Отчёт по лабораторной работе №4
по теме „Градиентные методы"

Выполнили: Порошин и Заболотские
Подгруппа: 2

1. Постановка задачи:

а) Реализовать градиентный метод первого порядка с постоянным шагом и решить задачу поиска минимума

б) Реализовать градиентный метод второго порядка с выбором шага по методу Пшеничного. Показать в ходе вычислительного эксперимента сверхлинейную скорость сходимости (для точности 0.01)

Задача: $f(x,y) = x^2 + 5y^2 + \sin(4x+5y) + 3x + 2y$

$(x,y) - ? \; : \; f(x,y) \to \min$

2. Исследование применимости метода

а) Градиентный метод первого порядка

Воспользуемся следующей теоремой о сходимости метода:

Th Пусть ф-я f дифференцируема и ограничена снизу на $\mathbb{R}^n$, а её градиент удовлетворяет условию Липшица:
$\| f'(x_1) - f'(x_2) \| \le M \| x_1 - x_2 \|$; $x_1, x_2 \in \mathbb{R}^n$, $M > 0$.
Тогда при произвольной начальной точке $x^0 \in \mathbb{R}^n$ имеем:
$$\lim_{k\to\infty} \| f'(x^k) \| = 0$$

Теорема показывает, что любая предельная точка x^* последовательности $\{x^k\}$, генерируемой градиентным методом является стационарной точкой минимизируемой функции f. Если же в условиях теоремы функция f выпукла, то любая стационарная точка x^* есть точка минимума.

$f = x^2 + 5y^2 + \sin(4x+5y) + 3x + 2y$ $\qquad$ $\operatorname{grad} f = (2x + 4\cos(4x+5y) + 3,\; 10y + 5\cos(4x+5y) + 2)$

Ограниченность снизу:

$$x^2 + 5y^2 + \sin(4x+5y) + 3x + 2y = x^2 + 3x + \frac{9}{4} - \frac{9}{4} + 5y^2 + 2y + \frac{1}{5} - \frac{1}{5} + \sin(4x+5 =$$

$$= \left(x + \frac{3}{2}\right)^2 + \left(\sqrt{5}y + \frac{1}{\sqrt{5}}\right)^2 - \frac{9}{4} - \frac{1}{5} + \sin(4x+5y) \geqslant 0 - \frac{9}{4} + 0 - \frac{1}{5} - 1$$

Условие Липшица:

Непрерывная дифференцируемость функции влечёт локальное выполнение условия Липшица.

б) Градиентный метод второго порядка

Воспользуемся следующей теоремой о сходимости метода:

Th Если для минимизации функции $f(x)$, удовлетворяющей условиям:

$$(*) \quad m\|y\|^2 \le (f''(x)y, y) \le M\|y\|^2, \; m>0 \qquad \forall x,y \in E^n$$

(такие ф-ии ограничены снизу и у них $\exists!$ (·) минимума x^*), используется метод Ньютона, в котором α_k выбирается из условия:

$$f(x) - f(x_k) \le \varepsilon\alpha(f'_k, p_k); \quad 0<\varepsilon<\frac{1}{2}$$

($p_k = -(f''(x_k))^{-1} f'(x_k)$ — направление спуска $f(x)$; то независимо от выбора начальной точки x_0, последовательность $\{x_k\}$ сходится к точке минимума со сверхлинейной скоростью:

$$\|x_{N+l} - x_*\| \le C\lambda_N \dots \lambda_{N+l} \quad \forall l \ge 0, \; \lambda_i \xrightarrow[i\to\infty]{} 0$$

$(N, C < \infty; \lambda_{N+} < 1)$

Существование обратной матрицы вторых производных и её ограниченность

$$f''(x,y) = \begin{array}{cc} f''_x & f''_y \\ \left(\begin{array}{c} 2-16\sin(4x+5y) \\ -20\sin(4x+5y) \end{array}\right. & \left.\begin{array}{c} -20\sin(4x+5y) \\ 10-25\sin(4x+5y) \end{array}\right) \end{array}$$

Для этого достаточно $\det f''(x,y) \ne 0 \Leftrightarrow$

$$\neg\left([2-16\sin(4x+5y)][10-25\sin(4x+5y)] + [20\sin(4x+5y)]^2 = 0\right] \Leftrightarrow$$

$$\neg\left(10[42 - 21\sin(4x+5y) - 40\cos(4x+5y)] = 0\right) \text{ — истина}$$

Т.е. $\det f''(x,y) \ne 0 \;\; \forall x,y \in E^2$, где E — некоторая окрестность точки $(-0.1, 1)$ (пояснение в результатах)

Выполнение неравенства $(*)$

Мы имеем непрерывные и ограниченные вторые производные функции $f(x)$ на некоторой области, значит:

$$\exists\, m = \inf_{E^2} f''(x,y); \; M = \sup_{E^2} f''(x,y)$$

3 Описание алгоритма

а) Градиентный метод первого порядка с постоянным шагом

Дано: x_0 – начальная точка, ε – необходимая точность

Алгоритм:

1° $x_{k+1} = x_k - \nabla f(x_k)$

2° Если $\|x_{k+1} - x_k\| < \varepsilon$ – выходим из алгоритма

3° переходим к шагу 1°, k++

б) Градиентный метод второго порядка с выбором шага по методу Пшеничного.

Дано: x_0 – начальное приближение, ε – необход. точность

Алгоритм:

1° $p_k = - f''(x_k) \cdot f'(x_k)$

2° Находим a_k:

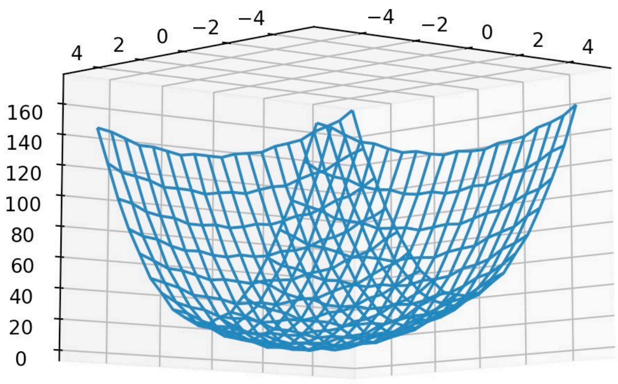
1. $a_k = 1$
2. $x = x_k + p_k \cdot a_k$
3. Если $f(x) - f(x_k) > \varepsilon a (f', p_k)$ выходим иначе $a /= 2$
4. возвр. к шагу 2.

3° $x_{k+1} = x_k + a_k p_k$

4° Если $\|x_{k+1} - x_k\| < \varepsilon$ – выходим из алгоритма иначе возвр. к шагу 1°

4 Результаты

Прежде чем анализировать результаты метода, зададимся вопросом поиска начального приближения.
Для этого построим график функции:



Так же проанализировав аналитический вид ф-ции — сумма параболоида и периодической ф-ции можно сделать вывод о наличии глобального минимума в окрестности точки $(-0.1, -1)$, которую и возьмем за начальное приближение

а) Градиентный метод

$\varepsilon = 0.1$

$\bar{x} = (-1.57, -0.33)$

$f(\bar{x}) = -3.3$

$\varepsilon = 0.001$

$\bar{x} = (-1.648, -0.247)$

$f(\bar{x}) = -3.418$

$\varepsilon = 0.0001$

$\bar{x} = (-1.6478, -0.2371)$

$f(\bar{x}) = -3.4183$

Проверка результатов

$$f(x,y) = x^2 + 5y^2 + \sin(4x + 5y) + 3x + 2y$$

$$\frac{\partial f}{\partial x} = 2x + 4\cdot\cos(4x+5y) + 3$$

$$\frac{\partial f}{\partial y} = 10y + 5\cos(4x+5y) + 2$$

$$\frac{\partial^2 f}{\partial x^2} = 2 - 16\sin(4x+5y)$$

$$\frac{\partial^2 f}{\partial y^2} = 10 - 25\sin(4x+5y)$$

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = -10\sin(4x+5y)$$

Траектория $(\bar{x}_k, f(x_k))$ в ходе выполнения метода градиента

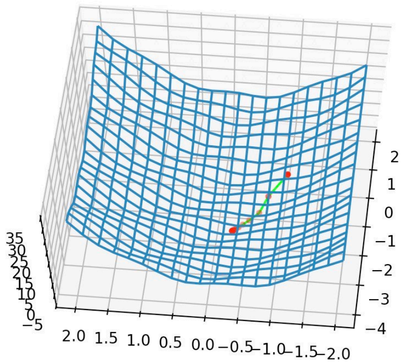
Необход. и дост. усл. лок. минимума

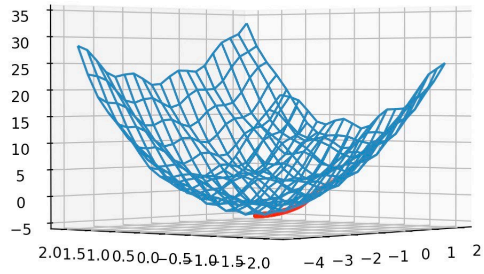
$$\frac{\partial f}{\partial x}(x,y) = 8.50e-6 \approx 0$$

$$\frac{\partial f}{\partial y}(x,y) = -5.48e-6 \approx 0$$

$$H(x,y) = 527.766$$

Графическая иллюстрация:





На первом рисунке изображены шаги метода (каждый 100 ый). Второй рисунок иллюстрирует достижение глобального минимума.

б) Градиентный метод 2го порядка с выбором шага по способу Пшеничного.

$\varepsilon = 0.1$

$\bar{x} = (-1.64, -0.24)$

$f(\bar{x}) = -3.41$

$\varepsilon = 0.001$

$\bar{x} = (-1.643, -0.240)$

$f(\bar{x}) = -3.418$

$\varepsilon = 0.0001$

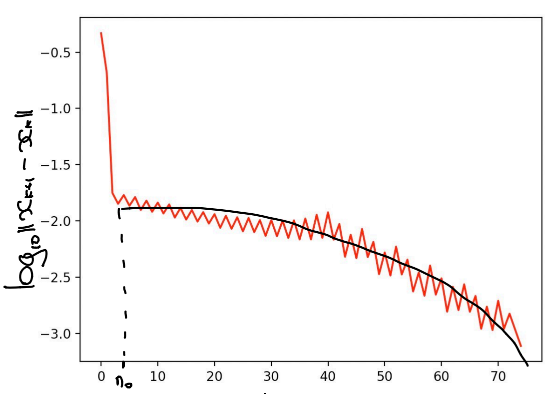
$\bar{x} = (-1.6479, -0.2375)$

$f(\bar{x}) = -3.4183$

Подтверждение сверхлинейной скорости сходимости для точности 0.01

Будем рассматривать поведение величины $\log\|x_{k+1} - x_k\|$ на каждом шаге. В случае линейной скорости сходимости график зависимости данной величины от номера шага будет линейным, причем константа q из опр. лин. сход. будет характеризовать угол наклона.

Рассмотрим результаты полученные в ходе эксперимента:



Используется точность 0.001 так как для $\varepsilon = 0.01$ минимум достигается слишком быстро.

Из графика видно, что начиная с шага n_0 график имеет нелинейную скорость убывания step

Результаты вычислений двумя методами из одной точки

Возьмем нач. точку $(-1.2, -0.31)$; Точность $\varepsilon = 0.01$

Результаты:

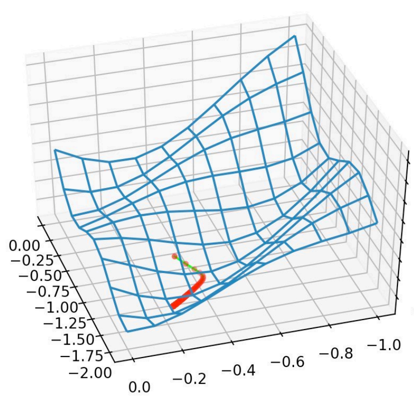
Метод 1го порядка

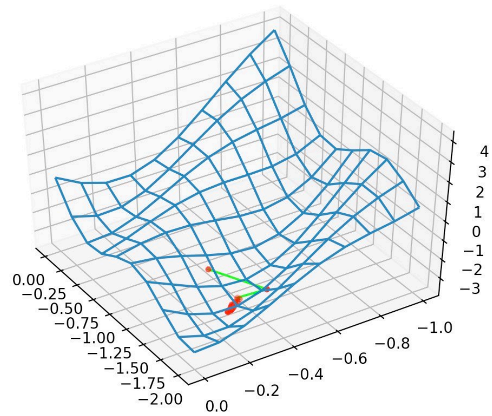
$x_{min} = (-1.64, -0.23)$
$f(x_{min}) = -3.14$

Метод 2го порядка с выб. по способу Пшеничного

$x_{min} = (-1.64, -0.23)$
$f(x_{min}) = -3.14$

Т.е. видно что оба метода достигли необходимой точности. Но интересно пронаблюдать траектории:

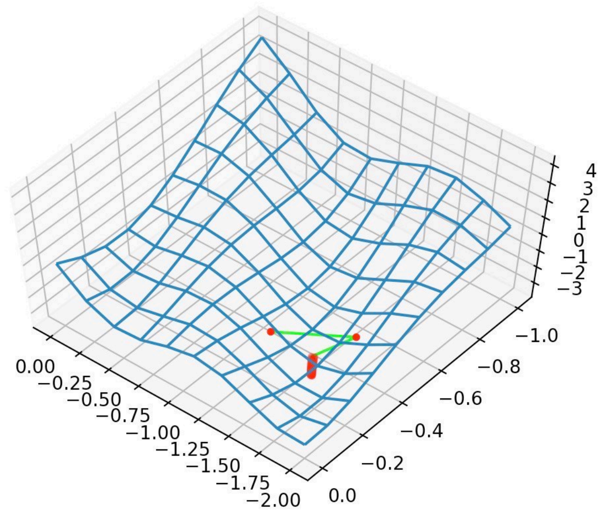




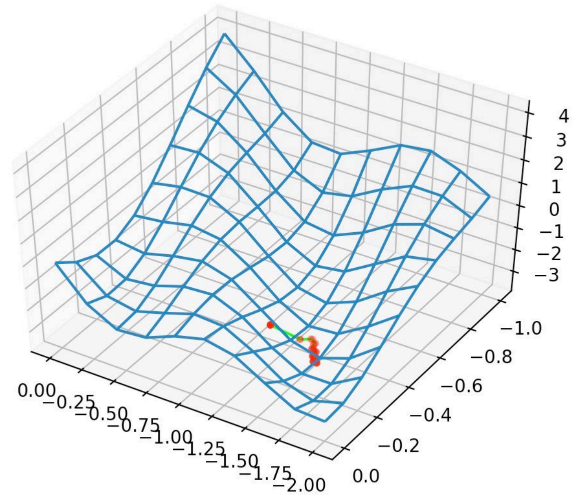
Видно, что градиентный метод 1го порядка плавно спустился с горки в овраг, когда метод 2го порядка двигался скачками.

Это можно объяснить тем, что у м. 1 порядка "скорость" зависит только от значения градиента, в силу того, что шаг постоянный. А метод с выбором шага по способу Пшеничного исходит из максимально возможного шага и уменьшает до тех пор, пока не получит выполнение убывания функции.

В подтверждение слов, рассмотрим поведение 2го метода с разными параметрами поиска шага (скорость убывания ε)



Мелкое дробление $\varepsilon = 0.4$



Крупное дробление $\varepsilon = 0.05$