One Method for Convex Optimization on Square

I. Kuruzov[1] F. Stonyakin[1,2]

[1]Moscow Institute of Physics and Technology

[2]V.I.Vernadsky Crimean Federal University

COG, 2019

Method Description

Task:

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

where f is a convex function, $Q = [a,b] \times [c,d] \in \mathbb{R}^2$.

One iteration

- Find minimum with accuracy δ on central horizontal segment in square and calculate gradient at this point

Method Description

Task:

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

where f is a convex function, $Q = [a,b] \times [c,d] \in \mathbb{R}^2$.

One iteration

- Find minimum with accuracy δ on central horizontal segment in square and calculate gradient at this point
- Choose rectangle which anti-gradient looks in

Method Description

Task:

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

where f is a convex function, $Q = [a,b] \times [c,d] \in \mathbb{R}^2$.

One iteration

- Find minimum with accuracy δ on central horizontal segment in square and calculate gradient at this point
- Choose rectangle which anti-gradient looks in
- Analogically for vertical segment in the rectangle

Plan

Let f be convex and has L-Lipschitz continuous gradient.

$$\text{sign}\, f'_y(x_0) = \text{sign}\, f'_y(x_{current})$$

For this it is sufficient:

$$|f'_y(x_0) - f'_y(x_{current})| \leq |f'_y(x_0)|$$

Let f be convex and has L-Lipschitz continuous gradient.

$$\text{sign } f'_y(x_0) = \text{sign } f'_y(x_{current})$$

For this it is sufficient:

$$|f'_y(x_0) - f'_y(x_{current})| \leq |f'_y(x_0)|$$

$$\boxed{\delta < \frac{|f'_y(x_0)|}{L}}$$

$$\text{sign } f'_y(x_0) = \text{sign } f'_y(x_{current})$$

For this it is sufficient:

$$|f'_y(x_0) - f'_y(x_{current})| \leq |f'_y(x_{current})|$$

$$\text{sign } f'_y(x_0) = \text{sign } f'_y(x_{current})$$

For this it is sufficient:

$$|f'_y(x_0) - f'_y(x_{current})| \leq |f'_y(x_{current})|$$

$$\boxed{\delta < \frac{|f'_y(x_{current})|}{L}}$$

Convergence

Function f is convex. Size of square Q is equal to a. One takes a center of a current square as approximate solution.

Estimate through Lipschitz function constant

Let function f be L_f-Lipschitz continuous. Then for accuracy ϵ on function it is sufficient:

$$N = \left\lceil \log_2 \frac{L_f a}{\sqrt{2}\epsilon} \right\rceil. \tag{1}$$

Estimate through Lipschitz gradient constant

Let function f have L_g-Lipschitz continuous gradient. Moreover, solution of initial task is **internal point**. Then for accuracy ϵ on function it is sufficient:

$$N = \left\lceil \frac{1}{2} \log_2 \frac{L_g a^2}{4\epsilon} \right\rceil. \tag{2}$$

Convergence

If following inequallity is met

$$\frac{2L_f^2}{L_g} \geq \epsilon$$

then the estimate through L_g is better than through L_f.

- Quadratic Function

$$f(x,y) = (Ax+By)^2 + Cy^2 + Dx + Ey$$

$$(x^*, y^*) \in Q$$

$$L_f = \max_{(x,y)\in Q} \|\nabla f(x,y)\|$$

$$L_g = \max |\lambda (H(f))|$$

Experiments

Iterations Number

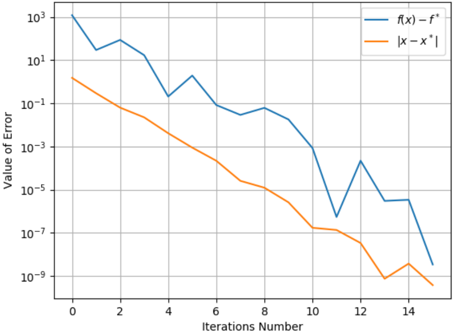
Theoretical Iteration Number through function constant 40

Theoretical Iteration Number through gradient constant 20

Experiments

Iterations Number

Theoretical Iteration Number through function constant 40
Theoretical Iteration Number through gradient constant 20



Experiments

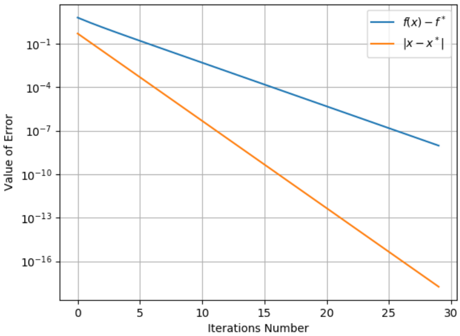
Iterations Number

$$f(x,y) = (x+y)^2 + x^2,\ Q = [1,2]^2,\ (x^*, y^*) = 1, 1$$

Theoretical Iteration Number through function constant 30
Theoretical Iteration Number through gradient constant 14

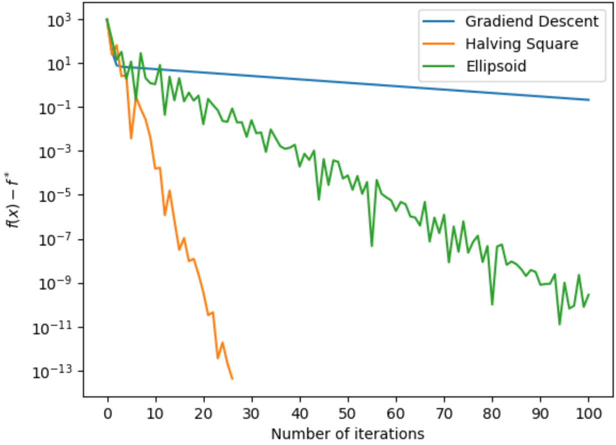
$$f(x,y) = (x+y)^2 + x^2,\ Q = [1,2]^2,\ (x^*, y^*) = 1, 1$$

Theoretical Iteration Number through function constant 30
Theoretical Iteration Number through gradient constant 14



Experiments

Comparison Of Methods



Work time

Gradiend Descent 0.0045

Halving Square 0.0209

Ellipsoid 0.0283

Summary

- Strategy for solution on segment
- Convergence Results
- Experiments for this method
  - Iterations number
  - Comparison with different methods