Homework 3 of DSC 291 Introduction to Differential Privacy (Fall 2024)

University of California, San Diego

Assigned on Nov 12, 2024(Tuesday)

Due at 11:59 pm on Nov 27, 2024 (Wednesday before Thanksgiving)

Notes:

- Be sure to read "Policy on Academic Integrity" on the course syllabus.
- There are ***[100 points]*** in this homework, and a bonus ***[15 points]***.
- You need to submit your homework via Gradescope.
- Contract the instructor if you spot typos. Any updates or correction will be posted on the course Announcements page and piazza, so check there occasionally.

---

0. Acknowledgment *[0 points]*

For each question in this HW, please list all your collaborators and reference materials (beyond those specified on the website) that were used for this homework.

1. **List of Collaborators** List the names of all people you have collaborated with and for which question(s).
2. **List of Acknowledgements.** If you find an assignment's answer or use a another source for help, acknowledge for which question and provide an appropriate citation (there is no penalty, provided you include the acknowledgement). If not, then write "none".

1 General facts about Differential Privacy [22 pts]

Please answer True or False and provide a short (one sentence explanation). (2 pts each)

(a) Differential privacy prevents all harms that could incur to an individual when a dataset including this individual is analyzed (differentially privately).

(b) Differentially Private machine learning algorithm cannot possibly predict my attribute accurately because otherwise it violates the DP guarantee.

(c) Differential Privacy prevents attackers from identifying Bob, even if the rest of the dataset (except Bob) is known.

(d) If one can implement homomorphic encryption computationally efficiently to train a machine learning model, then it provides stronger privacy guarantee than differential privacy.

(e) DP algorithms must add noise.

(f) Deterministic algorithms cannot be differentially private.

(g) (ϵ, δ)-Approximate differential privacy provides DP guarantee with probability $1 - \delta$.

(h) k-anonymity could fail completely upon composition, whereas differential privacy degrades more gracefully under composition.

(i) It is often worthwhile to exploit the sparsity in the problem of differentially private histogram release. Adding noise only to those elements of the histogram that are non-zero helps to improve the utility of the DP release.

(j) Gaussian mechanism is the only mechanism satisfying a linear upper bound of the Renyi differential privacy (i.e., Concentrated Differential Privacy).

(k) Objective perturbation mechanism dominates output perturbation mechanism in differentially private (convex) empirical risk minimization.

2 Concentrated and Renyi DP [32 pts]

(I have covered many of these in the lectures. The point of this question is to make sure you understand the lecture. If you don't follow the lecture, feel free to read the textbook, understand the proof and write them down.)

(a) (8 pts) Rewrite the definition of Renyi-divergence of $\mathcal{M}(x)$ and $\mathcal{M}(x')$ for two fixed neighboring datasets x and x' in terms of the moments generating function of the privacy loss random variable.

(b) (8 pts) Show that if $\mathcal{M}_1$ satisfies ρ_1-zCDP and $\mathcal{M}_2$ satisfies ρ_2-zCDP conditioning on any possible output of $\mathcal{M}_1$, then the composition $(\mathcal{M}_1, \mathcal{M}_2)$ satisfy $(\rho_1 + \rho_2)$-zCDP.

(c) (8 pts) Show that if $\mathcal{M}$ satisfies $(\alpha, \epsilon(\alpha))$-Renyi Differential privacy for all $\alpha > 1$, then $\mathcal{M}$ also satisfies (ϵ, δ)-DP for any $0 < \delta < 1$ and

$$\epsilon(\delta) = \min_{\alpha > 1} \epsilon(\alpha) + \frac{\log(1/\delta)}{\alpha - 1}.$$

(Hint: using tail bound lemma and the Chernoff trick.)

(d) (8 pts) Derive the optimal α and the expression of $\epsilon(\delta)$ above when $\mathcal{M}$ is a composition of k-Gaussian mechanisms, each satisfying ρ-zCDP.

3 Differentially private Linear Regression [24 pts]

In this question, you will see the connection of various mechanisms for differentially private linear regression and how they are related to each otehr. The goal is to solve

$$\min_{\theta} \sum_{i=1}^{n} (x_i^T \theta - y_i)^2 + \lambda \|\theta\|^2 = \|X\theta - y\|^2 + \lambda \|\theta\|^2$$

This this question, we do not concern choosing the parameter of the randomization to achieve a particular (ϵ, δ)-DP. You should just write down the form of the output using the native parameters of the algorithms.

(a) (8 pts) Write down the output perturbation mechanism that adds noise to the solution in terms of X, y, λ, you may assume the noise you add is $\mathcal{N}(0, \sigma^2 I_d)$

(b) (8 pts) Express the objective perturbation algorithm as a data-dependent noise-adding procedure. You may assume the noise vector $b \sim \mathcal{N}(0, \sigma^2 I_d)$

(c) (8 pts) Express the posterior sampling algorithm as a data-dependent noise adding procedure. You may just express things as scale parameter γ such that we are outputting

$$\hat{\theta} \sim P(\theta | X, y) \propto \exp(-\gamma(\sum_{i=1}^{n} (x_i^T \theta - y_i)^2 + \lambda \|\theta\|^2)).$$

(Hint: the above samples from a multivariate Gaussian mechanism. What is the mean and what is the covariance matrix?)

4 Local sensitivity and median [24 pt + 10 pt bonus]

Let dataset x contains the list of scores $[0, 100]$ students in a class gets. The goal is to release median. We will consider two simple mechanisms of interest $\mathcal{M}(x)$ with the attempt to output the actual median, and you should provide an error bound in terms of $|\mathcal{M}(x) - \text{Median}(x)|$.

The local and global sensitivity of a query f is given by

$$\text{LS}_f(x) = \max_{x' \text{ s.t. } x' \text{ is a neighbor of } x} |f(x) - f(x')|.$$

and

$$\text{GS}_f = \max_x \text{LS}_f(x).$$

respectively.

Let f be the median and consider the *add-remove* neighbor relations. (Notice that the neighbor relationship is *different from* the midterm!)

1. (8 pts) What is the local sensitivity of f? (Hint: It should depends on the order statistic)

2. (8 pts) Prove, using the definition of ε-DP, that no $var\epsilon$-DP mechanism that can return $|\mathcal{M}(x) - \text{Median}(x)| < \text{LS}_f(x)/2$ with probability more than $e^{\varepsilon}/2$ for all input x.

   (Hint: If such a mechanism exists, then the utility guarantee needs to apply to both neighbor datasets. Also, if x has a local sensitivity of Δ, the local sensitivity of a neighbor x' is at least Δ.)

3. (8 pts) Come up with an example in which adding Laplace noise with $b = \text{LS}_f(x)/\epsilon$ does not satisfy differential privacy.

4. (Bonus 5pts) Recall the exponential mechanism from the midterm. Write down the utility function to maximize (minus Total Average Deviation objective function to minimize), work out the global sensitivity and the corresponding suboptimality bound implied by the utility guarantee of the exponential mechanism. Assume the data x satisfies that there are n students in the class, and the score of $[10, 20, 30, ..., 90]$ with $n/9$ students each. (Assume n is an integer multiple of 9)

   Work out an upper bound of $|\mathcal{M}(x) - \text{Median}(x)|$ implied by the utility guarantee as a function of n, ε.

5. (Bonus 5 pts)Prove an error bound for the Laplace mechanism that release the histogram with bins $[0, \tau], [\tau, 2\tau], ..., [100 - \tau, 100]$, then calculate the median using the noisy histogram. Work out an upper bound of $|\mathcal{M}(x) - \text{Median}(x)|$ implied by the utility guarantee as a function of x, ε. Comment on the effect of τ.

   (hint: The bound should depend on x and the number of data points in each bin of the histogram.)

5 Privacy amplification by sampling [Bonus 15 pts]

Consider a dataset $x \in \{0, 1\}^n$ (represented as a binary bit vector), and the following algorithm:

1. Randomly select one coordinate of this dataset.
2. Run randomized response on the selected coordinate with parameter ϵ (i.e., output the correct value with probability $e^{\epsilon}/(1 + e^{\epsilon})$)

In this question, we will work out the privacy parameter of this algorithm under the "Replace-One" neighboring relationship.

(a) (5 pts) We will describe a dataset x under the assumption that the ordering of the coordinates does not matter. Let $0 \leq m \leq n$ be the number of 1s . Without loss of generality, assume the last coordinate is what differs between x and its neighbor x'. What is the probability of outputting 1?

(b) (5 pts) Show that x and x' are ϵ' indistinguishable. Parameterize ϵ' by n, m, ϵ.

(c) (5 pts) Show that this algorithm is ϵ''-DP. Work out the parameter ϵ''. Notice that you cannot use the general theorem of privacy amplification by sampling, but you may use it to check if you solution is correct.