Homework2 DP

April 2020

1 Basics of DP (25 pts)

Given a database $x = (1, 3, 2, 5, 4)$ and the domain of databases is $\{1, 2, 3, 4, 5\}^5$, design ϵ-differentially private mechanisms corresponding to the following queries where $\epsilon = 0.1$.

1.1

$q(x) = \Sigma_{i=1}^{5} x_i$.

1.2

$q(x) = \max_{i \in [1,5]} x_i$.

2 Sensitivity Analysis (25 pts)

2.1

Given a function $F(x, t)$, $x \in \mathcal{X}$, $t \in \mathbb{N}$, define $\Delta_t = \max_{x \simeq y} \|F(x, t) - F(y, t)\|$ ($x \simeq y$ denotes that x and y are neighboring inputs). If $\Delta_0 = 0$ and

$$\Delta_t \leq \begin{cases} \Delta_{t-1} + 2L\eta_t, & if\ t = 1 + jm,\ j \in \mathbb{N} \\ \Delta_{t-1}, & otherwise, \end{cases} \tag{1}$$

where L, m, η_t are given parameters. Show that $\Delta_T \leq 2L\Sigma_{j=0}^{k-1}\eta_{1+jm}$ where $T = km$.

2.2

If (1) is replaced by

$$\Delta_t \leq \begin{cases} (1 - n\gamma)\Delta_{t-1} + 2L\eta, & if\ t = 1 + jm,\ j \in \mathbb{N} \\ (1 - n\gamma)\Delta_{t-1}, & otherwise, \end{cases} \tag{2}$$

where L, γ, m, η are given parameters. Show that $\Delta_T \leq 2L\eta\Sigma_{j=0}^{k-1}(1-n\gamma)^{(k-j)m-1}$ where $T = km$.

3 Composition (25 pts)

The algorithm of differentially private stochastic gradient decent is presented in Fig. 1. In this question, assume that two inputs X and Y are neighbouring inputs if X can be obtained from Y by removing or adding one element (e.g., $X = (x_1, \ldots, x_N)$ and $Y = (x_1, \ldots, x_{N-1})$ are neighbouring inputs). Answer the following questions.

3.1

Prove that each update in Algorithm 1 (i.e., lines 6-15) is (ϵ, δ)-DP if $\sigma^2 \geq 2\ln(1.25/\delta)/\epsilon^2$ for $\epsilon, \delta \in (0, 1)$ and $q \triangleq L/N = 1$.

3.2

Given $(\epsilon, \delta) = (1.25, 10^{-5})$, $q = 1$ and $T = 10000$, calculate σ in Algorithm 1 with the composition theorem (Theorem 3.16 in the textbook) such that Algorithm is (ϵ, δ)-differentially private.

3.3

Calculate σ in Algorithm 1 with the advanced composition theorem (Theorem 3.20 in the textbook) under the setting above (choose $\delta^{'} = \delta$ while using Theorem 3.20).

Algorithm 1 Differentially private SGD (Outline)

Input: Examples $\{x_1, \ldots, x_N\}$, loss function $\mathcal{L}(\theta) = \frac{1}{N}\sum_i \mathcal{L}(\theta, x_i)$. Parameters: learning rate η_t, noise scale σ, group size L, gradient norm bound C.

Initialize θ_0 randomly

for $t \in [T]$ **do**

Take a random sample L_t with sampling probability L/N

Compute gradient

For each $i \in L_t$, compute $\mathbf{g}_t(x_i) \leftarrow \nabla_{\theta_t}\mathcal{L}(\theta_t, x_i)$

Clip gradient

$\bar{\mathbf{g}}_t(x_i) \leftarrow \mathbf{g}_t(x_i) / \max\left(1, \frac{\|\mathbf{g}_t(x_i)\|_2}{C}\right)$

Add noise

$\tilde{\mathbf{g}}_t \leftarrow \frac{1}{L}\left(\sum_i \bar{\mathbf{g}}_t(x_i) + \mathcal{N}(0, \sigma^2 C^2 \mathbf{I})\right)$

Descent

$\theta_{t+1} \leftarrow \theta_t - \eta_t \tilde{\mathbf{g}}_t$

Output θ_T and compute the overall privacy cost (ε, δ) using a privacy accounting method.

Figure 1: Differentially private Stochastic gradient decent

4 Local DP (25 pts)

In Random Response, consider the input data $X = (x_1, \ldots, x_n)$ and $x_i \overset{i.i.d}{\sim} B(p)$ for $i \in [1, n]$, where $B(p)$ is the Bernoulli distribution with probability p. If $p = 0.5$ and $\epsilon = 0.2$, calculate the expectation of $\Sigma_{x \in \tilde{X}} x$, where $\tilde{X}$ is the result of the Random Response. If $p = 0.1$ or 0.9, what is the expectation? Comparing the results with the expectation of $\Sigma_{x \in X} x$, what can you find?

5 *Random Subsampling (20 pts)

Given a dataset $x \in \mathcal{X}^n$, and $m \in \{0, 1, \ldots, n\}$, a random m-subsample of x is a new (random) dataset $x^{'} \in X^m$ formed by keeping a random subset of m rows from x and throwing out the remaining $n - m$ rows. Similarly to Problem 3, assume that two inputs are neighbours if one can be obtained from the other by removing or deleting one element.

5.1

Show that for every $n \in \mathbb{N}, |X| \geq 2, m \in \{1, \ldots, n\}, \epsilon > 0$, and $\delta < m/n$, the algorithm $A(x)$ that outputs a random m-subsample of $x \in \mathcal{X}^n$ is not (ϵ, δ)-differentially private.

5.2

Although random subsamples do not ensure differential privacy on their own, a random subsample does have the effect of "amplifying" differential privacy. Let $A : \mathcal{X}^m \to \mathbb{R}$ be any algorithm. We define the algorithm $A^{'}(x) : \mathcal{X}^n \to \mathbb{R}$ as follows: choose $x^{'}$ to be a random m-subsample of x, then output $A(x^{'})$.

Prove that if A is (ϵ, δ)-differentially private, then $A^{'}$ is $(\frac{(e^\epsilon - 1)m}{n}, \frac{\delta m}{n})$-differentially private. Thus, if we have an algorithm with the relatively weak guarantee of 1-differential privacy, we can get an algorithm with ϵ-differential privacy by using a random subsample of a dataset that is larger by a factor of $1/(e^\epsilon - 1) = O(1/\epsilon)$.