

Contents

1	Description Of A Task	2
2	Algorithm correctness	3
2.1	Zero Error	3
2.2	Nonzero Error	3
2.3	Undifferentiable convex function	3
3	Error's Value	4
4	Number of iterations	7
5	Tests	8
5.1	About function <i>solve</i>	8
5.2	Tests for iterations number	8

1 Description Of A Task

Let's consider a following task:

$$\min_{(x,y)} \{f(x,y) | (x,y) \in Q\},$$

where f is a convex function, Q - is a square on the plane.

Let's consider a following method. One solves task of minimization for a function $g(x) = f(x, y_0 = \frac{a}{2})$ on a segment $[0, a]$ with an accuracy δ on function. After that one calculates a sub-gradient in a received point and chooses the rectangle which the sub-gradient "does not look" in. Similar actions are repeated for a vertical segment. As a result we have the square decreased twice. Let's find a possible value of error δ_0 for task on segment and a sufficient iteration's number N to solve the initial task with accuracy ϵ on function.

Let's describe an algorithm formally. See pseudo-code [1](#).

Algorithm 1 Algorithm of the method

```

1: function METHOD(convex function  $f$ , square  $Q = [a, b] \times [c, d]$ )
     $x_0 := \text{solve}(g = f(\cdot, \frac{c+d}{2}), [a, b], \delta)$ 
     $g = \text{subgradient}(f, (x_0, \frac{c+d}{2}))$ 
2:   if  $g[1] > 0$  then
3:      $Q := [a, b] \times [c, \frac{c+d}{2}]$ 
4:   else
5:      $Q := [a, b] \times [\frac{c+d}{2}, d]$ 
6:   end if
     $y_0 := \text{solve}(g = f(\frac{a+b}{2}, \cdot), [c, d], \delta)$ 
     $g := \text{subgradient}(f, (\frac{a+b}{2}, y_0))$ 
7:   if  $g[0] > 0$  then
8:      $Q := [a, \frac{a+b}{2}] \times [c, d]$ 
9:   else
10:     $Q := [\frac{a+b}{2}, b] \times [c, d]$ 
11:   end if
12:   if StopRec() == False then
13:     Method( $f$ ,  $Q$ )
14:   end if return  $(\frac{a+b}{2}, \frac{c+d}{2})$ 
15: end function

```

2 Algorithm correctness

Let's \mathbf{x}_0 is solution of the task on segment, Q_1 is choosed rectangle, Q_2 is not choosed rectangle.

2.1 Zero Error

Lemma 1. *If the optimization task on segment is solved with zero error and the f is convex and differentiable at a point-solution, rectangle with solution of initial task was choosed correct.*

Proof. From sub-gradient definition, $\mathbf{x}^* \in \{x | (\mathbf{g}(\mathbf{x}_0), \mathbf{x}_0 - \mathbf{x}^*) \geq 0\}$. Lemma's statement follows from it and a fact that the first (or the second for vertical segment) gradient's component in point \mathbf{x}_0 is zero. \square

2.2 Nonzero Error

Theorem 2.1. *Let's the f has continuous derivative on the square. Then there is a neighbourhood of a solution of optimization task on segment such as a choice of rectangle will not change if one use any point from the neighbourhood.*

Proof. Let's consider a case when we work with horizontal segment. Case with vertical segment is considered analogously. Then we are interesting in $f'_y(x_0, y_0)$. If \mathbf{x}_0 is not solution of initial task, then $f'_y(x_0, y_0) \neq 0$.

From a continuity of the derivative:

$$\lim_{\delta \rightarrow 0} f'_y(x_0 + \delta, y_0) = f'_y(x_0, y_0)$$

Therefore,

$$\exists \delta_0 : \forall \mathbf{x}_\delta \in B_{\delta_0}(x_0) \times y_0 \Rightarrow \text{sign}(f'_y(\mathbf{x}_\delta)) = \text{sign}(f'_y(\mathbf{x}_0))$$

From it and lemma 1 theorem's statement follows. \square

2.3 Undifferentiable convex function

The method does not work for all convex functions even for zero error on segment. Let's consider following example.

Example 1.

$$f(x, y) = \max\{x - 2y, y - 2x\}, Q = [-1, 1]^2 \quad (1)$$

Function f is convex as maximum of affine functions on x and y . A solution of task on horizontal segment $[-1, 1] \times \{0\}$ is point $(0, 0)$. Its subdifferential is

$$\partial f(0, 0) = \text{conv} \left\{ (1, -2)^\top, (-2, 1)^\top \right\}.$$

So if one takes a subgradient $(-2, 1)^\top$ then a bottom rectangle will be choosed. But optimal value is point $(1, 1)$ and there is not it in choosed rectangle. Therefore, this method cannot give a solution of initial task with error less than $\frac{1}{2}$.

3 Error's Value

From derivative continuously we have following obvious result:

Lemma 2. *If f has continuous derivative. If $|f'_y(x, y_0)| > 0$ for all x on horizontal segment, then the second gradient's component has same sign at all points of segment. If $|f'_x(x_0, y)| > 0$ for all y on vertical segment, then the first gradient's component has same sign at all points of segment.*

Example 2. All functions f of the following type meet conditions of written above lemma:

$$f(x, y) = \psi(x) + \phi(y),$$

where ψ, ϕ are convex and differentiable functions.

Example 3. Let's illustrate that we can not always take any point from segment. Let's consider following task:

$$\min \left\{ (x - y)^2 + x^2 \middle| Q = [0, 1]^2 \right\}$$

On segment $[0, 1] \times \{\frac{1}{2}\}$ this task has solution $f^* = f\left(\frac{1}{4}, \frac{1}{2}\right) = \frac{1}{8}$. Derivative on y at this point is $f'_y\left(\frac{1}{4}, \frac{1}{2}\right) = \frac{1}{2}$ but at the point $\left(1, \frac{1}{2}\right)$ is equal to -1 . We can see that in this case rectangle will selected non-correctly.

Let's find possible value of error's value δ_0 . Rectangles are defined correctly for a horizontal optimization task, if:

$$\forall \delta : |\delta| < \delta_0 \Rightarrow f'_y(\mathbf{x}_0)f'_y(x_0 + \delta, y_0) > 0 \quad (2)$$

Analogically, for a vertical segment:

$$\forall \delta : |\delta| < \delta_0 \Rightarrow f'_x(\mathbf{x}_0)f'_x(x_0, y_0 + \delta) > 0 \quad (3)$$

Theorem 3.1. *Let the function f be convex and differentiable and a current rectangle be $[a, b] \times [c, d]$.*

For horizontal segment: *There is f''_{xy} on the segment. Rectangle is defined in point $(x_0 + \delta, y_0)$ correctly if one meet following condition:*

$$\delta_0 < \frac{|f'_y(\mathbf{x}_0)|}{\max_{t \in [a, b]} |f''_{xy}(t, y_0)|} \quad (4)$$

For vertical segment: *There is f''_{yx} on the segment. Rectangle is defined in point $(x_0, y_0 + \delta)$ correctly if one meet following condition:*

$$\delta_0 < \frac{|f'_x(\mathbf{x}_0)|}{\max_{t \in [c, d]} |f''_{yx}(x_0, t)|} \quad (5)$$

Proof. Let's prove this statement for horizontal segment.

Rewrite the condition (2) using Taylor formula:

$$\forall \delta : |\delta| \leq \delta_0 \Rightarrow f'_y(\mathbf{x}_0) (f'_y(\mathbf{x}_0) + f''_{xy}(\mathbf{x}_0 + (\theta\delta, 0)^\top) \delta) > 0,$$

where $\theta \in (0, 1)$

Using the written above inequality we have a following inequality for δ_0 :

$$\delta_0 < \frac{|f'_y(\mathbf{x}_0)|}{\max_{\theta \in [-1, 1]} |f''_{xy}(x_0 + \theta\delta_0, y_0)|}$$

It and an obvious inequality $\max_{\theta \in [-1, 1]} |f''_{xy}(x_0 + \theta\delta_0, y_0)| < \max_{t \in [a, b]} |f''_{xy}(t, y_0)|$ proves (3). Inequality (4) are proved similar. \square

Example 4. All positive semidefinite quadratic form meet conditions of written above theorem:

$$B(x, y) = Ax^2 + 2Bxy + Cy^2 + Dx + Ey + F,$$

$$B''_{xy} = B''_{yx} = 2B < \infty$$

where $A \geq 0, AC - B^2 \geq 0$. Also estimate for δ has some sense (estimate is non zero if \mathbf{x}_0 is not optimal solution). Also one can easy show that this estimate is accurate for task from example 3.

Theorem 3.2. *Let function f be convex and has L -Lipschitz continuous gradient and a point \mathbf{x}_0 is a solution of optimization's task on a current segment.*

The current segment is horizontal and $\exists M > 0 : \Rightarrow |f'_y(\mathbf{x}_0)| \geq M$ or the current segment is vertical and $\exists M > 0 : \Rightarrow |f'_x(\mathbf{x}_0)| \geq M$. Then rectangle is defined correctly if the possible value of error is less than $\frac{M}{L}$.

Proof. Condition (2) is met if there is a derivative $f'_y(x_0 + \delta, y_0)$ in a neighbourhood of $f'_y(\mathbf{x}_0)$ with radius $|f'_y(\mathbf{x}_0)|$:

$$|f'_y(\mathbf{x}_0) - f'_y(x_0 + \delta, y_0)| < |f'_y(\mathbf{x}_0)|$$

The L -Lipschitz continuity gives following inequality:

$$|f'_y(\mathbf{x}_0) - f'_y(x_0 + \delta, y_0)| \leq L|\delta|$$

Therefore the following possible value is sufficient to select rectangle correctly:

$$\delta_0 < \frac{M}{L} \leq \frac{|f'_y(\mathbf{x}_0)|}{L}$$

Statement for vertical segment is proved similarly. \square

4 Number of iterations

Following estimates are correct if each iterations was correct (a rectangle is selected correctly on each iterations).

Theorem 4.1. *If function f is convex and L -Lipschitz continuous, then for to solve initial task with accuracy ϵ on function one has to take a center of a current square as proximal solution and make following iteration's numbers:*

$$N = \left\lceil \log_2 \frac{La}{\sqrt{2}\epsilon} \right\rceil \quad (6)$$

where a is a size of the initial square Q .

Proof.

$$|f(\mathbf{x}^*) - f(\mathbf{x})| < L|\mathbf{x}^* - \mathbf{x}|$$

After N iterations we have a square with size $\frac{a}{2^N}$. That's why if we choose a square's center as proximal solution we have following estimates:

$$|\mathbf{x}^* - \mathbf{x}| \leq \frac{a}{\sqrt{2}} 2^{-N}$$

$$|f(\mathbf{x}^*) - f(\mathbf{x})| < \frac{1}{\sqrt{2}} La 2^{-N}$$

Therefore, for accuracy epsilon following number of iterations is sufficient:

$$N > \log_2 \frac{La}{\sqrt{2}\epsilon}$$

It proves the estimate (8). \square

There are functions which estimates from written above theorem are very accurate for.

Example 5. Let's consider following task with positive constant A :

$$\min \{A(x+y) | Q = [0, 1]^2\}$$

If one take a center of a current solution as proximal solution one have value $\frac{A}{2^N}$ after N iterations. Therefore, for accuracy ϵ one has to $\lceil \log_2 \frac{A}{\epsilon} \rceil$. For this function $L = 2A$. Therefore, estimate (8) is accurate for such tasks with little error that not more one iteration.

But any convex function is locally Lipschitz continuous at all $x \in \text{int } Q$. Therefore, we have following theorem.

Theorem 4.2. *If function f is convex and a solution $x^* \in \text{int } Q$, then for to solve initial task with accuracy ϵ on function one has to take a center of a current square as proximal solution and make following iteration's numbers:*

$$N = \left\lceil \log_2 \max \left\{ \frac{a}{\epsilon_0(\mathbf{x}^*)}, \frac{La}{\sqrt{2}\epsilon} \right\} \right\rceil \quad (7)$$

where a is a size of the initial square Q , $\epsilon_0(x^*)$ is size of neighbourhood of x^* which f is L -Lipschitz continuous in, $\Delta f = f(x_0) - f(x^*)$, x_0 is a center of square Q .

5 Tests

5.1 About function *solve*

If you look at pseudocode [1](#) you can find function *solve*. This function solves task of minimization on segment. A cost of one iteration depends on a choice of its implementation. In our tests we will use gradient descent with step $h_k = \frac{h}{k} = \frac{a}{4k}$, where a is size of current segment.

5.2 Tests for iterations number

Let's make tests for the estimate [\(8\)](#).

Functions $-\sin \frac{\pi x}{a}$ and $-\sin \frac{\pi y}{b}$ are convex on square $Q = [0, 1]^2$ when $a, b \geq 1$. Therefore, a function $-A \sin \frac{\pi x}{a} - B \sin \frac{\pi y}{b}$ is convex for all $A, B \geq 0$ as cone combination of convex function.

Functions x^n are convex and monotonously non-decreasing on $[0, 1]$ for all $n \in \mathbb{N}$ that's why functions $(-A \sin \frac{\pi x}{a} - B \sin \frac{\pi y}{b} + A + B + D)^n$ are convex for all $D \geq 0$.

Therefore, following function is convex:

$$f(x, y) = -A_1 \sin \frac{\pi x}{a_1} - B_1 \sin \frac{\pi y}{b_1} + \sum_{n=2}^N \left(-A_n \sin \frac{\pi x}{a_n} - B_n \sin \frac{\pi y}{b_n} + A_n + B_n + D_n \right)^n,$$

where $A_i, B_i, D_i \geq 0$ and $a_i, b_i \geq 1$ for all $i = \overline{1, n}$.

The function f is differentiable infinite times and we can use it to test the method.

Let's take $a_1 = \dots = a_n = a$ and $b_1 = \dots = b_n = b$:

$$f(x, y) = -A_1 \sin \frac{\pi x}{a} - B_1 \sin \frac{\pi y}{b} + \sum_{n=2}^N \left(-A_n \sin \frac{\pi x}{a} - B_n \sin \frac{\pi y}{b} + A_n + B_n + D_n \right)^n,$$

where $A_i, B_i, D_i \geq 0$ for all $i = \overline{1, n}$ and $a, b \geq 1$.

We have following estimates for derivatives on horizontal and vertical segments:

$$|f'_x| \Big|_{x=x_0} \geq \left(A_1 + \sum_{n=2}^N n A_n D_n^{n-1} \right) \frac{\pi}{a} \left| \cos \frac{\pi x_0}{a} \right|$$

$$|f'_y|_{y=y_0} \geq \left(B_1 + \sum_{n=2}^N n B_n D_n^{n-1} \right) \frac{\pi}{b} \left| \cos \frac{\pi y_0}{b} \right|$$

Therefore this functions met conditions of lemma 2 and we can take gradient at any point of segment. One can see examples of functions f on fig. 1

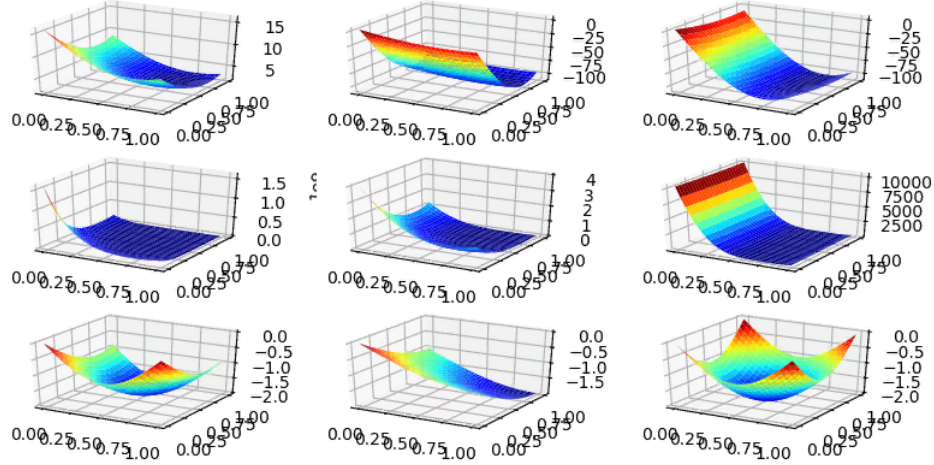


Figure 1: Examples of tests functions

We will use a and b from $[1, 2]$ and $N = 2$. Let's solve task of minimization function f with different parameters on square $[0, 1]^2$ through new method and compares number of iteration with estimate (8).

Result of the test you can see on fig. 2.

On graphic N is number of done iterations for accuracy ϵ , $\log_2 \frac{La}{\sqrt{2}\epsilon}$ is the parameter of tests tasks. Line $N = \log_2 \frac{La}{\sqrt{2}\epsilon}$ is our estimates (8).

We can see that there are tests points under this line. It confirms our estimates (8).

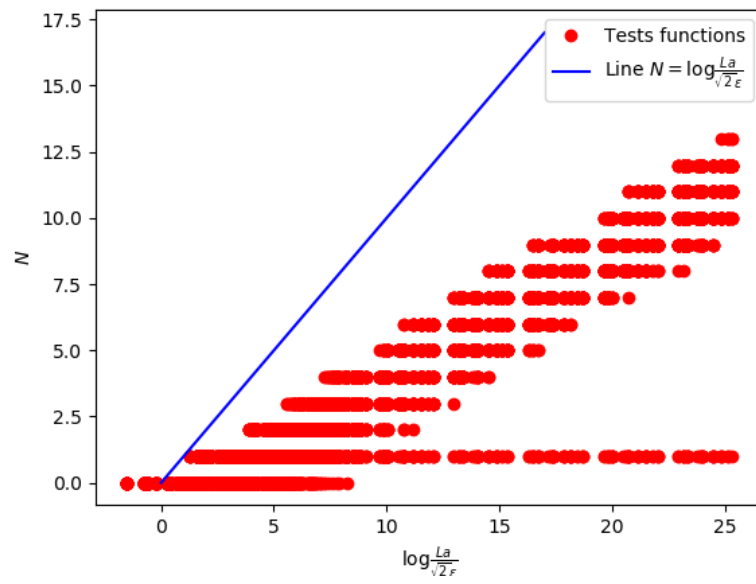


Figure 2: Tests results for iterations number