CPSC 368 / CPSC 516

Assignment due by Tuesday, 28 February 2023, 9 a.m.

Instructions:

- These problem sets are meant to be worked on alone – no groups.
- Each answer should be supported by a rigorous mathematical proof.
- You are strongly encouraged to type the solution in LaTeX.
- You are strongly encouraged to discover the solutions by yourself as that will help you understand the techniques better and will help you do well in the exam. In case you take help of any kind it has to be clearly acknowledged for each problem.

Problems:

P.1. In this problem we analyze a gradient descent algorithm for minimizing a twice-differentiable convex function $f : \mathbb{R}^n \to \mathbb{R}$, which satisfies for every $x \in \mathbb{R}^n$, $mI \preceq \nabla^2 f(x) \preceq MI$ for some $0 < m \leq M$.

The algorithm starts with some $x_0 \in \mathbb{R}^n$ and at every step $t = 0, 1, 2, \ldots$ it chooses the next point

$$x_{t+1} := x_t - \alpha_t \nabla f(x_t),$$

where α_t is chosen to minimize the value $f(x_t - \alpha \nabla f(x_t))$ over all $\alpha \in \mathbb{R}$ while fixing x_t. Let $y^\star := \min\{f(x) : x \in \mathbb{R}^n\}$.

(a) Prove that

$$\forall x, y \in \mathbb{R}^n, \quad \frac{m}{2} \|y - x\|_2^2 \leq f(y) - f(x) + \langle \nabla f(x), x - y \rangle \leq \frac{M}{2} \|y - x\|_2^2 .$$

(b) Prove that

$$\forall x \in \mathbb{R}^n, \quad f(x) - \frac{1}{2m} \|\nabla f(x)\|_2^2 \leq y^\star \leq f(x) - \frac{1}{2M} \|\nabla f(x)\|_2^2 .$$

(c) Prove that for every $t = 0, 1, 2 \ldots$

$$f(x_{t+1}) \leq f(x_t) - \frac{1}{2M} \|\nabla f(x_t)\|_2^2 .$$

(d) Prove that for every $t = 0, 1, 2 \ldots$

$$f(x_t) - y^\star \leq \left(1 - \frac{m}{M}\right)^t (f(x_0) - y^\star) .$$

What is the number of iterations t required to reach $f(x_t) - y^\star \leq \varepsilon$?

(e) Consider a linear system $Ax = b$, where $b \in \mathbb{R}^n$ is a vector and $A \in \mathbb{R}^{n \times n}$ is a symmetric positive definite matrix such that $\frac{\lambda_n(A)}{\lambda_1(A)} \leq \kappa$ (where $\lambda_1(A)$ and $\lambda_n(A)$ are the smallest and the largest eigenvalues of A respectively). Use the above framework to design an algorithm for approximately solving the system $Ax = b$ with logarithmic dependency on the error $\varepsilon > 0$ and polynomial dependency on κ. What is the running time?

P.2. Let $G = (V, E)$ be an undirected graph with n vertices and m edges. Let $B \in \mathbb{R}^{n \times m}$ be the vertex-edge incidence matrix of G. Assume that G is connected and let $\Pi := B^\top (BB^\top)^+ B$. Prove that, given a vector $g \in \mathbb{R}^m$, if we let x_g denote the projection of g on the subspace $K := \{x \in \mathbb{R}^m : Bx = 0\}$, then it holds that

$$x_g = g - \Pi g.$$

Recall that the projection of a point $h \in \mathbb{R}^n$ on a closed convex nonempty set $K \subseteq \mathbb{R}^n$ is defined as the unique point $x_h \in K$ that minimizes the Euclidean distance to h:

$$x_h := \operatorname*{argmin}_{y \in K} \|y - h\|_2.$$