Fall 2021 MATH 5720 Homework 7

(*) When submitting on GradeScope, please indicates pages for each question.

1. Let $f(x_1, x_2) = x_1^2 + x_2^2 - x_1x_2 - 2x_1$. Perform the exact line search at $x = (1, 1)$ in direction $d = (2, -1)$.

Hint: write $f(x + td)$ as a function of one variable t, then find the solution t of the problem

$$\min_{t>0} f(x + td).$$

2. Let $f(x_1, x_2) = x_1^2 + x_2^2 - x_1x_2 - 2x_1$. Perform the backtracking procedure at $x = (1, 1)$ in direction $d = (2, -1)$ with $s = 1$, $\alpha = 0.7$, $\beta = 0.5$. Express your answer in the following form

$$\alpha \nabla f(x)^T d = \ldots$$

t	$\frac{f(x+td)-f(x)}{t}$
1	2
0.5	
$\vdots$	

3. **(source location problem)** Suppose we are given m locations of sensors $a_1, a_2, \ldots, a_m \in \mathbb{R}^n$ and approximate distances between the sensors and an unknown "source" located at $x \in \mathbb{R}^n$:

$$d_i \approx \|x - a_i\|.$$

The problem is to find an estimate x given the locations $a_1, \ldots, a_m$ and the approximate distances $d_1, \ldots, d_m$. A natural approach is to consider the nonlinear least squares problem

$$\min_{x\in\mathbb{R}^n} \Big\{ f(x) = \sum_{i=1}^{m} (\|x - a_i\| - d_i)^2 \Big\}.$$

We will denote the set of sensors by $\mathcal{A} := \{a_1, \ldots, a_m\}$.

(a) Find the gradient $\nabla f(x)$.

(b) Show that the optimality condition $\nabla f(x) = 0$ ($x \notin \mathcal{A}$) is the same as

$$x = \frac{1}{m}\Big(\sum_{i=1}^{m} a_i + \sum_{i=1}^{m} d_i \frac{x - a_i}{\|x - a_i\|} \Big).$$

(c) Show that the corresponding fixed point method

$$x_{k+1} = \frac{1}{m}\Big(\sum_{i=1}^{m} a_i + \sum_{i=1}^{m} d_i \frac{x_k - a_i}{\|x_k - a_i\|} \Big).$$

is a gradient method, assuming that $x_k \notin \mathcal{A}$ for all $k \in \mathbb{N}$.

4. In $\mathbb{R}^3$, consider the set of points $\{\mathbf{a}_i\}$ and the weight vectors ω that are given by

$$\begin{pmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \mathbf{a}_3 \\ \mathbf{a}_4 \\ \mathbf{a}_5 \\ \mathbf{a}_6 \\ \mathbf{a}_7 \\ \mathbf{a}_8 \\ \mathbf{a}_9 \end{pmatrix} = \begin{pmatrix} -10 & 10 & 0 \\ 0 & 30 & -10 \\ 20 & 20 & 10 \\ 30 & 0 & 5 \\ 25 & -5 & 20 \\ -20 & -25 & 5 \\ 30 & 25 & -10 \\ -20 & 20 & 10 \\ 0 & -15 & 0 \end{pmatrix}, \quad \text{and} \quad \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \\ \omega_4 \\ \omega_5 \\ \omega_6 \\ \omega_7 \\ \omega_8 \\ \omega_9 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 1/2 \\ 1/2 \\ 1/3 \\ 1 \\ 3/2 \\ 1 \\ 1. \end{pmatrix}$$

Consider the Fermat Weber problem

$$\min_{\mathbf{x}\in\mathbb{R}^3} \quad f(\mathbf{x}) = \sum_{i=1}^{9} \omega_i \|\mathbf{x} - \mathbf{a}_i\|.$$

Set $\mathbf{x}_0 = (0, 0, 0)$. Find the iteration $\mathbf{x}_1$ by the Weiszfeld's method. Then use the Weiszfeld's method to solves the problem.

What criteria can be used to verify if a solution is approximately attained? You may use Julia/Matlab to solve the problem.

Extra Problems (not graded)

E.1. Let $a = (1, 1)$, $b = (2, -1)$, $c = (-1, 0)$. Define

$$f : \mathbb{R}^2 \to \mathbb{R} : x \mapsto \|x - a\| + \|x - b\| + \|x - c\|$$

(a) Check if $d = (2, 0)$ is a descent direction of f at $x = c$.

(b) Perform the backtracking procedure at $x = c$ in direction $d = (2, 0)$ with parameters $s = 1$, $\alpha = 0.9$, and $\beta = 0.5$.

E.2. Let $f \in C_L^{1,1}(\mathbb{R}^n)$ (i.e., f is continuously differentiable and the gradient ∇f is Lipschitz continuous with modulus L) and let $(x_k)_{k\in\mathbb{N}}$ be a sequence generated by the gradient method with a constant stepsize $t_k = \frac{1}{L}$. Assume that $x_k \to x^*$ and that $\nabla f(x_k) \neq 0$ for all $k \geq 0$. Prove that x^* is not a local maximum point.

Hint: Use Proposition on sufficient decrease of gradient methods to show that $f(x_k)$ is a strictly decreasing sequence, i.e.,

$$f(x^1) > f(x^2) > \cdots > f(x^k) > \cdots$$

which contradicts the fact that x^* is a local maximizer.

E.3. Give an example of a function $f \in C_L^{1,1}(\mathbb{R})$ and a starting point $x_0 \in \mathbb{R}$ such that the problem

$$\min f(x)$$

has an optimal solution but the gradient method with constant stepsize $t = \frac{2}{L}$ diverges.