Homework # 02

Prof. Elena Gryazina, Mazhar Ali (TA), Oleg Khamisov (TA)

(**Total**: **35 points**)

Submission Deadline: 17^{th} February (Monday) 16:00 PM

Question 01

(**2 point**) Find minimum (i.e. both point x^* and function value $f^* = f(x^*)$) of function with respect to parameters a, b, c :

$$f(x) = ax^2 + bx + c$$

Question 02:

(**1 point**) What are dimensions of gradient of a function $h(x) = f(Ax)$, constructed of function $f : \mathbb{R}^m \to \mathbb{R}$ and matrix $A \in \mathbb{R}^{m\times k}$, so the dimension $\nabla h(x)$ is?

Question 03:

(**3 point**) Prove that for a strongly convex function with parameter μ holds,

$$\frac{\mu}{2}||x - x^*||_2^2 \leq f(x) - f(x^*)$$

Question 04

(**3 point**) Derive a) gradient $\nabla f(x)$ and b) Hessian matrix $\nabla^2 f(x)$ (both in vector form) for the function $f(x) = (x, c)^2$; $x \in \mathbb{R}^n$.

Question 05

(**3 point**) Derive Hessian matrix $\nabla_x^2 f(x)$ for the function $f(x) = g(Ax+b)$, assuming differentiable $g : \mathbb{R}^m \to \mathbb{R}$, with dimensions $A \in \mathbb{R}^{m\times n}; b \in \mathbb{R}^m$; $x \in \mathbb{R}^n$.

Question 06:

(**3 point**) Solve optimal step-size problem for the quadratic function, with symmetric positive definite matrix $A \succ 0 \in \mathbb{R}^{n\times n}$, and $x, b, d \in \mathbb{R}^n$. Your goal is to find optimal γ^* for given A, b, d, x. The resulting expression must be written in terms of inner products $(\ldots, \ldots)$

$$f(\gamma) = (A(x + \gamma d), x + \gamma d) + (b, x + \gamma d) \to \min_{\gamma\in\mathbb{R}}$$

Question 7:

(**3 point**) Derive subgradient (subdifferential) for the function $f(x) = [x^2 - 1]_+, x \in \mathbb{R}$. (do not write subgradient method).

Question 8:

(**3 point**) Let $f : \mathbb{R}^n \to \mathbb{R}$ be given by $f(x) = \frac{1}{2}x^\top Qx - x^\top b$, where $b \in \mathbb{R}^n$ and Q is a real symmetric matrix positive definite $n \times n$ matrix. Suppose that we apply steepest descent (or gradient descent) method to this function with $x^0 \neq Q^{-1}b$. Show that method converges in one step that is $x^1 = Q^{-1}b$, if and only if x^0 is chosen such that $g^0 = Qx^0 - b$ is an eigenvector of Q.

Question 9:

(**5 point**) Find the minimizer of,

$$f(x, y) = x^2 + xy + 10y^2 - 22y - 5x$$

numerically by steepest descent.

1. For each iteration, record the values of x, y and f and include in a table.
2. Plot the values on a contour plot.
3. Explore different starting values, such as $(1, 10), (10, 10), (10, 1)$. Does the number of steps depend significantly on the starting guess?

Question 10

(**9 point**) Let the cost function of the unconstrained optimization problem of interest be

$$f(x) = \frac{1}{4}(x_1 - 1)^2 + \sum_{i=2}^{n} \left(2x_{i-1}^2 - x_i - 1\right)^2$$

Consider two different scenarios, first $n = 3$, and then $n = 10$. Recall that the steepest descent algorithm is,

$$x^{k+1} = x^k - \alpha^k \nabla f(x^k), \quad \alpha^k \in \mathbb{R}_{>0}$$

1. Using $x^0 = [-1.5, 1, \dots, 1]^\top$ write out the first iteration of the steepest descent algorithm and obtain the optimum value for α^0. What is the value of x^1 if you implement α^0? Verify that $f(x^1) < f(x^0)$?
2. Write a code to find the minimizer of $f(x)$ using steepest descent algorithm with starting point of $x^0 = [-1.5, 1, \dots, 1]^\top$ and using
   - $\alpha^k = \text{argmin } f(x^k - \alpha\nabla f(x^k))$.
   - constant $\alpha = 0.1$.
   - constant $\alpha = 0.5$.
   - constant $\alpha = 1.0$. (Use $||x^{k+1} - x^k|| \leq 10^{-6}$ as the stopping condition for your algorithm.)
3. How many steps it takes for the algorithm to converge for each choices of the step sizes above?
4. Use fminbnd from Matlab or an equivalent function from the programming language of your choice to solve the problem. How many steps this algorithm takes to converge?