CS 325: Private Data Analysis **Fall 2023**

Lecture 5: Approximate Differential Privacy

Lecturer: Di Wang *Scribes: Di Wang*

Note: *LaTeX template courtesy of UC Berkeley EECS dept.*

5.1 Approximate DP

Recall that in ϵ-DP we require the multiplication notion of similarity between distributions, that is

$$\mathbb{P}(A(D) \in E) \leq e^{\epsilon}\mathbb{P}(A(D') \in E) \tag{5.1}$$

for every event E. Intuitively this requirement seems too stringent. For example, suppose there is an event E such that $\mathbb{P}(A(D) \in E) = 0$. Then we know that $\mathbb{P}(A(D') \in E) = 0$ for every neighboring dataset and this is necessary. Since if $\mathbb{P}(A(D') \in E) = \frac{1}{2^{80}}$ and we see an outcome in E then we can distinguish whether the input data is D or D'. However, this may not be an issue since E only occurs with extremely low probability. Thus, we wish to explore a slightly more permissive variant of differential privacy that captures the intuition that these sorts of highly disclosure but extremely low probability events should be allowed. Specifically, we will modify the definition of differential privacy to allow a hybrid additive-multiplicative definition:

Definition 5.1 *A randomized algorithm $A : \mathcal{X}^n \mapsto \mathcal{Y}$ is (ϵ, δ)-DP for size n datasets if for every pair of neighboring datasets $D \sim D'$, for all $\mathbb{E} \subseteq \mathcal{Y}$,*

$$\mathbb{P}(A(D) \in E) \leq e^{\epsilon}\mathbb{P}(A(D') \in E) + \delta. \tag{5.2}$$

We often call (ϵ, δ)-DP approximate DP and ϵ-DP pure DP.

Intuitively, we can think δ as the probability that the two probabilities is not upper bounded by e^{ϵ}, or the probability of a "total privacy failure". Thus, δ should be as small as possible. However, the question is how large could δ be? Next we will show that $\delta \ll \frac{1}{n}$.

Lemma 5.2 *For (ϵ, δ)-DP algorithm, $\delta = o(\frac{1}{n})$.*

Proof: We consider the "Name and Shame Algorithm" $\mathcal{A}$: Given a dataset $D = \{x_1, \cdots, x_n\}$, for each x_i, let

$$y_i = \begin{cases} x_i, \text{ w.p } \delta \\ \text{nothing, w.p } 1 - \delta. \end{cases} \tag{5.3}$$

The algorithm $\mathcal{A}$ will return $(y_1, \cdots, y_n)$. We can easily see that $\mathcal{A}$ is $(0, \delta)$-DP. Moreover, with probability $1 - (1 - \delta)^n \approx \delta n$ the algorithm will release one individuals record. Thus if $\delta \neq o(\frac{1}{n})$, we have with constant probability the algorithm will leak someone's information, which is not private. ■

The relaxed definition of DP has many similar properties as in the pure DP:

Lemma 5.3 *Approximate DP satisfies the following properties:*

- (ϵ, δ)*-DP is closed under post-processing.* (ϵ, δ)*-DP satisfies (adaptive) composition. Running one mechanism satisfying* (ϵ_1, δ_1)*-DP followed by another mechanism satisfying* (ϵ_2, δ_2)*-DP satisfies* $(\epsilon_1 + \epsilon_2, \delta_1 + \delta_2)$*-DP.*

Note that, similar to pure DP, approximate DP satisfies a composition property where composing T mechanisms gives us a combined mechanism that is $(\epsilon T, \delta T)$-DP. However, unlike the pure DP, the composition bound could be improved.

So far we have introduced ϵ and (ϵ, δ)-DP and some of their properties, such as the **postprocessing property**, the **composition property** and the **group privacy property**. In this lecture, we will given another important property, subsampling property, which is super useful, especially in Machine Learning.

Theorem 5.4 ([2, 3]) *Let A be an (ϵ, δ)-DP algorithm. Now we construct the algorithm B as follows: On input $D = \{x_1, \cdots, x_n\}$, first we construct a new sub-sampled dataset D_S where each $x_i \in D_s$ with probability q. Then we run algorithm A on the dataset D_S. Then $B(D) = A(D_S)$ is $(\tilde{\epsilon}, \tilde{\delta})$-DP, where $\tilde{\epsilon} = \ln(1+(e^\epsilon - 1)q)$ and $\tilde{\delta} = q\delta$.*

Proof: The full proof could be found at Theorem in [5]. Here we show a simpler case, that is we will show it is $(O(\tilde{\epsilon}), \tilde{\delta})$-DP. For simplicity we will assume the range of $A(\cdot)$ is discrete. Thus, by the definition of approximate DP, our goal is to show that for any event E:

$$\frac{\mathbb{P}(B(D) \in E) - \tilde{\delta}}{\mathbb{P}(B(D') \in E)} \leq \exp(O(\tilde{\epsilon})). \tag{5.4}$$

We denote S as the index set of D_S and D, D' differs in the i-th item. Thus we have

$$\begin{aligned}
&\frac{\mathbb{P}(B(D) \in E) - \tilde{\delta}}{\mathbb{P}(B(D') \in E)} \\
&= \frac{\mathbb{P}(i \in S)\mathbb{P}(A(D_S) \in E | i \in S) + \mathbb{P}(i \notin S)\mathbb{P}(A(D_S) \in E | i \notin S) - \tilde{\delta}}{\mathbb{P}(i \in S)\mathbb{P}(A(D'_S) \in E | i \in S) + \mathbb{P}(i \notin S)\mathbb{P}(A(D'_S) \in E | i \notin S)} \\
&= \frac{q\mathbb{P}(A(D_S) \in E | i \in S) + (1-q)\mathbb{P}(A(D_S) \in E | i \notin S) - \tilde{\delta}}{q\mathbb{P}(A(D'_S) \in E | i \in S) + (1-q)\mathbb{P}(A(D'_S) \in E | i \notin S)} \\
&= \frac{\sum_{R \subseteq [n] \setminus \{i\}} \mathbb{P}(S \setminus \{i\} = R)[q\mathbb{P}(A(D_S) \in E | S = R \cup \{i\}) + (1-q)\mathbb{P}(A(D_S) \in E | S = R) - \tilde{\delta}]}{\sum_{R \subseteq [n] \setminus \{i\}} \mathbb{P}(S \setminus \{i\} = R)[q\mathbb{P}(A(D'_S) \in E | S = R \cup \{i\}) + (1-q)\mathbb{P}(A(D'_S) \in E | S = R)]} \\
&\leq \max_{R \subseteq [n] \setminus \{i\}} \frac{q\mathbb{P}(A(D_S) \in E | S = R \cup \{i\}) + (1-q)\mathbb{P}(A(D_S) \in E | S = R) - \tilde{\delta}}{q\mathbb{P}(A(D'_S) \in E | S = R \cup \{i\}) + (1-q)\mathbb{P}(A(D'_S) \in E | S = R)} \\
&\leq \max_{R \subseteq [n] \setminus \{i\}} \frac{q(e^\epsilon \mathbb{P}(A(D_S) \in E | S = R) + \delta) + (1-q)\mathbb{P}(A(D_S) \in E | S = R) - \tilde{\delta}}{(1-q)\mathbb{P}(A(D_S) \in E | S = R)} \\
&= \frac{qe^\epsilon + 1 - q}{1-q} \\
&\leq O(1 + q(e^\epsilon - 1)).
\end{aligned}$$

■

When the original algorithm satisfies $\epsilon \leq 1$, the theorem shows that the new algorithm has privacy parameter $\tilde{\epsilon} \approx q\epsilon$. For example, if we start with from a $(1, \delta)$-algorithm, then we can get $(\epsilon', \epsilon'\delta)$-DP algorithms by running on a subsample of roughly ϵ' times the size of the original.

5.2 Some Mechanisms

In proving pure DP, we focused on events E that were just singletons, $E = \{y\}$, and relied on the equivalence

$$\begin{aligned} &\forall y \in \mathcal{Y}, \mathbb{P}(A(D) = y) \leq e^{\epsilon}\mathbb{P}(A(D') = y) \\ \iff &\forall E \subseteq \mathcal{Y}, \mathbb{P}(A(D) \in E) \leq e^{\epsilon}\mathbb{P}(A(D') \in E) \end{aligned}$$

However, this is not true for approximate DP, and we have to consider all sets $E \subseteq \mathcal{Y}$. However, to prove approximate DP, it is enough to prove that if we draw y from $A(D)$ then with high probability we have $\mathbb{P}(A(D) = y) \leq e^{\epsilon}\mathbb{P}(A(D') = y)$. We can capture this idea with the following lemma and we will often use it to prove a mechanism is (ϵ, δ)-DP.

Lemma 5.5 *For a mechanism $A : \mathcal{X}^n \mapsto \mathcal{Y}$, a pair of neighboring data $D \sim D'$, defines the sets*

$$\textit{Good} = \{y \in \mathcal{Y} : \frac{\mathbb{P}(A(D) = y)}{\mathbb{P}(A(D') = y)} \leq e^{\epsilon}\}, \textit{Bad} = \mathcal{Y} - \textit{Good}. \tag{5.5}$$

If $\mathbb{P}(A(D) \in \textit{Bad}) \leq \delta$ for every pair of neighboring datasets, then A is (ϵ, δ)-DP. Note that is $A(D)$ and $A(D')$ are continuous distributions then we just replace the probability to the probability density functions.

Proof: To proof the statement, we fix a pair of neighboring datasets and an aribitrary event $E \subseteq \mathcal{Y}$. Then we can calculate

$$\begin{aligned} \mathbb{P}(A(D) \in E) &= \mathbb{P}(A(D) \in E \cap \text{Good}) + \mathbb{P}(A(D) \in E \cap \text{Bad}) && (5.6)\\ &\leq \mathbb{P}(A(D) \in E \cap \text{Good}) + \mathbb{P}(A(D) \in \text{Bad}) && (5.7)\\ &\leq \int_{E \cap \text{Good}} p_{A(D)}(y)dy + \delta && (5.8)\\ &\leq e^{\epsilon} \int_{E \cap \text{Good}} p_{A(D')}(y)dy + \delta && (5.9)\\ &\leq e^{\epsilon}\mathbb{P}(A(D') \in E) + \delta. && (5.10) \end{aligned}$$

■

It is notable that (ϵ, δ)-DP does not imply that $\mathbb{P}(A(D) \in \text{Bad}) \leq \delta$, as you might expect, although a version of this equivalence does not hold but with slightly weaker parameters.

5.2.1 Truncated Laplace Mechanism

For a given (one-dimensional) query function $f : \mathcal{X}^n \mapsto \mathbb{R}$, recall the Laplace mechanism:

$$A(D) = f(D) + \text{Lap}(\frac{\Delta}{\epsilon}) \tag{5.11}$$

can ensure ϵ-DP, where a Laplace distribution $\text{Lap}(\lambda)$ has the density function of $p_\lambda(y) = \frac{1}{2\lambda}\exp(-\frac{|y|}{\lambda})$. We will show that adding some truncated Laplace noise could ensure (ϵ, δ)-DP.

For $\lambda, \tau > 0$, the truncated Laplace distribution $\text{Lap}(\lambda, \tau)$ is defined by the probability density function

$$p(y) = \begin{cases} \frac{1}{Z}\exp(-\frac{|y|}{\lambda}), |y| \leq \tau \\ 0, |y| > \tau, \end{cases} \tag{5.12}$$

where $Z = \int_{-\tau}^{\tau} \exp(-\frac{|y|}{\lambda})dy$ is the normalizing constant.

Theorem 5.6 *For a fixed ϵ, δ, define the truncated Laplace mechanism $A : \mathcal{X}^n \mapsto \mathbb{R}$ as*

$$A(D) = f(D) + Lap(\lambda, \tau),$$

where $\lambda = \frac{\Delta}{\epsilon}$, $\tau = \frac{\Delta}{\epsilon}\ln(1 + \frac{e^\epsilon - 1}{2\delta})$ and $Z = 2\lambda(1 - \exp(-\frac{\tau}{\lambda})) = 2\frac{\Delta}{\epsilon}(1 - \frac{1}{1+\frac{e^\epsilon-1}{2\delta}})$. Then the truncated Laplace mechanism is (ϵ, δ)-DP if ϵ, δ is small enough.

Proof: First we can see that by setting λ, τ, Z, the truncated Laplace distributions is indeed a distribution. Next, we will find the "Bad" set in Lemma 6.4. For convenience we assume that $f(D) = 0, f(D') > 0$ and thus $f(D') \leq \Delta$. Thus, the density function of $A(D)$, $p_{A(D)}(\cdot)$, is just the density function of $\text{Lap}(\lambda, \tau)$, the density function of $A(D')$ is

$$p_{A(D')}(y) = \begin{cases} \frac{1}{Z}\exp(-\frac{|y - f(D')|}{\lambda}), |y - f(D')| \leq \tau \\ 0, |y - f(D')| > \tau. \end{cases} \tag{5.13}$$

Now we consider different cases.

Case 1: $y > \tau + f(D')$ or $y < -\tau$.

In this case, $p_{A(D')}(y) = p_{A(D)}(y) = 0$. Thus the ratio is $1 \leq e^\epsilon$. That is $y \in$ Good.

Case 2: $\tau < y \leq f(D') + \tau$.

In this case $p_{A(D)}(y) = 0$ and $p_{A(D')}(y) \neq 0$. Thus, the ratio is 0 and $y \in$ Good.

Case 3: $f(D') - \tau < y \leq \tau$ (Since we have $f(D') \leq \Delta < 2\tau$ if ϵ, δ are small enough)

In this case, $\frac{p_{A(D)}(y)}{p_{A(D')}(y)} \leq e^\epsilon$. Thus, $y \in$ Good.

Case 4: $-\tau \leq y \leq f(D') - \tau$.

In this case, $p_{A(D)}(y) \neq 0$ and $p_{A(D')}(y) = 0$, so the ratio is $+\infty > e^\epsilon$. Thus, we have $y \in$ Bad.

Next, we will show that $\mathbb{P}(A(D) \in [-\tau, f(D') - \tau]) \leq \delta$. To show this we have (we have $f(D') - \tau \leq 0$ if ϵ, δ are small enough)

$$\begin{aligned}
\mathbb{P}(A(D) \in [\tau, f(D') - \tau]) &= \int_{-\tau}^{f(D')-\tau} \frac{1}{Z}\exp(-\frac{|y|}{\lambda})dy \\
&= \int_{-\tau}^{f(D')-\tau} \frac{1}{Z}\exp(\frac{y}{\lambda})dy \\
&= \frac{\lambda}{Z}\exp(\frac{y}{\lambda})|_{-\tau}^{f(D')-\tau} \\
&= \frac{1}{2(1 - \exp(-\frac{\tau}{\lambda}))}\exp(-\frac{\tau}{\lambda})(\exp(\frac{f(D')}{\lambda}) - 1) \\
&\leq \frac{1}{2(\exp(\frac{\tau}{\lambda}) - 1)}(\exp(\epsilon) - 1) = \delta.
\end{aligned}$$

Thus, by Lemma 6.4 we have it is (ϵ, δ)-DP. ■

5.2.2 Gaussian Mechanism

We will see that adding noise from a Gaussian distribution, rather than a Laplace distribution, satisfies approximate DP. We first consider the one-dimensional case. Recall that the Gaussian distribution with mean μ and standard deviation σ denoted bu $\mathcal{N}(\mu, \sigma^2)$ is defined by the density function

$$p(y) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y-\mu)^2}{2\sigma^2}}. \tag{5.14}$$

We will start to show that adding Gaussian noise with mean 0 and variance $\sigma^2 = O(\frac{\Delta^2 \log \frac{1}{\delta}}{\epsilon^2})$ satisfies (ϵ, δ)-DP. Note that this standard deviation is larger than the Laplace mechanism by a factor of $O(\log \frac{1}{\delta})$. However, as we will see later in the high dimensional case, the noise given by the Gaussian mechanism could be smaller than the Laplace mechanism.

Theorem 5.7 *For a given function $f : \mathcal{X}^n \mapsto \mathbb{R}$, we defined the Gaussian mechanism as*

$$A(D) = f(D) + \mathcal{N}(0, \frac{32\Delta^2 \log \frac{2}{\delta}}{9\epsilon^2}).$$

Then for any $\epsilon \leq 1$ and $\delta > 0$, the Gaussian mechanism is (ϵ, δ)-DP.

Note that, unlike the Laplace mechanism, which works for any value $\epsilon > 0$, for analysis of the Gaussian mechanism we need ϵ to be small.

Proof: Like in the previous proof, the idea is to show that $\mathbb{P}(A(D) \in \text{Bad}) \leq \delta$. We assume $f(D) = 0$ and denote $p_{A(D)}$ and $p_{A(D')}$ as the density function of $A(D)$ and $A(D')$ respectively.

$$\begin{aligned}
\ln(\frac{p_{A(D)}(y)}{p_{A(D')}(y)}) &= \ln \frac{\exp(-\frac{y^2}{2\sigma^2})}{\exp(-\frac{(y-f(D'))^2}{2\sigma^2})} \\
&= \frac{(y - f(D'))^2 - y^2}{2\sigma^2} \\
&= \frac{f(D')^2 - 2yf(D')}{2\sigma^2} \\
&\leq \frac{9\epsilon^2}{64 \log \frac{2}{\delta}} + \frac{9|y|\Delta\epsilon^2}{32\Delta^2 \log \frac{2}{\delta}} \\
&\leq \frac{\epsilon}{4} + \frac{9|y|\Delta\epsilon^2}{32\Delta^2 \log \frac{2}{\delta}}
\end{aligned}$$

where the last inequality is due to $\log \frac{2}{\delta} > 1$. We note that if $|y| \leq \sqrt{2 \log \frac{2}{\delta}}\sigma = \frac{8\Delta \log \frac{2}{\delta}}{3\epsilon}$ then

$$\frac{\epsilon}{4} + \frac{9|y|\Delta\epsilon^2}{32\Delta^2 \log \frac{2}{\delta}} \leq \frac{\epsilon}{4} + \frac{3\epsilon}{4} = \epsilon.$$

That is $\{y : |y| \leq \sqrt{2 \log \frac{2}{\delta}}\sigma\} \subseteq$ Good. And thus,

$$\text{Bad} \subseteq \{y : |y| > \sqrt{2 \log \frac{2}{\delta}}\sigma\}.$$

We recall the following lemma of Gaussian distribution

Lemma 5.8 *If $y \sim \mathcal{N}(0, \sigma^2)$, the we have*

$$\mathbb{P}(|y| > \sigma t) \leq 2\exp(-\frac{t^2}{2}).$$

Thus, we have

$$\mathbb{P}(A(D) \in \text{Bad}) \leq \mathbb{P}(\{y : |y| > \sqrt{2\log\frac{2}{\delta}}\sigma\}) \leq \delta. \tag{5.15}$$

Thus, by Lemma 6.4, we can see that the Gaussian mechanism is (ϵ, δ)-DP. ■

The previous analysis is quite lose, actually, we can show the following theorem for the Gaussian mechanism. The proof can be found at the Appendix A of [4]

Theorem 5.9 (Gaussian Mechanism) *For a given function $f : \mathcal{X}^n \mapsto \mathbb{R}$, we defined the Gaussian mechanism as*

$$A(D) = f(D) + \mathcal{N}(0, \frac{2\Delta^2 \log\frac{1.25}{\delta}}{\epsilon^2}).$$

Then for any $\epsilon \in (0, 1)$ and $\delta > 0$, the Gaussian mechanism is (ϵ, δ)-DP.

In the previous part we have seen two different (ϵ, δ)-DP mechanisms. However, we still have not found any application or problem which allow us to get lower error. Next, we will show that the multiplicative version of the Gaussian mechanism can do that if we want to approximate some function $f : \mathcal{X}^n \mapsto \mathbb{R}^k$. Before that we will first define the ℓ_2-norm sensitivity of a query function.

Definition 5.10 *For a function $f : \mathcal{X}^n \mapsto \mathbb{R}^k$, we define its ℓ_2-norm sensitivity as*

$$\Delta_2 = \max_{D \sim D'} \|f(D) - f(D')\|_2. \tag{5.16}$$

One important thing to remember is that the ℓ_2-norm sensitivity is never more than the ℓ_1-norm. In fact we have $\Delta_2 \leq \Delta_1 \leq \sqrt{k}\Delta_2$. For example if $\mathcal{X} = \{0, 1\}^k$ and $f(D) = \sum_{i=1}^n x_i$, then $\Delta_1 = k$ while $\Delta_2 = \sqrt{k}$.

The spherical multivariate Gaussian mechanism in $\mathbb{R}^k$ with mean $\mu = (\mu_1, \cdots, \mu_k)^T$ and variance σ^2 is denoted by $\mathcal{N}(\mu, \sigma^2\mathbb{I}_k)$ is defined by the density function

$$p(y) = \frac{1}{(2\pi\sigma^2)^{\frac{k}{2}}}\exp(-\frac{\|y - \mu\|_2^2}{2\sigma^2}).$$

Actually, it is just the random variable $Z = (Z_1, \cdots, Z_k)$ where each $Z_i \sim \mathcal{N}(\mu_i, \sigma^2)$ and independent to each other.

Lemma 5.11 *If $Z \sim \mathcal{N}(\mu, \sigma^2\mathbb{I}_d)$, the we have*

- $\mathbb{E}(\|Z\|_2^2) = k\sigma^2$.
- $\mathbb{E}(\|Z\|) \leq \sqrt{k}\sigma$.
- $\mathbb{E}(\max\{|Z_1|, |Z_2|, \cdots, |Z_k|\}) \leq O(\sigma\sqrt{\log k})$.
- *For any vector v, the dot product $Z \cdot v$ is distributed as $\mathcal{N}(0, \|v\|_2^2\sigma^2)$.*

Theorem 5.12 *For a given function $f : \mathcal{X}^n \mapsto \mathbb{R}^k$, we defined the Gaussian mechanism as*

$$A(D) = f(D) + \mathcal{N}(0, \frac{32\Delta^2 \log \frac{2}{\delta}}{9\epsilon^2}\mathbb{I}_k).$$

Then for any $\epsilon \leq 1$ and $\delta > 0$, the Gaussian mechanism is (ϵ, δ)-DP.

Proof: We assume $f(D) = 0$ and denote $p_{A(D)}$ and $p_{A(D')}$ as the density function of $A(D)$ and $A(D')$ respectively.

$$\begin{aligned}
\ln(\frac{p_{A(D)}(y)}{p_{A(D')}(y)}) &= \ln \frac{\exp(-\frac{\|y\|_2^2}{2\sigma^2})}{\exp(-\frac{\|y-f(D')\|_2^2}{2\sigma^2})} \\
&= \frac{\|y - f(D')\|_2^2 - \|y\|_2^2}{2\sigma^2} \\
&= \frac{\|f(D')\|_2^2 - 2y \cdot f(D')}{2\sigma^2} \\
&\leq \frac{9\epsilon^2}{64 \log \frac{2}{\delta}} + \frac{-9y \cdot f(D')\epsilon^2}{32\Delta^2 \log \frac{2}{\delta}} \\
&\leq \frac{\epsilon}{4} + \frac{-9y \cdot f(D')\epsilon^2}{32\Delta^2 \log \frac{2}{\delta}}
\end{aligned}$$

By Lemma 6.10 we have that $-y \cdot f(D') \sim \mathcal{N}(0, \|f(D')\|_2^2\sigma^2)$. Thus,

$$\mathbb{P}(|-y \cdot f(D')| > \sqrt{2 \log \frac{2}{\delta}}\|f(D')\|_2\sigma) \leq \delta.$$

We note that if $|-y \cdot f(D')| \leq \sqrt{2 \log \frac{2}{\delta}}\|f(D')\|_2\sigma$ then

$$\frac{\epsilon}{4} + \frac{-9y \cdot f(D')\epsilon^2}{32\Delta^2 \log \frac{2}{\delta}} \leq \frac{\epsilon}{4} + \frac{3\epsilon}{4} = \epsilon.$$

That is $\{y : |-y \cdot f(D')| \leq \sqrt{2 \log \frac{2}{\delta}}\|f(D')\|_2\sigma\} \subseteq$ Good. And thus,

$$\text{Bad} \subseteq \{y : |-y \cdot f(D')| \leq \sqrt{2 \log \frac{2}{\delta}}\|f(D')\|_2\sigma\}.$$

$$\mathbb{P}(A(D) \in \text{Bad}) \leq \mathbb{P}(\{y : |-y \cdot f(D')| \leq \sqrt{2 \log \frac{2}{\delta}}\|f(D')\|_2\sigma\}) \leq \delta. \tag{5.17}$$

Thus, by Lemma 6.4, we can see that the Gaussian mechanism is (ϵ, δ)-DP. ■

5.2.3 Optimal Parameters of Gaussian Mechanism

In the previous section we study the Gaussian mechanism. However there are still two issues: Firstly, whether the above value of σ provides the minimal amount of noise required to obtain (ϵ, δ)-DP with Gaussian perturbations; Secondly, what happens in the case where $\epsilon > 1$. In this section we will solve these two problems. All the results are given by [1].

Let us go back to Lemma 5.5. For a neighboring datasets $D \sim D'$, if we denote $p_{A(D)}(y)$ and $p_{A(D')}(y)$ as the density function of $A(D)$ and $A(D')$ respectively. And we denote the privacy loss as $\ell_{D,D'}(y) = \ln(\frac{p_{A(D)}(y)}{p_{A(D')}(y)})$. We could think $L_{D,D'} = \ell_{D,D'}(Y)$ as the transformation of the output random variable $Y = A(D)$. Then Lemma 5.5 tells us

$$\forall D \sim D', \mathbb{P}[L_{D,D'} \geq \epsilon] \leq \delta. \tag{5.18}$$

For the Gaussian mechanism with variance σ^2, we have the following lemma:

Lemma 5.13 *The privacy loss $L_{D,D'}$ follows a distribution $\mathcal{N}(\eta, 2\eta)$ with $\eta = \frac{D^2}{2\sigma^2}$ where $D = \|f(D) - f(D')\|_2$.*

Proof:

$$\begin{aligned}
\ell_{D,D'}(y) &= \ln(\frac{p_{A(D)}(y)}{p_{A(D')}(y)}) \\
&= \ln(\frac{\exp(-\frac{\|y-f(D)\|_2^2}{2\sigma^2})}{\exp(-\frac{\|y-f(D')\|_2^2}{2\sigma^2})}) \\
&= \frac{\|y - f(D')\|_2^2 - \|y - f(D)\|_2^2}{2\sigma^2} \\
&= \frac{\|f(D') - f(D)\|_2^2}{2\sigma^2} + \frac{(y - f(D)) \cdot (f(D) - f(D'))}{\sigma^2}.
\end{aligned}$$

Note that since $y \sim A(D) = f(D) + \mathcal{N}(0, \sigma^2 \mathbb{I}_k)$, by the property of Gaussian distribution we have $\ell_{D,D'}(y) \sim \mathcal{N}(\frac{\|f(D')-f(D)\|_2^2}{2\sigma^2}, \frac{\|f(D)-f(D')\|_2^2}{\sigma^2})$ ■

Thus, the previous lemma tells us the privacy loss of the Gaussian mechanism is still a Gaussian distribution. Since $\eta > 0$ and $\mathbb{P}(L_{D,D'} > 0) \geq \frac{1}{2}$. It is not possible to use this sufficient condition to prove that Gaussian mechanism achieves $(0, \delta)$-DP for any $\delta < \frac{1}{2}$, which could hold actually.

Theorem 5.14 *A Gaussian mechanism with $\sigma = \frac{\Delta}{2\delta}$ is $(0, \delta)$-DP.*

Thus, our question is, what is the sufficient and necessary condition of (ϵ, δ)-DP? We will show it via the following theorem:

Theorem 5.15 *A mechanism $A : \mathcal{X}^n \mapsto \mathcal{Y}$ is (ϵ, δ)-DP if and only if the following holds for every $D \sim D'$:*

$$\mathbb{P}[L_{D,D'} \geq \epsilon] - e^{\epsilon}\mathbb{P}[L_{D',D} \leq -\epsilon] \leq \delta \tag{5.19}$$

Proof: Note that the definition of (ϵ, δ)-DP is equivalent to

$$\int_E p_{A(D)}(y) - e^{\epsilon} p_{A(D')}(y) \leq \delta, \forall E \subseteq \mathcal{Y}, D \sim D'.$$

We denote $E_* = \{y \in \mathcal{Y} : p_{A(D)}(y) \geq e^{\epsilon} p_{A(D')}(y)\}$ and $E^* = \mathcal{Y} - E_*$. Thus we have for any event E,

$$\begin{aligned}
\int_E p_{A(D)}(y) - e^{\epsilon} p_{A(D')}(y) &\leq \int_{E \cap E_*} p_{A(D)}(y) - e^{\epsilon} p_{A(D')}(y) + \int_{E \cap E^*} p_{A(D)}(y) - e^{\epsilon} p_{A(D')}(y) \\
&\leq \int_{E \cap E_*} p_{A(D)}(y) - e^{\epsilon} p_{A(D')}(y) \\
&\leq \int_{E_*} p_{A(D)}(y) - e^{\epsilon} p_{A(D')}(y).
\end{aligned}$$

Since the above inequality holds for every event, we have (ϵ, δ)-DP is equivalent to $\int_{E_*} p_{A(D)}(y) - e^{\epsilon} p_{A(D')}(y) \leq \delta$.

In the following we will show that

$$\mathbb{P}[L_{D,D'} \geq \epsilon] = \int_{E_*} p_{A(D)}(y) \tag{5.20}$$

and

$$\mathbb{P}[L_{D',D} \leq -\epsilon] = \int_{E_*} p_{A(D')}(y). \tag{5.21}$$

$$\begin{aligned}
\mathbb{P}[L_{D,D'} \geq \epsilon] &= \mathbb{P}_{y \sim A(D)}[p_{A(D)} \geq e^{\epsilon} p_{A(D')}] \\
&= \int_{\{p_{A(D)(y)} \geq e^{\epsilon} p_{A(D')(y)}\}} p_{A(D)(y)} dy \\
&= \int_{E_*} p_{A(D)}(y)
\end{aligned}$$

$$\mathbb{P}[L_{D',D} \leq -\epsilon] = \mathbb{P}_{y \sim A(D')}[e^{\epsilon} p_{A(D')} \leq p_{A(D)}] = \int_{E_*} p_{A(D')}(y).$$

Thus, we finish the proof. ■

By Theorem 5.15 and Lemma 5.13 we have

Lemma 5.16 *The Gaussian mechanism is (ϵ, δ)-DP if and only if for all $D \sim D'$*

$$\mathbb{P}(\mathcal{N}(\eta, 2\eta) \geq \epsilon) - e^{\epsilon} \mathbb{P}(\eta, 2\eta) \leq -\epsilon) \leq \delta.$$

If we denote the CDF of the standard Gaussian distribution as

$$\Phi(t) = \mathbb{P}(\mathcal{N}(0, 1) \leq t)$$

Then it is equivalent to

$$\Phi(\frac{D}{2\sigma} - \frac{\epsilon\sigma}{D}) - e^{\epsilon} \Phi(-\frac{D}{2\sigma} - \frac{\epsilon\sigma}{D}) \leq \delta \tag{5.22}$$

Lemma 5.17 *If we denote the function*

$$h(D) = \Phi(\frac{D}{2\sigma} - \frac{\epsilon\sigma}{D}) - e^{\epsilon} \Phi(-\frac{D}{2\sigma} - \frac{\epsilon\sigma}{D})$$

Then $h(D)$ is monotonically increasing.

In total we have the following theorem:

Theorem 5.18 *Let $f : \mathcal{X}^n \mapsto \mathbb{R}^k$ be a function with global ℓ_2-norm sensitivity Δ. For any $\epsilon \geq 0$ and $\delta \leq 1$, the Gaussian mechanism $A(D) = f(D) + Z$ with $Z \sim \mathcal{N}(0, \sigma^2 \mathbb{I}_k)$ is (ϵ, δ)-DP if and only if*

$$\Phi(\frac{\Delta}{2\sigma} - \frac{\epsilon\sigma}{\Delta}) - e^{\epsilon} \Phi(-\frac{\Delta}{2\sigma} - \frac{\epsilon\sigma}{\Delta}) \leq \delta. \tag{5.23}$$

We can use the above theorem to show how much we need to add when $\epsilon \geq 1$.

Theorem 5.19 *Let $f : \mathcal{X}^n \mapsto \mathbb{R}^k$ be a function with global ℓ_2-norm sensitivity Δ. Suppose $\epsilon > 0$ and $0 < \delta < \frac{1}{2} - e^{-3\epsilon}/\sqrt{4\pi\epsilon}$. If the Gaussian mechanism $A(D) = f(D) + Z$ with $Z \sim \mathcal{N}(0, \sigma^2\mathbb{I}_k)$ is (ϵ, δ)-DP then $\sigma \geq \frac{\Delta}{\sqrt{2\epsilon}}$.*

Proof: Recall that when $\sigma = \frac{\Delta}{\sqrt{2\epsilon}}$, then the mechanism will be (ϵ, δ)-DP with $\delta = \Phi(0) - e^{\epsilon}\Phi(-\sqrt{2\epsilon}) > \frac{1}{2} - e^{-3\epsilon}/\sqrt{4\pi\epsilon}$. Thus, it is impossible to achieve (ϵ, δ)-DP with $\delta = \Phi(0) - e^{\epsilon}\Phi(-\sqrt{2\epsilon})$ without increasing the variance of perturbation. ■

References

[1] Borja Balle and Yu-Xiang Wang. Improving the gaussian mechanism for differential privacy: Analytical calibration and optimal denoising. In *International Conference on Machine Learning*, pages 394–403. PMLR, 2018.

[2] Raef Bassily, Adam Smith, and Abhradeep Thakurta. Private empirical risk minimization: Efficient algorithms and tight error bounds. In *2014 IEEE 55th Annual Symposium on Foundations of Computer Science*, pages 464–473. IEEE, 2014.

[3] Amos Beimel, Shiva Prasad Kasiviswanathan, and Kobbi Nissim. Bounds on the sample complexity for private learning and private data release. In *Theory of Cryptography Conference*, pages 437–454. Springer, 2010.

[4] Cynthia Dwork, Aaron Roth, et al. The algorithmic foundations of differential privacy. *Found. Trends Theor. Comput. Sci.*, 9(3-4):211–407, 2014.

[5] Ninghui Li, Wahbeh Qardaji, and Dong Su. On sampling, anonymization, and differential privacy or, k-anonymization meets differential privacy. In *Proceedings of the 7th ACM Symposium on Information, Computer and Communications Security*, pages 32–33, 2012.