Homework 1

Tuesday, March 16, 2021 3:07 PM

Problem 1: (a) Let $f: \mathbb{R}^n \to \mathbb{R}$ be a convex differentiable function. Prove that $\forall\, x, y \in \mathbb{R}^n$, $f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle$

(b) Let $f: \mathbb{R}^n \to \mathbb{R}$ be a second-order differentiable function that is L-smooth i.e. $\|H_f(x)\| \leq L \;\; \forall\, x \in \mathbb{R}^n$. Prove that $\forall\, x, y \in \mathbb{R}^n$

$$f(y) \leq f(x) + \langle \nabla f(x), y - x \rangle + \frac{L}{2} \|y - x\|_2^2$$

(c) Let $f: \mathbb{R}^n \to \mathbb{R}$ be a second-order differentiable function that is μ-strongly convex i.e. $H_f(x) \succeq \mu I \;\; \forall\, x \in \mathbb{R}^n$. Prove that $\forall\, x, y \in \mathbb{R}^n$

$$f(y) \geq f(x) + \langle \nabla f(x), y - x \rangle + \frac{\mu}{2} \|y - x\|_2^2$$

Problem 2: Let $f: \mathbb{R}^n \to \mathbb{R}$ be L-smooth & μ-strongly convex. Prove that gradient descent with an appropriate choice of the step size η in $T = O\left(\frac{L}{\mu} \log\left(\frac{f(x_0) - f(x^*)}{\varepsilon}\right)\right)$ iterations will output x_T s.t.

$$f(x_T) - f(x^*) \leq \varepsilon.$$

(Try to bound the multiplicative decrease in $f(x_t) - f(x^*)$ in each iteration)

Problem 3 [Coordinate descent, exercise 6.7 in Nisheeth's book]:

Let $f: \mathbb{R}^n \to \mathbb{R}$ be a convex twice differentiable function s.t. $\frac{\partial^2 f}{\partial x_i^2} \leq \beta_i \;\; \forall\, i, x$.

Let $\beta = \sum_{i=1}^{n} \beta_i$. Let $x \in \mathbb{R}^n$ & consider the update rule:

$$x' = x - \frac{1}{\beta_i} \frac{\partial f}{\partial x_i} e_i$$

where i is chosen at random from $[n]$ according to the distribution $\left(\frac{\beta_1}{\beta}, \ldots, \frac{\beta_n}{\beta}\right)$.

(and $e_i = (0, 0, \ldots, \overset{i}{1}, 0 \ldots 0)$ is the ith standard unit vector)

Prove that $$\mathbb{E}[f(x')] \leq f(x) - \frac{1}{2\beta} \|\nabla f(x)\|_2^2$$

(This can be further used to design a randomized gradient descent like algorithm but gets a bit messier).

Problem 4 [Solving LPs using MWU]:

Consider the LP feasibility problem where one is supposed to find a feasible point x to the system of linear inequalities.

$$\langle a_i, x\rangle \leq b_i \quad \forall\, i = 1, \dots, m.$$

(We will use MWU algorithm to design an algorithm that outputs $\tilde{x}$ that is ε-approximately feasible i.e. $\langle a_i, \tilde{x}\rangle \leq b_i + \varepsilon \;\forall\, i$ whenever the above LP is feasible.)

Assume the following oracle: given $p \in \Delta_m$, it outputs an x s.t.

$$\sum_i p_i \langle a_i, x\rangle \leq \sum_i p_i b_i$$

(if there exists one). Furthermore assume that the returned x always satisfies

$$\max_i |\langle a_i, x\rangle - b_i| \leq G.$$

Design an algorithm to find an ε-approximately feasible $\tilde{x}$ using only $O\left(G^2 \log(m)/\varepsilon^2\right)$ calls to the oracle.