Privacy in Statistics and Machine Learning — Spring 2021
In-class Exercises for Lecture 10 (ERM)
March 4/5, 2021

Adam Smith and Jonathan Ullman

Problems with marked with an asterisk (*) are more challenging or open-ended.

1. Consider the Euclidean median problem in d-dimensional space: given $x_1, x_2, ..., x_n \in \mathbb{R}^d$, defined by $\ell(w; x) = \|w - x\|_2$. (If the norm were squared, the minimizer would be the mean, but the square root changes its behavior.)

   What is the gradient of the individual loss function? What conditions if any do we need on C (the set of acceptable w) and $\mathcal{U}$ (the set of allowed x) for the loss to be Lipschitz?

   Is the loss convex?

2. Show that if the individual loss ℓ is G-Lipschitz on a set C, then so is the overall loss L from equation (1) in the notes (assuming no additional regularizer Λ).

3. Show that ERM for G-Lipschitz losses on a set of diameter R can be reduced by rescaling to ERM for 1-Lipschitz losses for a set of diameter 1. Show that the excess risk increases by a factor of GR due to the rescaling. (*Reminder:* Re-scaling the input domain changes G as well as R, since it compresses distances.)

4. Consider the following wacky idea: given a G-Lipschitz loss function $\ell : C \times \mathcal{U} \to \mathbb{R}$, you decide to optimize $L(w; \mathbf{x})$ differentially privately by running the exponential mechanism with score

$$q(w; \mathbf{x}) = -\|\nabla L(w; \mathbf{x})\|_2 .$$

   (a) Show that this score function is $\frac{G}{n}$-sensitive, so sampling from $p(w) \propto \exp(-\frac{\varepsilon n}{2G} q(w; \mathbf{x}))$ is $(\varepsilon, 0)$-DP.

   (b) Suppose you run this algorithm to optimize the median's objective function given by $\ell(w; x) = |w - x|$, with w and the x_i's restricted to the interval $C = \mathcal{U} = [0, 1]$. What algorithm from class or homework do you recover?

   (c) Suppose you run this algorithm to optimize the mean's objective function given by $\ell(w; x) = (w - x)^2$, with w and the x_i's restricted to the interval $C = \mathcal{U} = [0, 1]$. What algorithm from class or homework do you recover?

(Continued on next page.)

Exercises about convex sets and functions

5. Which of these operations, when applied to a set of convex functions, always produces a convex function?
   - Sum
   - Min
   - Max
   - Median

6. Show that the set of minima of a convex function is a convex set.

7. Show that for any convex set $C \subseteq \mathbb{R}^d$, the function $f(x) = \min_{w \in C} \|x - w\|_2$ is a convex function on all of $\mathbb{R}^d$.

8. Show that for any convex set $C \subseteq \mathbb{R}^d$, the *projection* function $\Pi_C(x) = \arg\min_{w \in C} \|x - w\|_2$ is (a) a well-defined function from $\mathbb{R}^d$ to $\mathbb{R}^d$ (that is, the minimizer is unique), (b) 1-Lipschitz, meaning that for all $x, y \in \mathbb{R}^d$, we have
$$\|\Pi_C(x) - \Pi_C(y)\|_2 \leq \|x - y\|_2$$

9. Prove Jensen's inequality (Exercise 3.4 in the notes). [*Hint:* Let $\mu = \mathbb{E}(X)$ and let $g_\mu(\cdot)$ be an affine lower bound to f such that $f(\mu) = g_\mu(\mu)$. What is $\mathbb{E}\left(g_\mu(X)\right)$?]

10. The level set of a function f at a is the set $\{w \in C : f(w) \leq a\}$. Show that if f is convex, then all of its level sets are convex. Show via a counterexample that the converse is false.

11. **(*)** Suppose $f : \mathbb{R} \to \mathbb{R}$ is convex.

    (a) Show that the subgradients of f are monotone, namely: for every $x_1, x_2 \in \mathbb{R}$ such that $x_1 < x_2$, if $y_1 \in \partial f(x_1)$ and $y_2 \in \partial f(x_2)$, then $y_1 \leq y_2$. (It might be easier to first prove this for f that is differentiable.)

    (b) Show that if f is convex and 1-Lipschitz on a finite interval (say $[-1, 1]$), then it can be written a constant plus a convex combination of absolute value functions. Specifically, show that there is a constant a and a distribution P on $[-1, 1]$ such that for all x, $f(x) = a + \mathop{\mathbb{E}}_{Y \sim P}(|x - Y|)$.