Contents

1 Description Of A Task

Let's consider a following task:

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

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 $g(x) = f\left(x, y_0 = \frac{a}{2}\right)$ 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

function METHOD(convex function f, square $Q = [a, b] \times [c, d]$)
    $x_0 := solve(g = f(\cdot, \frac{c+d}{2}), [a, b], \delta)$
    $g = subgradient(f, (x_0, \frac{c+d}{2}))$
    if g[1] $>$ 0 **then**
        $Q := [a, b] \times [c, \frac{c+d}{2}]$
    else
        $Q := [a, b] \times [\frac{c+d}{2}, d]$
    end if
    $y_0 := solve(g = f(\frac{a+b}{2}, \cdot), [c, d], \delta)$
    $g := subgradient(f, (\frac{a+b}{2}), y_0)$
    if g[0] $>$ 0 **then**
        $Q := [a, \frac{a+b}{2}] \times [c, d]$
    else
        $Q := [\frac{a+b}{2}, b] \times [c, d]$
    end if
    if StopRec() $==$ False **then**
        Method(f, Q)
    end ifreturn $(\frac{a+b}{2}, \frac{c+d}{2})$
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. □

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:

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

Therefore,

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

From it and lemma 1 theorem's statement follows.

□

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.

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

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

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

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

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

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

Analogically, for a vertical segment:

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

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

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:*

$$\boxed{\delta_0 < \frac{|f'_y(\boldsymbol{x}_0)|}{\max\limits_{t\in[a,b]} |f''_{xy}(t, y_0)|}} \tag{4}$$

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:*

$$\boxed{\delta_0 < \frac{|f'_x(\boldsymbol{x}_0)|}{\max\limits_{t\in[c,d]} |f''_{yx}(x_0, t)|}} \tag{5}$$

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

Rewrite the condition (1) using Taylor formula:

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

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

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

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

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

Theorem 3.2. *Let's function f is convex and has L-Lipschitz continuous gradient.*

And on horizontal segment $\exists M_1 : \forall \boldsymbol{x} \Rightarrow |f'_y(\boldsymbol{x})| > M_1$. Then rectangle is defined correctly if the possible value of error is not more $\frac{M_1}{L}$.

And on vertical segment $\exists M_2 : \forall \boldsymbol{x} \Rightarrow |f'_x(\boldsymbol{x})| > M_2$. Then rectangle is defined correctly if the possible value of error is not more $\frac{M_2}{L}$

Proof. Condition (1) 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 $\left|f'_y(\mathbf{x}_0)\right|$:

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

The L-Lipschitz continuity gives following inequality:

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

Theorem's estimate for vertical segment follows from two written above inequality. Inequality 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 make any following iteration's numbers:*

$$N = \left\lfloor \log_2 \frac{La}{\sqrt{2}\epsilon} + 1 \right\rfloor, \tag{4.1}$$

$$N = \left\lfloor \log_2 \frac{\Delta f}{\epsilon} + 1 \right\rfloor, \tag{4.2}$$

where a is a size of the initial square Q, $\Delta f = f(x_0) - f(x^*)$, x_0 is a center of 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 squares center as proximal solution we have following estimates:

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

$$|f(\mathbf{x}^*) - f(\mathbf{x})| < La2^{-N}$$

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

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

It proves the estimate (4.1).

$$|f(x_1) - f(x_2)| < L|\mathbf{x}_1 - \mathbf{x}_2|$$

In particular, if x_0 is a squares center and $\mathbf{x}^*$ is a solution then

$$\Delta f = f(x_0) - f(x^*) < L|\mathbf{x}_0 - \mathbf{x}^*| \leq L\frac{a}{\sqrt{2}}$$

Using it and estimate (4.1) we have estimate (4.2).

□

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

Theorem 4.2. *If function f is convex and a solution $x^* \in \operatorname{int} Q$, then for to solve initial task with accuracy ϵ on function one has to make following iteration's number:*

$$N = \left\lfloor \log_2 \max\left\{\frac{a}{\epsilon_0(\boldsymbol{x}^*)}, \frac{La}{\sqrt{2}\epsilon}\right\} + 1 \right\rfloor, \tag{4.3}$$

$$N = \left\lfloor \log_2 \max\left\{\frac{a}{\epsilon_0(\boldsymbol{x}^*)}, \frac{\Delta f}{\epsilon}\right\} + 1 \right\rfloor, \tag{4.4}$$

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