of GCPMwAU ($|\Phi_k(y^*) - \Phi_{k-1}(x_k)| \leq \sigma - \frac{1}{k}$) will eventually happen. This is a contradiction. Therefore, GCPMwAU will terminate after finitely many iterations. □

The following proposition shows the convergence of GCPMwAU with an exact minimizer, whenever two consecutive iterations have the same solution.

Proposition 4.5. *Let x_k and x_{k+1} be obtained from Step 4 of GCPMwAU and $x_k = x_{k+1}$. Then, x_k is an exact global minimizer of the function f over C.*

Proof. Since $x_k = x_{k+1}$, $\max_{0\leq i\leq k} h_i = \max_{0\leq i\leq k+1} h_i$ and consequently $\Phi_k = \Phi_{k+1}$. Therefore,

$$\Phi_k(x_k) = \Phi_{k+j}(x_k) = \Phi_{k+1+j}(x_{k+1}) = \Phi_{k+1}(x_{k+1}), \tag{4.8}$$

for all $j = 0, 1, 2, \ldots$. On the other hand, by Lemma 4.3 part (iii) we have

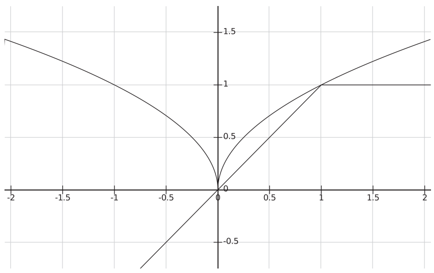
$$\Phi_{k+j+1}(x_k) \geq \min_{x\in C} f(x) \geq \Phi_{k+j}(x_k) - \frac{1}{k+j} - \epsilon_{k+j}, \tag{4.9}$$

for all $j = 0, 1, 2, \ldots$. As $j \to +\infty$, it follows from (4.8) and (4.9) that

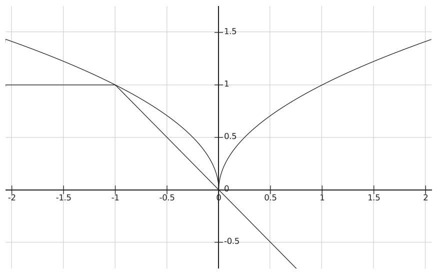
$$\min_{x \in C} f(x) = \Phi_k(x_k).$$

This completes the proof. $\square$

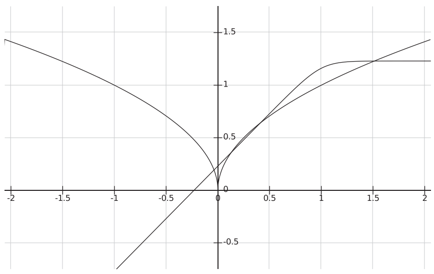
In the rest of this section and through the following geometric example, we illustrate how GCPMwAU works.



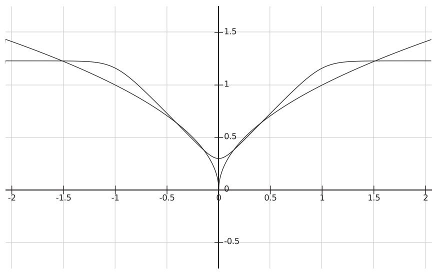
(A) The original function f and the subgradient h_0 at $x_0 = 1$.



(B) The original function f and the subgradient h_1 at $x_1 = -1$.



(C) The original function f and the smooth estimator of function h_0.



(D) The original function f and the smooth estimator of $\max\{h_0, h_1\}$.

FIGURE 1. The algorithm terminates here due to $x_2 = 0 = x_3$.

Example 4.6. Let $f(x) := \sqrt{|x|}$ and $C := [-1, 1]$. Let $x_0 = 1$ and $h_0 \in \partial^*_{\mathcal{H}_{n+1}} f(x_0)$ be defined by

$$h_0(x) = \begin{cases} x & x \leq 1, \\ 1 & x > 1. \end{cases}$$

Choose $p_0 = 10$ and $\epsilon_0 = 0$ in Step 2,

$$\Phi_0(x) := \frac{1}{10} \ln \Big(\frac{1}{\exp(-10) + \exp(-10x)} \Big) + \frac{\ln(10)}{10}.$$

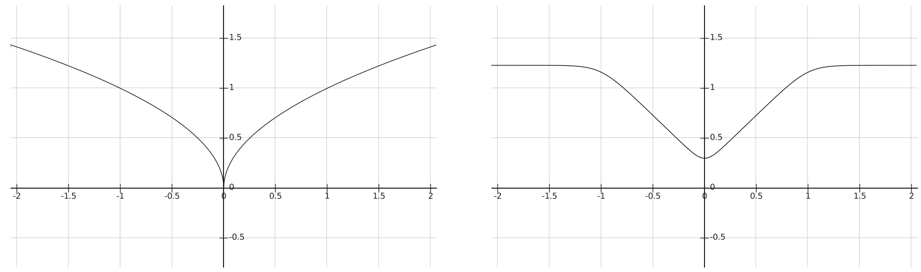
Figure 1 part (a) and (c) depict the function h_0 and its smooth estimator, respectively. The minimizer of Φ_0 on $[-1, 1]$ is $x_1 = -1$. So, we take $h_1 \in \partial^*_{\mathcal{H}_{n+1}} f(x_1)$ defined by

$$h_1(x) = \begin{cases} -x & x \geq -1, \\ 1 & x < -1. \end{cases}$$

For $p_1 = 10$ and $\epsilon_1 = 0$,

$$\Phi_1(x) := \frac{1}{10} \ln \Big(\frac{1}{\exp(-10) + \exp(-10x)} + \frac{1}{\exp(-10) + \exp(10x)} \Big) + \frac{\ln(10)}{10}.$$

Figure 1 part (b) and (d) depict the function h_1 and the smooth estimator Φ_1, respectively. Therefore, the minimizer of Φ_1 over C is $x_2 = 0$. Now, we take $h_2 = 0$ of $\partial^*_{\mathcal{H}_{n+1}} f(x_2)$. So $\max\{h_0, h_1, h_2\} = \max\{h_0, h_1\}$. This means that $\Phi_1 = \Phi_2$ and consequently $x_3 = 0 = x_2$. It follows from Proposition 4.5 that $x_2 = 0$ is the minimizer of f over $[-1, 1]$ which we have already known.



(A) The original function f. (B) The smooth function Φ_1.

FIGURE 2. The exact global minimizer is obtained by Φ_1.

It is worthy mentioning that the function $f(x) = \sqrt{|x|}$ is neither smooth nor locally Lipschitz at 0. Moreover, comparing parts (a) and (b) of Figure 2, GCPMwAU gives us a smooth function which has the same minimizer at the point 0. Nevertheless, we emphasis here that the shape of Φ_k entirely depends on both the abstract subgradients choosing in Step 1 and the values of p_k and ϵ_k picking in Step 2.

Acknowledgments

The first author is deeply grateful of Martin-Luther University Halle-Wittenberg for their warm hospitality during this research.

References

[1] M. Yu. Andramonov, A. M. Rubinov and B. M. Glover, Cutting angle methods in global optimization. Applied Mathematics Letters, **12** (1999), 95-100.

[2] M. Yu. Andramonov, A. M. Rubinov and B. M. Glover, Cutting angle method for minimizing increasing convex-along-rays functions, Research Report 7/97, SITMS, University of Ballarat, 1997.

[3] A. M. Bagirov and A. M. Rubinov, Cutting angle method and a local search. J. Glob. Optim. **27** (2003), 193-213.

[4] G. Beliakov, Cutting angle method. A tool for constrained global optimization. Optimization Methods and Software, **19** (2004) , 137-151.

[5] G. Beliakov, Extended cutting angle method of global optimization. Pacific Journal of Optimization, **4** (2008), 153-176.

[6] G. Beliakov, Geometry and combinatorics of the cutting angle method. Optimization, **52** (2003), 379-394.

[7] S. Benson and J. Moreé, A limited-memory variable-metric algorithm for bound constrained minimization, Numerical Analysis Report P909-0901, ANL, Argonne, IL, USA, 2001.

[8] D. P. Bertsekas, Constrained Optimization and Lagrange Multiplier Methods, Academic Press, New York, 1982.

[9] O. Burak, The modified cutting angle method for global minimization of increasing positively homogeneous functions over the unit simplex. Journal of Industrial and Management Optimization, **5** (2009), 825-834.

[10] A. R. Conn, N. I. M. Gould and P. L. Toint, a FORTRAN package for large-scale nonlinear optimization (Releas A), no. 17 in Springer Series in Computational Mathematics, Springer-Verlage, New York, 1992.

[11] A. Ferrer, A. Bagirov and G. Beliakov, Solving DC programs using the cutting angle method, J. Glob. Optim. **61** (2015), 71-89.

[12] B. A. Murtagh and M. A. Saunders, MINOS 5.1 User's guide. Technical Report SOL-83-20R, Stanford University, 1987.

[13] J. Nocedal and S. J. Wright, Numerical Optimization, 2nd edition. Series in Operations Research. Springer, Heidelberg, 2006.

[14] D. Pallaschke and S. Rolewicz, Foundations of Mathematical Optimization (Convex analysis without linearity) Kluwer Academic Publishers, Dordrecht, 1997.

[15] E. Polak and J. O. Royset, Algorithms for finite and semi-infinite min—max—min problems using adaptive smoothing techniques. J. Optim. Theory Appl. **119** (2003), 421-457.

[16] E. Polak, J. O. Royset and R. S. Womersle, Algorithms with adaptive smoothing for finite minimax problems. J. Optim. Theory Appl. **119** (2003), 459-484.

[17] A. M. Rubinov, Abstract Convexity and Global Optimization. London: Kluwer Academic; 2000.

[18] A. M. Rubinov and M. Yu. Andramonov, Lipschitz Programming via increasing convex-along-rays functions. Optimization Methods and Software, **10** (1999), 763-781.

[19] A. M. Rubinov and M. Yu. Andramonov, Minimizing increasing star-shaped functions based on abstract convexity. J. of Glob. Optim. **15** (1999), 19-39.

[20] A. Sheykhi and A. R. Doagooei, Radiant separation theorem and minimum-type subdifferentials. J. Optim. Theory Appl. **174** (2017), 693-711.

[21] A. Tsoukalas, P. Parpas and B. Rustem, A smoothing algorithm for finite min-max-min problems. Optim Lett. **3** (2009), 49-62.

[22] C. Zhu, R. H. Byrd, P. Lu and J. Nocedal, Algorithm 778: L-BFGS-B, FORTRAN subroutines for larg scale bound constrained optimization, ACM Transaction on Mathematical Software, **23** (1997), 550-560.

Manuscript received February 4 2018
revised April 6 2018

A. R. Doagooei
Department of Applied Mathematics, Shahid Bahonar University of Kerman, Kerman, Iran
E-mail address: `doagooei@uk.ac.ir`

C. Tammer
Institute for Mathematics, Martin-Luther University, Halle/Saale, Germany
E-mail address: `christiane.tammer@mathematik.uni-halle.de`