Применение градиентного метода

Постановка задачи. Записать градиентный метод решения задачи

4134

$$\min(b \cdot x^2 + b \cdot y^2)$$

$$x_1 = -1, \qquad y_1 = -1$$

Постоянный шаг

$$\alpha_t = \frac{1}{3}, \alpha_t = \frac{2}{3}, \alpha_t = 1.$$

$$b = \frac{4}{10}.$$

$$\min(0.4 \cdot x^2 + 0.4 \cdot y^2)$$

Что Вы можете сказать относительно сходимости метода при разных α_t? Насколько существенно начальное приближение x_1, y_1 – как оно влияет на сходимость метода? на скорость сходимости? Исследуйте наискорейший спуск.

Решение. Градиентный метод:

$$z_{t+1} = z_t - \alpha_t \cdot (\nabla f)(z_t),$$

$$f = 0.4 \cdot x^2 + 0.4 \cdot y^2, \qquad \nabla f = \begin{pmatrix} \frac{\partial}{\partial x}(0.4 \cdot x^2 + 0.4 \cdot y^2) \\ \frac{\partial}{\partial y}(0.4 \cdot x^2 + 0.4 \cdot y^2) \end{pmatrix} = \begin{pmatrix} 0.8x \\ 0.8y \end{pmatrix},$$

$$z_t = \begin{pmatrix} x_t \\ y_t \end{pmatrix}.$$

Тогда

$$\begin{pmatrix} x_{t+1} \\ y_{t+1} \end{pmatrix} = \begin{pmatrix} x_t \\ y_t \end{pmatrix} - \alpha_t \begin{pmatrix} 0.8x_t \\ 0.8y_t \end{pmatrix}$$

или в виде системы:

$$\begin{cases} x_{t+1} = x_t - \alpha_t \cdot 2x_t = (1 - 2\alpha_t)x_t \\ y_{t+1} = y_t - \alpha_t \cdot 2y_t = (1 - 2\alpha_t)y_t \end{cases}$$

Для **постоянного шага**: $\alpha_t = \frac{1}{3}$. Тогда

$$\begin{cases} x_{t+1} = \left(1 - 0.8 \cdot \dfrac{1}{3}\right) x_t = 0.7333\ xt \\ y_{t+1} = \left(1 - 0.8 \cdot \dfrac{1}{3}\right) y_t = 0.7333\ yt \end{cases}, \qquad t = 1,2, \dots$$

Выпишем несколько итераций:

t	1	2	3	4
x_t	0.8	0.7333	0.53777	0.39431
y_t	0.8	0.7333	0.53777	0.39431

Для **постоянного шага**: $\alpha_t = \frac{2}{3}$. Тогда

$$\begin{cases} x_{t+1} = \left(1 - 0.8 \cdot \dfrac{2}{3}\right) x_t = 0.46666\ xt \\ y_{t+1} = \left(1 - 0.8 \cdot \dfrac{2}{3}\right) y_t = 0.46666\ yt \end{cases}, \qquad t = 1,2, \dots$$

t	1	2	3	4
x_t	0.8	0.21777	0.101162	0.04742
y_t	0.8	0.7333	0.101162	0.04742

Для **постоянного шага**: $\alpha_t = 1$. Тогда

$$\begin{cases} x_{t+1} = (1 - 0.8 \cdot 1)x_t = 0.2\ xt \\ y_{t+1} = (1 - 0.8 \cdot 1)y_t = 0.2yt \end{cases}, \qquad t = 1,2, \dots$$

t	1	2	3	4
x_t	0.2	0.04	0.008	0.0016
y_t	0.2	0.04	0.008	0.0016

Метод Наискорейшего спуска

$f(X)=0.4*x_1^2+0.4*y^2$

Итерация №1.

$X^0=(-1;-1)$.

Вычислим значение функции в начальной точке $f(X_0) = 0.8$.

В качестве направления поиска выберем вектор градиент в текущей точке:

$$\nabla f(X) = \begin{pmatrix} 0.8x_1 \\ 0.8y \end{pmatrix}$$

Значение градиента в точке X_0:

$$\nabla f(X_0) = \begin{pmatrix} -0.8 \\ -0.8 \end{pmatrix}$$

Проверим критерий остановки:

$|\nabla f(X_0)| < \varepsilon$

Имеем:

$|\nabla f(X_0)| = \sqrt{0.8^2 + 0.8^2} = 1.131 > 0$

Сделаем шаг вдоль направления антиградиента.

$$X_1 = X_0 - \lambda_1 \nabla f(X_0) = \begin{pmatrix} -1 \\ -1 \end{pmatrix} - \lambda_1 \begin{pmatrix} -0.8 \\ -0.8 \end{pmatrix} = \begin{pmatrix} 0.8\lambda_1-1.0 \\ 0.8\lambda_1-1.0 \end{pmatrix}$$

Вычислим значение функции в новой точке.

$f(X_1) = 0.4*(0.8*\lambda_1-1.0)^2+0.4*(0.8*\lambda_1-1.0)^2$

Найдем такой шаг, чтобы целевая функция достигала минимума вдоль этого направления. Из необходимого условия существования экстремума функции ($f'(X)=0$):

$1.024*\lambda_1-1.28 = 0$

Получим шаг: $\lambda_1 = 1.25$

Выполнение этого шага приведет в точку:

$$X_1 = \begin{bmatrix} -1 \\ -1 \end{bmatrix} - 1.25 \begin{bmatrix} -0.8 \\ -0.8 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Итерация №2.

X^1=(0;0).

Вычислим значение функции в точке $f(X_1) = 0$.

Значение градиента в точке X_1:

$$\nabla f(X_1) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Проверим критерий остановки:

$|\nabla f(X_1)| < \varepsilon$

Имеем:

$|\nabla f(X_1)| = 0>0$

Сделаем шаг вдоль направления антиградиента.

$$Y = X_1 - \lambda_2 \nabla f(X_1) = \begin{bmatrix} 0 \\ 0 \end{bmatrix} - \lambda_2 \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Вычислим значение функции в новой точке.

$f(Y) = 0.4*(0)^2+0.4*(0)^2$

Найдем такой шаг, чтобы целевая функция достигала минимума вдоль этого направления. Из необходимого условия существования экстремума функции (f'(X)=0):

0 = 0

Данное уравнение не имеет действительных корней.

Функция не имеет глобального экстремума.