CS7545, Spring 2023: Machine Learning Theory - Homework #4

Jacob Abernethy, Zihao Hu, and Guanghui Wang Due: Tuesday, April 25 at 11:59 p.m.

Homework Policy: *The due is at Tuesday 4/25, but everyone gets a free extension to 5/2.* Working in groups is fine, but *every student* must submit their own writeup. Please write the members of your group on your solutions. There is no strict limit to the size of the group but we may find it a bit suspicious if there are more than 4 to a team. Questions labelled with **(Challenge)** are not strictly required, but you'll get some participation credit if you have something interesting to add, even if it's only a partial answer.

1) **Generalized Minimax Theorem.** Let $X \subset \mathbb{R}^n$ and $Y \subset \mathbb{R}^m$ be convex compact sets. Let $f : X \times Y \to \mathbb{R}$ be some differentiable function with bounded gradients, where $f(\cdot, \mathbf{y})$ is convex in its first argument for all fixed $\mathbf{y}$, and $f(\mathbf{x}, \cdot)$ is concave in its second argument for all fixed $\mathbf{x}$. An ϵ-optimal point $(\mathbf{x}^*, \mathbf{y}^*)$ satisfies

$$\sup_{\mathbf{y} \in Y} f(\mathbf{x}^*, \mathbf{y}) - \inf_{\mathbf{x} \in X} f(\mathbf{x}, \mathbf{y}^*) \leq \epsilon.$$

- Prove that

$$\inf_{\mathbf{x} \in X} \sup_{\mathbf{y} \in Y} f(\mathbf{x}, \mathbf{y}) = \sup_{\mathbf{y} \in Y} \inf_{\mathbf{x} \in X} f(\mathbf{x}, \mathbf{y}).$$

- Give an efficient algorithm which finds a $O(T^{-1/2})$-optimal $(\mathbf{x}^*, \mathbf{y}^*)$ in T iterations.
- Assume $f(\cdot, \cdot)$ is strongly convex in its first argument and strongly concave in its second argument. Find a $O(\frac{\log T}{T})$-optimal solution after T iterations.