In the Name of God

Data Privacy (25100)

Problem Set 01

Fall 2023

Department of Electrical Engineering

Sharif University of Technology

Instructor: Dr. M.H. Yassaee

Soft Dedline: 5 Dey 1402 - 23:55
Hard Dedline: 9 Dey 1402 - 23:55

- Delivering Assignment with LaTeX has 15% bonus mark.

1 Differential Privacy and Reconstruction Attacks

Suppose A is an ε-differentially private algorithm that takes input $\mathbf{x}=(x_1, x_2, \ldots, x_n) \in \{0,1\}^n$ (so each person's secret information is just one bit). Consider an algorithm B that attempts to reconstruct the input from A 's output: on input $A(\mathbf{x})$, it outputs a guess $\tilde{\mathbf{x}}$. Show that, for every algorithm B : if $\mathbf{x}$ is selected uniformly at random from $\{0,1\}^n$, and the algorithm B has access only to the output of A (nothing else), then

$$\underset{\substack{\mathbf{x} \in r_r \{0,1\}^n \\ \mathbf{x}=B(A(\mathrm{x}))}}{\mathbb{E}} (\#\, \text{errors}(\tilde{\mathbf{x}}, \mathbf{x})) \geq \frac{n}{e^{\varepsilon}+1}$$

Here, # errors (y, x) denotes the number of positions in which two vectors disagree (also called the Hamming distance). [1]

Hints: Use linearity of expectation. The number of errors can be written as a sum of random variables E_i (for $i=1$ to n), where

$$E_i = \begin{cases} 1 & \text{if } \tilde{\mathbf{x}}_i = x_i \\ 0 & \text{otherwise} \end{cases}$$

What can you say about the conditional distribution of x_i given a particular output $A(\mathrm{x})=a$? How big or small can $\operatorname{Pr}(x_i=1 \mid A(\mathrm{x})=a)$ be? Given that, what is the largest possible probability that $E_i=1$? What does that tell you about E_i 's expected value? It might be helpful to think about what happens when A is the randomized response mechanism, though your final proof should apply to any ε-DP algorithm.

[1] In other words: when ε is small, differentially private algorithms do not allow for non-trivial reconstruction attacks. Even vith no output at all, an attacker can always guess about $\frac{n}{2}$ of the bits of x in expectation (for example, by guessing the all-zeros tring). The result above says that a attack based on differentially private output cannot do much better in expectation.

2 Approximate Differential Privacy

Prove that approximate differential privacy satisfies the following properties:

1. (ε, δ)-differential privacy is closed under post-processing.

2. (ε, δ)-differential privacy satisfies (adaptive) composition. Running one mechanism satisfying $(\varepsilon_1, \delta_1)$ DP followed by another mechanism satisfying $(\varepsilon_2, \delta_2) - DP$ satisfies $(\varepsilon_1 + \varepsilon_2, \delta_1 + \delta_2) - DP$.

Note that, similar to pure DP, approximate DP satisfies a composition property where composting T mechanisms, each with (ε, δ)-DP gives us a combined mechanism that is $(\varepsilon T, \delta T)$-DP. However, unlike pure DP, this composition bound can actually be improved considerably, and we can prove a bound more like $(\varepsilon\sqrt{T}, \delta T)$-DP, although the exact parameters are a bit weaker. This "strong composition" property is one of the most useful things about the approximate version of DP, and the next lecture is devoted to exploring this phenomenon in more detail.

3 Differentially Private LSR Problem

Consider the least squares regression problem, in which the dataset is $X = ((x_1, y_1), \ldots, (x_n, y_n))$, and each $x_i \in \mathcal{X} = [0,1]^d$, and $y_i \in [0,1]$, we consider functions of the type $f_\theta(x) = \theta^\top x$ for $\theta \in B_2(0, R)^d$, and the loss is given by

$$l\left(f_\theta(x), y\right) = \left(y - f_\theta(x)\right)^2.$$

The goal is to minimize the empirical loss $L(\theta, X) = \frac{1}{n}\sum_{i=1}^{n} l\left(f_\theta\left(x_i\right), y_i\right)$ over all $\theta \in B_2(0, R)$. Suppose we use the private stochastic gradient descent algorithm from the lecture notes to approximately minimize this loss. I.e. suppose we use the algorithm from the notes to give an (ε, δ)-deferentially private algorithm which outputs $\theta^{\text{priv}} = \frac{1}{T}\sum_{t=0}^{T-1}\theta^t$ such that, for all large enough datasets X,

$$\mathbb{E}L\left(\theta^{\text{priv}}, X\right) - \min_{\theta \in B_2(0,R)^d} L(\theta, X) \leq \alpha$$

What is the value of n_0 for which this inequality holds for all datasets X of size $n \geq n_0$? Justify your answer.
Hint: You can use the Cauchy-Schwarz inequality: for any two vectors $u, v \in \mathbb{R}^d$, $\left|u^\top v\right| \leq \|u\|_2 \|v\|_2$.

4 An Unknown Private Algorithm!

Suppose you are given a private dataset $X \in \mathcal{X}^n$, where $\mathcal{X} = \{1, \ldots, N\}$. I.e. the dataset X consists of n integers $x_1, \ldots, x_n$ between 1 and N. Describe an ε-differentially private algorithm, based on the exponential mechanism, which outputs a number $y \in \mathcal{X}$ such that if $n \geq \frac{C_1}{\varepsilon}\ln(|N|/\beta) + C_2$, then

$$\mathbb{P}\left(\min_{i=1}^{n} x_i \leq y \leq \max_{i=1}^{n} x_i\right) \geq 1 - \beta.$$

Above C_1 and C_2 are constants independent of n, N, β, and ε. Justify why your algorithm is ε-differentially private, and why it satisfies the property above. Specify the constants C_1, C_2 in your answer.