Домашнее задание №1.
Одномерная минимизация. Градиентный метод. Метод Ньютона.

1 (3 балла)

1. Реализуйте метод дихотомии, метод золотого сечения и метод Фибоначчи.
2. Выберите произвольную несимметричную относительно некоторой вертикальной оси унимодальную функцию.
3. Сравните сходимость методов по времени и по числу итераций, необходимых для достижения заданной точности.
4. Проанализируйте результаты.

2 (2 балла)

1. Проверьте на сильную выпуклость функцию $f(x) = \frac{1}{2}\|x\|_2^2$.
2. Докажите, что если градиент липшицев с константой L, то выполнено $\|\nabla^2 f(x)\|_2 \leq$ L.
3. Проверьте являются ли следующие функции L - гладкими, если да, то определите константу Липшица:
   - $f(x) = \frac{1}{2}\|x\|_2^2$;
   - $f(x) = \|x\|_2^3$;
   - $f(x) = x^3$ на отрезке $[1, 2]$.

3 (2 балла)

1. Покажите, что градиенты, полученные на двух последовательных шагах градиентного спуска с использованием правила наискорейшего спуска, ортогональны.
2. Покажите, что для $f(x) = x^\top x$ градиентный спуск с выбором шага по правилу наискорейшего спуска сходится за одну итерацию.

4 (2 балла)

Пусть $\varphi : \mathbb{R} \to \mathbb{R}$ возрастающая и выпуклая функция, $f : \mathbb{R}^n \to \mathbb{R}$ выпуклая, следовательно $g(x) = \varphi(f(x))$ тоже выпуклая функция. Также пусть f и g дважды диференцируемы.

1. Как связаны между собой задачи минимизации f и g?
2. Сравните градиентный метод и метод Ньютона для решения задач минимизации f и g. Как связаны направления, получаемые каждым из методов? Как связаны методы в случае использования наискорейшего спуска для выбора шага? Покажите экспериментально, что полученный результат является верным.

5 (2 балла)

Рассмотрим формулу обновления матрицы H_k в методе DFP:

$$s_k = x_{k+1} - x_k, \quad y_k = \nabla f(x_{k+1}) - \nabla f(x_k),$$

$$\mathrm{H}_{k+1} = \mathrm{H}_k - \frac{\mathrm{H}_k y_k y_k^\top \mathrm{H}_k}{\langle \mathrm{H}_k y_k, y_k \rangle} + \frac{s_k s_k^\top}{\langle y_k, s_k \rangle}.$$

Пусть $\langle y_k, s_k \rangle > 0$. (Когда мы можем гарантировать данное условие?) Докажите, что если H_k является положительно определенной матрицей, то H_{k+1} тоже будет положительно определенной.

6 (3 балла)

1. Для задачи

$$\min_x \quad \ln(e^x + e^{-x})$$

   - Запустите метод Ньютона с постоянным шагом $\alpha = 1$ из точки $x_0 = 1$ и $x_0 = 1.1$. В чем разница?
   - Используйте демпфированный метод Ньютона.
   - Постройте графики сходимости для каждой постановки эксперимента и оцените время работы каждого метода.

7 (3 балла)

Функция Розенброка

$$f(x_1, x_2) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$$
$$x^0 = (1.2, 1.2), \; x^0 = (-1.2, 1)$$

1. Аналитически найдите точку минимума.
2. Исследуйте зависимость сходимости градиентного спуска от начального приближения, шаг метода выбирать по правилу Армихо.

   Постройте график зависимости длины шага от итерации для каждой начальной точки. Какой вывод Вы можете сделать?
3. Реализуйте BFGS метод и сравните его сходимость с методом Ньютона.
4. Какой метод работает лучше всего и почему?

8 (3 балла)

Предложите свой квазиньютоновский метод, реализуйте его и сравните скорость его сходимости (по времени и по итерациям) с методом Ньютона и с методом BFGS для задачи

$$\min_x \quad \log\left(\sum_{i=1}^{m} \exp(a_i^\top x + b_i)\right) + \alpha ||x||_2^2.$$

Матрицу $\mathrm{A} \in \mathbb{R}^{m \times n}$, вектор b сгенерируйте случайным образом, коэффициент регуляризации α выберите сами, рассмотрите различные значения m и n.