Homework 2
CSCI 699: Privacy-Preserving Machine Learning

Instructor: Sai Praneeth Karimireddy
Due: Oct 27, 2024

Instructions: Answer the following questions clearly and concisely. Your proofs should be written in a manner such that if you show it to a fellow classmate, they will be able to follow along and understand it completely. Points for each question are indicated. Type your answers in Latex and submit the pdf.

Questions

Question 1: Private Mean Estimation

Suppose we are given $\{x_1, \ldots, x_n\}$ which i.i.d. random vectors $x_i \in \mathbb{R}^d$ which satisfy the following moment bounds for all $i \in [n]$

$$E[x_i] = \mu\,, \quad E\|x_i\|_2^2 \leq \sigma^2 \quad \text{, and} \quad E\|x_i\|_2^p \leq \alpha^p \text{ for some } p > 2\,.$$

Then consider the following estimator

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} y_i \text{ where } y_i = \frac{\tau}{\max(\|x_i\|_2, \tau)} x_i\,.$$

Assume the following fact is known about the bias introduced due to clipping:

$$E\|x_i\|_2^p \leq \alpha^p \Rightarrow \|E[y_i] - \mu\|_2 \leq \frac{\alpha^p}{\tau^{p-1}}\,.$$

Using the above stated fact, construct an (ε, δ)-DP algorithm which outputs $\mu_{private}$ with an error satisfying the error below (4 points)

$$E\|\mu_{private} - \mu\|_2^2 \leq O\left(\frac{\sigma^2}{n} + \alpha^2 \left(\frac{d}{n^2\varepsilon^2}\right)^{1-\frac{1}{p}} \log(1/\delta)^{1-\frac{1}{p}}\right)$$

Question 2: Private Learning (practical)

See this notebook here and answer questions / fill in missing code (6 points): `https://colab.research.google.com/drive/1L8U5cv3Cc-JI8vMTjPeC9RzStzjRsHNi?usp=sharing`.

Note: For question 1, you may respond in latex or in the colab notebook. Please attach a jupyter notebook when submitting your solution on Brightspace. Do not just submit a colab link.