Різні модифікації градієнтного методу

29 травня 2023 р.

Визначення

Визначення

Умова ліпшиця для градієнта з константою L

$$\|\nabla f(x_1) - \nabla f(x_2)\| \geq L\|x_1 - x_2\|$$

Визначення

Гладка функція — це функція, що має неперервну похідну на всій області визначення.

Визначення

Визначення

m-сильно опукла функція - функція, що задовільняє нерівність:

$$f(y) \geq f(x) + \nabla f(x)(y - x) + \frac{m}{2}\|y - x\|^2$$

Формалювання проблеми

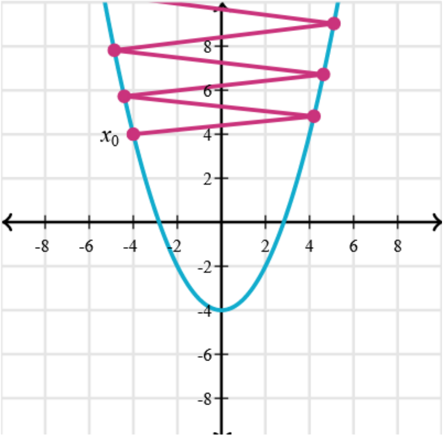
$\min\limits_{x \in \mathbb{R}^n} f(x)$, де f гладка і опукла функція. Часто ще додають сильну m-опуклість та умову Ліпшиця.

Згадаємо базовий градієнтний методу

$$x_{k+1} = x_k - \alpha \nabla f(x_k)$$

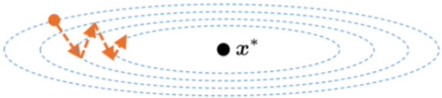
$\alpha = \frac{1}{L}, N = O(\frac{L}{m} \ln(\frac{\|x_0 - x^*\|^2}{\varepsilon}))$

Проблема незмінного кроку

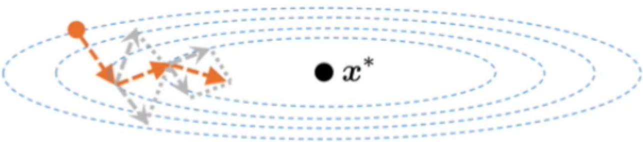


Змінний крок

$$x_{k+1} = x_k - \alpha_k \nabla f(x_k), \alpha_k \to 0$$



Метод важкого шара Поляка

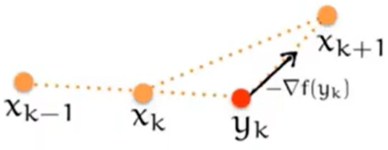


Метод важкого шара Поляка

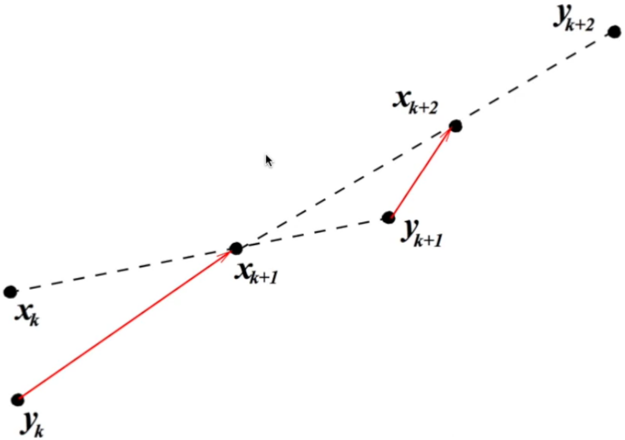
$x_{k+1} = x_k - \alpha_k \nabla f(x_k) + \beta(x_k - x_{k-1}), \beta$- масса шара

Метод Нестерова

$$x_{k+1} = y_k - \alpha_k \nabla f(y_k)$$
$$y_{k+1} = x_{k+1} + \beta(x_{k+1} - x_k)$$



Метод Нестерова



Метод Нестерова

Теорема

Для досягнення точності ε, отримання x_N, такого що $f(x_N) - f^* \leq \varepsilon$, методу Нестерова потрібно

- в опуломому випадку: $N = O\left(\frac{LR^2}{\sqrt{\varepsilon}}\right)$
- у сильно опуклому випадку $N = O\left(\sqrt{\frac{L}{\mu}} \log\left(\frac{1}{\varepsilon}\right)\right)$