Домашнее задание по теме "Ускоренные методы"

1. **Квадратичная функция**

   (a) Реализуйте метод Нестерова, метод сопряженных градиентов и сравните с ранее реализованным градиентным спуском.

   (b) Задайте три квадратичные функции ($n = 2, 10, 1000$) с разными числами обусловленности ($\varkappa = 1,\ 100,\ 10000$). Всего у вас получается 9 разных матриц A. Запустите на них методы следующим образом:

   - рассмотрите разные начальные точки;
   - рассмотрите разную точность для остановки;
   - для метода Нестерова рассмотрите разные начальные условия, а именно в одном случае $x_0 = y_0 = 0_n$, в другом $x_0 = 0_n,\ y_0 = 1_n$.

   Постарайтесь ответить на вопрос: как зависит поведение методов от числа обусловленности и от начальной точки?

Замечания 1. Критерий остановки В этом задании используйте следующий критерий остановки:

$$\frac{\|\nabla f(x_k)\|_2^2}{\|\nabla f(x_0)\|_2^2} \leq \varepsilon$$

Этот критерий задает относительную точность решения благодаря нормировке на $\|\nabla f(x_0)\|_2^2$.

2. Квадратичная функция Рассмотрим матрицу $\mathrm{A} \in \mathbf{S}_{++}^n$ и вектор $\mathrm{b} \in \mathbf{R}^n$ и зададим функцию

$$f(x) = \frac{1}{2} x^\top \mathrm{A} x + \mathrm{b}^\top x.$$

Её число обусловленности зависит от матрицы A и равняется $\varkappa = \frac{\lambda_{\max}(\mathrm{A})}{\lambda_{\min}(\mathrm{A})} = \frac{L}{\mu}$, где L– константа Липшица градиента, а μ - константа сильной выпуклости. В задании вам необходимо исследовать поведение метода в зависимости от числа обусловленности.

Сгенерировать случайную квадратичную задачу с заданным $\varkappa$ можно, например, так: взять случайные числа $\lambda_1,\ \lambda_2,\ \ldots, \lambda_n \in [1, \varkappa]$, так что $\min\limits_i \lambda_i = 1$, а $\max\limits_i \lambda_i = \varkappa$, и положить $\mathrm{A} = \mathrm{diag}\,(\lambda_1, \ldots, \lambda_n)$. Элементы вектора b можно взять произвольными, они на обусловленность не влияют.