

2nd Report of Mathematical Optimization: Theory and Algorithms

Tokyo Institute of Technology
Department of Mathematical and Computing Science
MITUHIRO FUKUDA

Due to August 15 (Saturday), 2020, 11:50PM

Submit the report to the OCW-i. It only accepts pdf format files.

This report will correspond to 70% of the final grade of this course.

Answer all the questions below in English.

1. (15 points) Show that the log-sum-exp function, $f(\mathbf{x}) = \ln \left(\sum_{i=1}^n e^{x_i} \right)$, is a convex function on \mathbb{R}^n .
2. (24 points) Given a convex set $S \subseteq \mathbb{R}^n$ and an arbitrarily norm $\|\cdot\|$ in \mathbb{R}^n , define the distance of a point $\mathbf{x} \in \mathbb{R}^n$ to the set S as

$$\text{dist}(\mathbf{x}, S) := \inf_{\mathbf{y} \in S} \|\mathbf{x} - \mathbf{y}\|.$$

Show that the distance function $\text{dist}(\mathbf{x}, S)$ is convex on \mathbf{x} . Notice that the convex set S does not need to be closed.

3. (20 points) Let $Q \subseteq \mathbb{R}^n$ a convex subset of \mathbb{R}^n . A function $f: Q \rightarrow \mathbb{R}$ is called *quasiconvex* for $Q \subseteq \mathbb{R}^n$ if ~~its domain Q is convex and~~ all of its level sets $L_\lambda := \{\mathbf{x} \in Q \mid f(\mathbf{x}) \leq \lambda\}$ are convex for $\forall \lambda \in \mathbb{R}$. Show that a differentiable function $f(\mathbf{x})$ is quasiconvex if and only if for any $\mathbf{x}, \mathbf{y} \in Q$ such that

$$f(\mathbf{y}) \leq f(\mathbf{x}) \Rightarrow \langle \nabla f(\mathbf{x}), \mathbf{y} - \mathbf{x} \rangle \leq 0.$$

4. (16 points) Answer “Yes” or “No” to the following questions. Justify your answers.
 - a. Is the function $f(\mathbf{x})$ of **problem (1) below** differentiable at all points of interior of Q ?
 - b. Are strongly convex functions also strictly convex functions?
 - c. Are functions with Lipschitz continuous gradients also Lipschitz continuous functions?
 - d. Is the Nesterov’s optimal gradient method the fastest among all gradient-based methods for *any* function of the class $\mathcal{F}_L^1(\mathbb{R}^n)$ or $\mathcal{S}_{\mu, L}^{1,1}(\mathbb{R}^n)$?
5. (25 points) Given an non-empty closed convex subset Q of \mathbb{R}^n and $f_i \in \mathcal{S}_{\mu, L}^1(Q)$ with $L > \mu > 0$ for $(i = 1, 2, \dots, m)$, consider the optimization problem

$$\begin{cases} \text{minimize} & f(\mathbf{x}) \equiv \max_{i=1,2,\dots,m} f_i(\mathbf{x}) \\ \text{subject to} & \mathbf{x} \in Q. \end{cases} \quad (1)$$

Given $\mathbf{x}_k \in \mathbb{R}^n$, we define

$$\begin{aligned} \mathbf{x}_f(\mathbf{x}_k; L) &:= \arg \min_{\mathbf{y} \in Q} \left\{ \max_{i=1,2,\dots,m} [f_i(\mathbf{x}_k) + \langle \nabla f_i(\mathbf{x}_k), \mathbf{y} - \mathbf{x}_k \rangle] + \frac{L}{2} \|\mathbf{y} - \mathbf{x}_k\|_2^2 \right\}, \\ \mathbf{g}_f(\mathbf{x}_k; L) &:= L(\mathbf{x}_k - \mathbf{x}_f(\mathbf{x}_k; L)). \end{aligned}$$

We have shown during the lectures that for any $\mathbf{x} \in Q$,

$$\max_{i=1,2,\dots,m} f_i(\mathbf{x}) \geq \max_{i=1,2,\dots,m} f_i(\mathbf{x}_f(\mathbf{x}_k; L)) + \langle \mathbf{g}_f(\mathbf{x}_k; L), \mathbf{x} - \mathbf{x}_k \rangle + \frac{1}{2L} \|\mathbf{g}_f(\mathbf{x}_k; L)\|_2^2 + \frac{\mu}{2} \|\mathbf{x} - \mathbf{x}_k\|_2^2.$$

a. If $\mathbf{x}^* \in Q$ is the optimal solution of (1), show that for $\mathbf{x}_k \in \mathbb{R}^n$, we have:

$$\langle \mathbf{g}_f(\mathbf{x}_k; L), \mathbf{x}_k - \mathbf{x}^* \rangle \geq \frac{1}{2L} \|\mathbf{g}_f(\mathbf{x}_k; L)\|_2^2 + \frac{\mu}{2} \|\mathbf{x}^* - \mathbf{x}_k\|_2^2.$$

b. Given $\mathbf{x}_0 \in Q$, consider the steepest descent update $\mathbf{x}_k := \mathbf{x}_{k-1} - \frac{1}{L} \mathbf{g}_f(\mathbf{x}_{k-1}; L)$. Show that

$$\|\mathbf{x}_k - \mathbf{x}^*\|_2^2 \leq \left(1 - \frac{\mu}{L}\right)^k \|\mathbf{x}_0 - \mathbf{x}^*\|_2^2.$$