Continuous Optimization
Homework 3

Lev Khoroshansky

Раздел 1. Теория

Задача 1. Зафиксируем произвольные векторы $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, а также константу $c > 0$. Обозначим $g(t) = \nabla f(\mathbf{x} + t\,c\,\mathbf{y})$, откуда $g'(t) = [\nabla^2 f(\mathbf{x} + t\,c\,\mathbf{y})(c\,\mathbf{y})]$. По теореме о среднем, существует такое $t_c \in [0, 1]$, что

$$\nabla f(\mathbf{x} + c\,\mathbf{y}) - \nabla f(\mathbf{x}) = g(1) - g(0) = g'(t_c) = [\nabla^2 f(\mathbf{x} + t_c\,c\,\mathbf{y})(c\,\mathbf{y})]$$

Возьмём норму и применим условие о липшицевости градиента:

$$\left\|[\nabla^2 f(\mathbf{x} + t\,c\,\mathbf{y})]\,\mathbf{y}\right\|_2 \leqslant L\,\|\mathbf{y}\|_2$$

Устремим c к нулю:

$$\left\|[\nabla^2 f(\mathbf{x})]\,\mathbf{y}\right\|_2 \leqslant L\,\|\mathbf{y}\|_2$$

Заметим, что это справедливо для любого вектора $\mathbf{y}$, откуда следует, что $\|\nabla^2 f(\mathbf{x})\| \leqslant L$.

Раздел 2. Градиентный спуск

Задача 1. Воспользуемся обозначениями из семинаров и положим $p_k := -\nabla f(x_k)$. Вспомним, что

$$\alpha_k = \operatorname*{argmin}_{\alpha \geqslant 0} \varphi_k(\alpha), \text{ где } \varphi_k(\alpha) = f(x_k + \alpha p_k)$$

Иными словами, α_k минимизирует функцию $\varphi_k(\alpha)$, откуда имеем

$$\nabla \varphi_k(\alpha_k)^T p_k = 0$$

Шаг происходит следующим образом:

$$x_{k+1} := x_k + \alpha_k p_k$$

Тогда нам нужно показать, что $\langle p_{k+1}, p_k \rangle = p_{k+1}^T\,p_k = 0$:

$$\begin{aligned}(x_{k+2} - x_{k+1})^T (x_{k+1} - x_k) &= \alpha_{k+1}\alpha_k\,p_{k+1}^T\,p_k = -\alpha_{k+1}\alpha_k\,\nabla f(x_{k+1})^T p_k \\ &= -\alpha_{k+1}\alpha_k \nabla f(x_k + \alpha_k p_k)^T p_k = -\alpha_{k+1}\alpha_k\,\underbrace{\nabla \varphi_k(\alpha_k)^T p_k}_{=0} = 0\end{aligned}$$

Задачи 2, 3, 4. См. в приложенном Jupyter ноутбуке.

Раздел 3. Метод Ньютона

Задачи 1. См. в приложенном Jupyter ноутбуке.