ISYE 8813: Algorithmic Foundations of Ethical ML **Fall 2021**

Lecture 6: Approx DP: Gaussian Mech and Advanced Comp

Lecturer: Juba Ziani

In previous lectures, we have seen that ε-differential privacy can be relaxed to an approximate, (ε, δ) version, where we need to guarantee

$$\Pr[\mathcal{M}(x) \in S] \leq \exp(\varepsilon) \Pr[\mathcal{M}(y) \in S] + \delta$$

on any two neighboring databases x and y and on any outcome set $S \subset \mathcal{R}$. However, all the mechanisms we have studied so far apply to $(\varepsilon, 0)$-differential privacy, and do not take advantage of the flexibility given by the additive parameter δ. In this lecture, we start by studying a mechanism that is not $(\varepsilon, 0)$-DP, but that is (ε, δ)-DP: the Gaussian mechanism.

1 The Gaussian Mechanism

As the name suggest, the Gaussian mechanism privatizes a statistic by adding Gaussian noise. However, the Gaussian mechanism requires a slightly different notion of sensitivity than the one that we have use for the multi-dimensional Laplace mechanism.

Definition 1 (ℓ_2-sensitivity)**.** *The ℓ_2-sensitivity of a function $f : \mathbb{N}^{\mathcal{X}} \to \mathbb{R}^d$ is given by*

$$\Delta_2 f \triangleq \max_{x,y \text{ neighbors}} \|f(x) - f(y)\|_2 = \sqrt{\sum_{i=1}^{d} (f_i(x) - f_i(y))^2}.$$

In the Laplace mechanism, we were adding noise according to the ℓ_1-sensitivity of our function, i.e. using the ℓ_1-norm. Note that those two norms are related: we know that for any vector $z \in \mathbb{R}^d$, we have

$$\|z\|_2 \leq \|z\|_1 \leq \sqrt{d}\|z\|_2.$$

We now remind the reader of the definition of the Gaussian distribution:

Definition 2. *The Gaussian distribution $N(\mu, \sigma^2)$ with mean μ and variance σ^2 has the following density:*

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right).$$

The Gaussian mechanism then simply adds well-chosen Gaussian noise to each coordinate of our vector-valued query $f(X)$. Formally,

Definition 3 (The Gaussian Mechanism)**.** *Let $f : \mathbb{N}^{|\mathcal{X}|} \to \mathbb{R}^d$. The Gaussian mechanism is then defined as*

$$\mathcal{M}_G(x) = f(X) + (Y_1, \ldots, Y_d),$$

where the Y_i's are drawn independently from $N(0, 2\ln(1.25/\delta) \cdot (\Delta_2 f)^2 / \varepsilon^2)$.

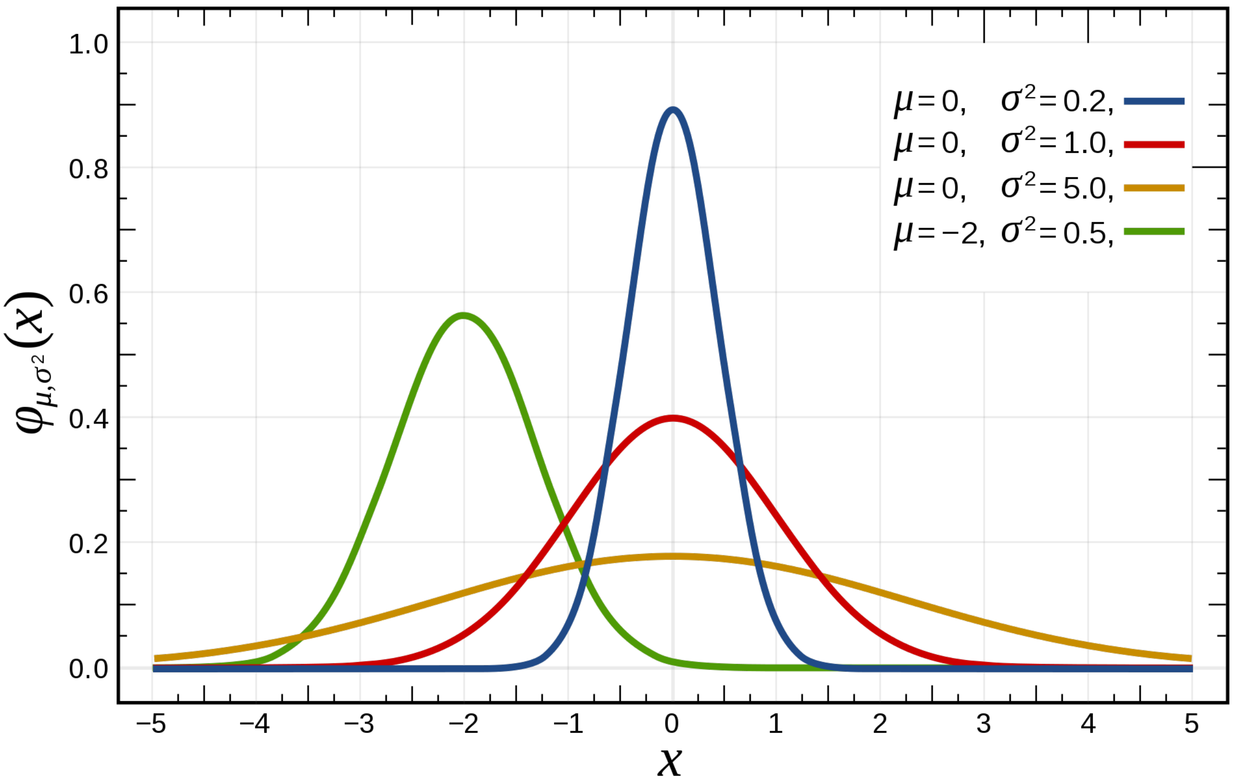

Figure 1: Gaussian pdf for different values of μ, σ

The Gaussian mechanism is not ε-differentially private for any $\varepsilon > 0$, no matter how you pick σ. This will be left as an exercise in the second problem set. However,

Theorem 4. *The Gaussian mechanism as described above is (ε, δ)-differentially private.*

Proof. We provide a partial proof of the result here. For a more careful and complete proof, please refer to Appendix A of [1].

Our goal is as usual to bound the following quantity across two neighboring databases x, y:

$$\frac{\Pr[\mathcal{M}(x) = s]}{\Pr[\mathcal{M}(y) = s]},$$

where here both $\mathcal{M}$ in s live in $\mathbb{R}^d$ and are d-dimensional vectors. Since the probability we will be working on are exponential, we can instead just work with the following random variable and aim to bound it by ε with probability at least $1 - \delta$, for $Z \sim N(0, 2\ln(1.25/\delta)$

As per the first problem set, this is sufficient to argue (ε, δ)-differential privacy:

$$\ln\left(\frac{\Pr\left[\mathcal{M}(x) = f(x) + Z\right]}{\Pr\left[\mathcal{M}(y) = f(x) + Z\right]}\right) = \ln\left(\frac{\exp\left(-\|Z\|_2^2/2\sigma^2\right)}{\exp\left(-\|f(y) - f(x) + Z)\|_2^2/2\sigma^2\right)}\right) \tag{1}$$

$$= \frac{1}{2\sigma^2}\left(-\|Z\|_2^2 + \|Z + v\|_2^2\right) \tag{2}$$

$$= \frac{1}{2\sigma^2}\left(-\|Z\|_2^2 + \|Z\|_2^2 + \|v\|^2 + 2Z^\top v\right) \tag{3}$$

$$= \frac{1}{2\sigma^2}\left(\|v\|^2 + 2Z^\top v\right) \tag{4}$$

where $v \triangleq f(x) - f(y)$. Let us focus on the 1D-case; there, we see that in absolute value, the above quantity is upper bounded by

$$\left|\frac{1}{2\sigma^2}\left(\|v\|^2 + 2Z^\top v\right)\right| \leq \frac{1}{2\sigma^2}\left(v^2 + 2|v||Z|\right)$$

$$\leq \frac{1}{2\sigma^2}\left(\Delta f^2 + 2\Delta f|Z|\right).$$

So first, we note that this is always less than ε under the condition that

$$|Z| \leq \sigma^2\varepsilon/\Delta f - \Delta f/2.$$

It only remains to show that $|Z| > \sigma^2\varepsilon/\Delta f - \Delta f/2$ with probability at most δ. Let us give some intuition on how to do this in the 1-dimensional case. By the traditional tail bounds of a Gaussian distribution, we have

$$\Pr\left[|Z| > t\right] \leq \frac{\sqrt{2}\sigma}{\sqrt{\pi}} \exp(-t^2/2\sigma^2).$$

We want $\delta \triangleq \frac{\sqrt{2}\sigma}{\sqrt{\pi}} \exp(-t^2/2\sigma^2)$, which can be rewritten (handwaving-ly)

$$t \sim \sigma\sqrt{\ln\left(\sigma/\delta\right)}.$$

With $\sigma \sim \frac{\Delta f}{\varepsilon}\sqrt{\ln(1/\delta)}$, we get $t \sim \frac{\Delta f}{\varepsilon}\ln\left(1/\delta\right)$. This roughly matches the desired lower bound on $|Z|$, which is, if we ignore the $\Delta f/2$ term,

$$\frac{\sigma^2\varepsilon}{\Delta f} \sim \frac{\Delta f}{\varepsilon}\ln(1/\delta).$$

(Here note that it makes sense to "ignore" the $\Delta f/2$ at a high level, as it is much smaller than $\Delta f/\varepsilon$ for small ε. To do this rigorously, we need to take the small terms that I ignored into account: t will contain a term that depends on $\ln \Delta f/\varepsilon$, which will lead to a slighty higher bound on $|Z|$ than what we desire; we will use the $\frac{\Delta f}{2}$ term to counteract its effect.)

Now, it is easy to see that that d-dimensional case reduces to the 1-D case. In particular, in the 1-D case, the term we are trying to bound is

$$\frac{1}{2\sigma^2}\left(v^2 + 2vZ\right),$$

which just follows a gaussian distribution with mean $\frac{v^2}{2\sigma^2}$ and variance $\frac{4v^2}{4\sigma^4}\sigma^2 \triangleq \frac{v^2}{\sigma^2}$ (follows from the fact that $a + bZ$ is still Gaussian, and $\mathbb{E}[Z] = a + b\mathbb{E}[Z] = a$, and $Var[Z] = b^2 \cdot Var[Z]$.) Now, in the multivariate case, we are interested instead in

$$\frac{1}{2\sigma^2}\left(\|v\|^2 + 2Z^\top v\right).$$

But note that $Z^\top v = \sum_i v_i Z_i$ is a weighted sum of *independent* Gaussian random variables, so is Gaussian itself. In particular, it has mean 0 and variance

$$4\sum_{i=1}^{d} v_i^2 = \|v\|^2.$$

So, we can rewrite $Z^\top v$ as $\|v\|Z'$ where $Z' \sim N(0,1)$, and we now just need to bound

$$\left|\frac{1}{2\sigma^2}\left(\|v\|^2 + 2\|v\|_2 Z'\right)\right| \leq \frac{1}{2\sigma^2}\left(\Delta f^2 + 2\Delta f|Z'|\right).$$

This is exactly the 1-D case. □

2 Advanced Composition

Let Δf be the sensitivity of query f, and let $g = (f, \ldots, f)$. We have that the ℓ_2-sensitivity of g is given by

$$\Delta g = \max_{x,y \text{ neighbors}} \sqrt{\sum_{i=1}^{d} |f(x) - f(y)|^2} = \max_{x,y \text{ neighbors}} \sqrt{d} \cdot |f(x) - f(y)| \leq \sqrt{d} \cdot \Delta f.$$

Now let us run the Gaussian mechanism on $(f(x), \ldots, f(x))$. To do so, we output $(f(x) + Z_1, \ldots, f(x) + Z_d)$ where $Z_i \sim N(0, \ln(1/\delta) \cdot \frac{(\Delta g)^2}{\varepsilon^2} = d\ln(1/\delta) \cdot \frac{(\Delta f)^2}{\varepsilon^2})$. Equivalently, one can see this as running the exponential mechanism d times, with parameter $\varepsilon' = \varepsilon/\sqrt{d}$: we then get that we have to pick

$$\sigma' \sim \ln(1/\delta)\frac{(\Delta f)^2}{\varepsilon'^{\,2}} = d\ln(1/\delta)\frac{(\Delta f)^2}{\varepsilon^2}$$

to obtain (ε, δ)-DP, as above. This, by the way, is exactly the composition of d Gaussian mechanisms with parameter ε'.

But now one may notice that if I were to apply the basic composition theorem, I would have that the composition of d Gaussian mechanisms with a ε' privacy parameter would have a privacy parameter of $d\varepsilon' = \sqrt{d}\varepsilon$. I.e., the basic composition theorem gives us a guarantee that is $\sqrt{d}$ worse than what we actually obtained!

This suggests that in the case of (ε, δ)-DP, the basic composition theorem may not be tight, and could have a dependency in $\sqrt{d}\varepsilon$ instead of ε when composing d queries. This exactly what is shown by the advanced composition theorem.

Theorem 5 (Advanced Composition)**.** *For all $\varepsilon, \delta, \delta' \geq 0$, the class of (ε, δ)-differentially private mechanisms satisfies $(\varepsilon', k\delta + \delta')$-differential privacy under k-fold adaptive composition for*

$$\varepsilon' = \sqrt{2k \ln(1/\delta')}\varepsilon + k\varepsilon\left(\exp(\varepsilon) - 1\right).$$

Proof. This is beyond the scope of the class. If interested, see [1]. □

Hence, this ability to shave off a factor of $\sqrt{d}$ in the privacy parameter is not a property of only the Gaussian mechanism; it is in fact a property of (ε, δ)-differential privacy!

A few notes about the definition:

- This theorem holds for *adaptive* composition. For a formal exposition, see [1] on page 49. But the idea is that you can pick the the mechanism you use in step j adaptively, i.e. as a function of the outputs of the previous mechanisms $\mathcal{M}_1$ to $\mathcal{M}_{j-1}$!

- Think of δ' as a parameter you get to choose. Different δ' lead to different guarantees: the smaller the δ', the better the second parameter in the composition theorem, but the first argument increases (however, it only increases at a very slow rate of $\ln(1/\delta')$, which is almost effectively constant. So, you can think of δ' as very small, and the dependency being roughly $k\delta$ as in basic composition.

- The dependency on the first argument is better than before. For ε small, we have roughly that $\exp(\varepsilon) - 1 \sim \varepsilon$, and so our guarantee becomes

  $$\varepsilon' \sim \sqrt{k}\varepsilon + k\varepsilon^2.$$

  The first term is better than the $k\varepsilon$ from the basic composition theorem by a factor of $\sqrt{k}$. The second term is better by a factor of ε (remember we think of ε small, often much smaller than 1).

- Imagine I want a mechanism that is $(\varepsilon, k\delta + \delta')$-DP for some δ by composing k mechanisms. Then I just need each mechanism to be $\left(\frac{\varepsilon}{2\sqrt{2k\ln(1/\delta')}}, \delta\right)$-DP. Indeed, we have that for small $\varepsilon \leq 1$ (in which case $\exp(\varepsilon) - 1 \leq 2\varepsilon$), the privacy parameter after composition is upper bounded by

  $$\sqrt{2k\ln(1/\delta')} \cdot \frac{\varepsilon}{2\sqrt{2k\ln(1/\delta')}} + 2k\frac{\varepsilon}{8k\ln(1/\delta')} \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{4\ln(1/\delta')}.$$

For δ' not too big (remember we want it to be very small), we have $\ln(1/\delta') \geq 1/2$ and the above bound is at most ε. That matches roughly what we saw earlier with the exponential mechanism, where we can use $\varepsilon/\sqrt{k}$ instead of the ε/k for the naive/basic composition theorem.

References

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