CS395T: Continuous Algorithms
Homework I

Kevin Tian

Due date: February 1, 2024, start of class (3:30 PM).

Please list all collaborators on the first page of your solutions. Unless we have discussed and I have specified otherwise, homework is not accepted if it is not turned in by hand at the start of class, or turned in electronically on Canvas by then. Send me an email to discuss any exceptions.

1 Problem 1

Let $f : \mathbb{R}^d \to \mathbb{R}$ be a convex function. For each of the following functions F derived from f, either prove F is always convex, or give a counterexample showing F may not always be convex.

(i) Let $g(x) := \mathbf{A}x + b$ for $\mathbf{A} \in \mathbb{R}^{d\times n}$, $b \in \mathbb{R}^d$,[1] and let $F(x) := f(g(x))$ for $x \in \mathbb{R}^n$.

(ii) Let $S \subseteq [d]$ and let F be a partial minimization of f over $[d] \setminus S$, i.e. for $x \in \mathbb{R}^S$, let $F(x) := \min_{y\in\mathbb{R}^{[d]\setminus S}} f(x, y)$ where $(x, y) \in \mathbb{R}^d$ is associated with coordinates of $[d]$ in the natural way. Assume that the minimum is always achieved for any $x \in \mathbb{R}^S$.

(iii) Let $g : \mathbb{R} \to \mathbb{R}$ be strictly convex, and let $F(x) = g(f(x))$.

(iv) Let $g : \mathbb{R} \to \mathbb{R}$ be convex and monotone nondecreasing, and let $F(x) = g(f(x))$.

2 Problem 2

In each of the following cases, given access to an algorithm $\mathcal{A}$ achieving a purported runtime on a class of functions $\mathcal{F}$, design another optimization algorithm $\mathcal{A}'$ which calls $\mathcal{A}$, and can optimize functions in $\mathcal{F}$ to ϵ additive error, in τ time for an arbitrarily small $\tau > 0$, for any $\epsilon > 0$.[2]

(i) $\mathcal{A}$ can optimize L-smooth functions over $\mathbb{R}^d$ to ϵ additive error in time $\frac{L^2}{\epsilon}$. Here, $\mathcal{F}$ is the class of L-smooth functions over $\mathbb{R}^d$, for some finite $L > 0$.

(ii) $\mathcal{A}$ can optimize L-Lipschitz functions over $\mathbb{B}(R)$ to ϵ additive error in time $\frac{LR^2}{\epsilon}$. Here, $\mathcal{F}$ is the class of L-Lipschitz functions supported on $\mathbb{B}(R)$, for some finite $L, R > 0$.

3 Problem 3

Prove the following statements about regularity in the ℓ_∞ norm.

(i) If $f : \mathbb{R}^d \to \mathbb{R}$ is twice-differentiable and 1-smooth in ℓ_∞, $\mathrm{Tr}(\nabla^2 f(x)) \le 1$ for all $x \in \mathbb{R}^d$.

(ii) If $f : \mathbb{R}^d \to \mathbb{R}$ is 1-strongly convex in ℓ_∞, then

$$\max_{x\in\mathcal{X}} f(x) - \min_{x\in\mathcal{X}} f(x) \ge \frac{d}{2}, \text{ for } \mathcal{X} := [-1, 1]^d.$$

[1] In other words, let g be an *affine* function.

[2] Clearly, this means such a purported runtime for $\mathcal{A}$ is impossible. In all cases, the culprit is non-*scale-invariant* runtimes; this problem is a lesson on common sanity checks which can be applied to runtimes claimed in papers.

4 Problem 4

Let $f : \mathbb{R}^d \to \mathbb{R}$ be a continuous and strictly convex function, let $r \geq 0$, and let $\mathcal{O}$ be an oracle which takes as input $x \in \mathbb{R}^d$ and returns the minimizer of f in $\mathbb{B}(x, r)$, i.e.

$$\mathcal{O}(x) = \text{argmin}_{x' \in \mathbb{B}(x,r)} f(x').$$

Suppose that $x^\star := \text{argmin}_{x \in \mathbb{R}^d} f(x)$ exists. Consider iterating the update, from some $x_0 \in \mathbb{R}^d$,

$$x_{t+1} \leftarrow \mathcal{O}(x_t), \text{ for all } 0 \leq t < T.$$

(i) Prove that for all $0 \leq t < T$, $\|x_{t+1} - x^\star\|_2 \leq \|x_t - x^\star\|_2$.

(ii) Let $R := \|x_0 - x^\star\|_2$. Prove that

$$f(x_T) - f(x^\star) \leq \left(1 - \frac{r}{R}\right)^T (f(x_0) - f(x^\star)).$$

5 Problem 5

Let $f : \mathbb{R}^d \to \mathbb{R}$ be convex and L-smooth for $L > 0$,[3] and let $x^\star \in \text{argmin}_{x \in \mathbb{R}^d} f(x)$. Consider the gradient flow ODE discussed in class, i.e. let $\frac{\mathrm{d}}{\mathrm{d}t} x_t = -\nabla f(x_t)$ for all $t \geq 0$.

(i) Prove that, for $\Phi(t) := \frac{1}{2} \|x_t - x^\star\|_2^2$, we have $\frac{\mathrm{d}}{\mathrm{d}t} \Phi(t) \leq 0$ for all $t \geq 0$.

(ii) Prove that, for $T \geq 0$ and $\bar{x} = \frac{1}{T} \int_0^T x_t \mathrm{d}t$, we have

$$f(\bar{x}) - f(x^\star) \leq \frac{\|x_0 - x^\star\|_2^2}{2T}.$$

[3] This is enough to conclude the ODE has a unique solution, by the Picard-Lindelöf theorem.