Chapter 4 — Algorithmic Tool: The Laplace Mechanism

Draft version: September 15, 2021

Contents

Acknowledgement. This book chapter is an extension from the course material jointly developed by Jonathan Ullman and Adam Smith. Please feel free to provide feedback.

1 A Second Example: The Laplace Mechanism

Another natural way to add randomness to a computation is to simply add noise to the output of some function f evaluated on the data. This function could just return a single real number (like a proportion or a sum), or it could be something more complex that returns a vector in $\mathbb{R}^d$ (such as the roughly 3 billion statistics produced by the US Census Bureau from its decennial census).

When does adding noise satisfy differential privacy? How does the choice of the function f we evaluate affect the amount of noise we must add? One basic idea is to look at how sensitive a function is to a change in one of its input records. We measure this via the *global sensitivity* of f:

Definition 1.1. Given a function $f : \mathcal{U}^n \to \mathbb{R}^d$, we define the *global sensitivity of f in the ℓ_1 norm* to be

$$GS_{f,\ell_1} = \sup_{\mathbf{x},\mathbf{x}' \text{ neighbors in } \mathcal{U}^n} \|f(\mathbf{x}) - f(\mathbf{x}')\|_1 . \tag{1}$$

(We often drop the subcript ℓ_1, and write simply GS_f.)

For some functions f, it is tricky to get our hands on the exact global sensitivity, and it is easier to work with an upper bound. A function has global sensitivity at most Δ (in the ℓ_1 norm) if for all pairs of neighboring data sets $\mathbf{x}, \mathbf{x}' \in \mathcal{U}^n$:

$$\|f(\mathbf{x}) - f(\mathbf{x}')\|_1 \leq \Delta . \tag{2}$$

The notation $\|\cdot\|_1$ refers to the ℓ_1 norm of a vector, which is sum of the absolute values of the vector's entries. For example, $\|(1, 0, -3)\|_1 = 4$, and $\|(1, 1, 1, 1, 1, 1)\|_1 = 6$. In 1 dimension, the ℓ_1 norm is just the absolute value.

Examples of global sensitivity A proportion $f(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^{n} \varphi(x_i)$, where $\varphi : \mathcal{U} \to \{0, 1\}$, has global sensitivity $GS_f = \frac{1}{n}$. The same is true if φ maps records to numbers in the interval $[0, 1]$.

To take another example, consider a *histogram*: given a data set $\mathbf{x} \in \mathcal{X}^n$ and a partition of $\mathcal{U}$ into d disjoint sets $B_1, ..., B_d$ (think of these as "bins" or "types" of items in $\mathcal{U}$), we count how many records there are of each type. $f(\mathbf{x}) = (n_1, n_2, ..., n_d)$ where $n_j = \#\left\{i : x_i \in B_j\right\}$. So, for example, if we wanted to compute the number of residents of each of the 50 US states from a census of the US population, we would be asking a hstogram query. The global sensitivity of a histogram query is at most 2, regardless of how many bins there are.

Exercise 1.2. Suppose we want to compute the average of a set of numbers known to lie in the interval $[0, R]$ on a data set of size n. What is the sensivitiy of the average?

The Laplace Mechanism If f has sensitivity Δ, we can satisfy ε-DP by adding noise from a *Laplace distribution* with scale $\frac{\Delta}{\varepsilon}$, independently to each entry of the output. The Laplace distribution, also called the double exponential distribution, is sort of a pointy Gaussian (Fig.1). The mean-0, scale-1 Laplace has density $h(y) = \frac{1}{2}e^{-|y|}$ for $y \in \mathbb{R}$. This distribution has expected absolute value 1,

and standard deviation $\sqrt{2}$. We can scale the distribution by a positive number $\lambda > 0$, to get the general form $\mathrm{Lap}(\lambda)$ with density

$$\mathrm{Lap}(\lambda): \quad \text{a distribution on } \mathbb{R} \text{ with p.d.f. } h_\lambda(y) = \frac{1}{2\lambda}\exp(-|y|/\lambda)$$

Figure 1 illustrates the probability density function for a few values of λ. If we translate the distribution to have mean μ, then the density becomes $\frac{1}{2\lambda}\exp(-|y-\mu|/\lambda)$.

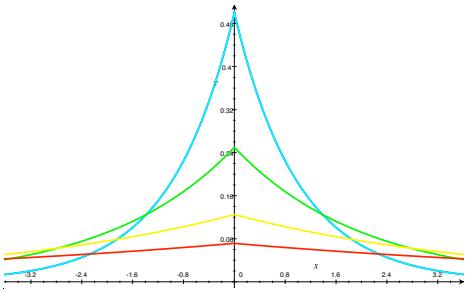

Figure 1: The probability density function of the Laplace distribution with $\lambda \in \{1, 2, 4, 7\}$.

The resulting algorithm (Algorithm 1) is called the Laplace mechanism[1], and is a basic building block for the design of many other differentially private algorithms. If used to estimate a proportion, it produces far more accurate estimates than randomized response (for the same privacy budget)—see Exercise 1.6.

Algorithm 1: Laplace mechanism $A_{\mathrm{Lap}}(\varepsilon, \mathbf{x})$

Input: Data set $\mathbf{x} = (x_1, \ldots, x_n) \in \mathcal{U}^n$ and parameter $\varepsilon > 0$.

1 Receive a query $f : \mathcal{U}^n \to \mathbb{R}^d$. Let GS_f denote its global sensitivity in the ℓ_1 norm;

2 **return** $f(\mathbf{x}) + (Z_1, \ldots, Z_d)$ where $Z_i \sim \mathrm{Lap}\left(\frac{GS_f}{\varepsilon}\right)$ are i.i.d.

Note that instead of the actual global sensitivity GS_f, we may use any upper bound $\Delta \geq GS_f$.

Theorem 1.3. *The Laplace mechanism is ε-differentially private.*

Proof. Fix two neighboring data sets $\mathbf{x}$ and $\mathbf{x}'$ in $\mathcal{U}^n$, and a query $f : \mathcal{U}^n \to \mathbb{R}^d$. Let $\Delta = GS_f$ be the ℓ_1 global sensitivity of f. Let $\mu = f(\mathbf{x})$ and $\mu' = f(\mathbf{x}')$. We know that the ℓ_1 norm of $\mu - \mu'$ is at most Δ. Comparing the output distributions of A_{Lap} on $\mathbf{x}$ and $\mathbf{x}'$ thus means comparing two Laplace distributions that have been shifted relative to one another by $\mu - \mu'$.

Because we add noise independently to each entry of the output, the density of the output at vector $\mathbf{y}$ on input $\mathbf{x}$ can be written as a product:

$$h_{\mathbf{x}}(\mathbf{y}) = \tfrac{\varepsilon}{2\Delta}e^{-\frac{\varepsilon}{\Delta}|y_1-\mu_1|} \times \tfrac{\varepsilon}{2\Delta}e^{-\frac{\varepsilon}{\Delta}|y_2-\mu_2|} \times \cdots \times \tfrac{\varepsilon}{2\Delta}e^{-\frac{\varepsilon}{\Delta}|y_d-\mu_d|}, \tag{3}$$

[1]Here and elsehwere in the course, the term "mechanism", inherited from literature on game theory, just means "algorithm".

which can be simplified to $h_{\mathbf{x}}(\mathbf{y}) = \left(\frac{\varepsilon}{2\Delta}\right)^d e^{-\frac{\varepsilon}{\Delta}\|\mathbf{y}-\mu\|_1}$. If we look at the ratio of the densities for $\mathbf{x}$ and $\mathbf{x}'$ at the same output $\mathbf{y}$, we get

$$\frac{h_{\mathbf{x}'}(\mathbf{y})}{h_{\mathbf{x}}(\mathbf{y})} = e^{-\frac{\varepsilon}{\Delta}(\|\mathbf{y}-\mu'\|_1 - \|\mathbf{y}-\mu\|_1)}. \tag{4}$$

By the triangle inequality, the difference $\|\mathbf{y}-\mu\|_1 - \|\mathbf{y}-\mu'\|_1$ is at most $\|\mu-\mu'\|_1$, which is at most Δ. We thus get:

$$\frac{h_{\mathbf{x}'}(\mathbf{y})}{h_{\mathbf{x}}(\mathbf{y})} \leq e^{\frac{\varepsilon}{\Delta}\|\mu-\mu'\|_1} \leq e^{\frac{\varepsilon}{\Delta}\cdot\Delta} = e^{\varepsilon}. \tag{5}$$

And that is enough to conclude the mechanism is ε-DP: For any measurable set E, we have $\Pr(A_{\mathsf{Lap}}(\varepsilon, \mathbf{x}) \in E = \int_{\mathbf{y}\in E} h_{\mathbf{x}}(\mathbf{y})$. So if the ratio of the densities is bounded everywhere by e^{ε}, then so is the ratio of the probabilities of any given event E. □

Thus, we can guarantee differential privacy by adding noise to the output of a function that scales with the function's sensitivity. Histograms, for example, fit the framework well: we can get away with adding error whose expected magnitude is $2/\varepsilon$ to each of the bin counts, regardless of the number of bins.

The following lemma provides useful bounds on the magnitude of the Laplace mechanism's error.

Lemma 1.4.

1. *If* $Z \sim \mathsf{Lap}(\lambda)$ *is a Laplace-distributed random variable, we have*
   (a) $\mathbb{E}(|Z|) = \lambda$
   (b) $\sqrt{\mathbb{E}(Z^2)} = \sqrt{2}\lambda$
   (c) *For every* $t > 0$: $\mathbb{P}(|Z| > t\lambda) \leq \exp(-t)$.
2. *Let* $Z_1, \ldots, Z_d$ *are i.i.d.* $\mathsf{Lap}(\lambda)$ *random variables, and let* $M = \max(|Z_1|, \ldots, |Z_d|)$.
   (a) $\mathbb{E}(\|(Z_1, \ldots, Z_d)\|_1) = d\lambda$
   (b) *For every* $t > 0$: $\mathbb{P}(M > \lambda(\ln(d) + t)) \leq \exp(-t)$.
   (c) $\mathbb{E}(M) \leq \lambda(\ln(d) + 1)$.

Exercise 1.5. Suppose we use the Laplace mechanism to estimate the number of individuals in a data set who reside in each of the 3,143 counties[2] in the US, using parameter $\varepsilon = 0.1$. What does Lemma 1.4 imply about the expected error of the count for Suffolk County, Mass.? What does it imply about the expectation of the largest error in the estimate of any county population?

Exercise 1.6. Suppose we use the Laplace mechanism to estimate the fraction $f(\mathbf{x}) = \frac{1}{n}\sum_{i=1}^{n}\varphi(x_i)$ where $\varphi : \mathcal{U} \to [0,1]$ is a predicate. Give an expression for the data set size $n(\alpha, \varepsilon)$ at which the Laplace mechanism's root mean square error $\sqrt{\mathbb{E}\left((A(\mathbf{x}) - f(\mathbf{x}))^2\right)}$ drops below α (when run with privacy parameter ε). How does this compare to the analogous calculation for Randomized Response? What happens to $n(\alpha, \varepsilon)$ when α is halved? When ε is halved?

Notice that the accuracy of the Laplace mechanism is pretty bad when ε is very small. Suppose we want to estimate an fraction (as in the Exercise 1.6). If we set $\varepsilon = 1/n$, then the standard deviation of the Laplace mechanism is $\sqrt{2} \cdot \frac{\Delta}{1/n} = \sqrt{2}$ (since $\Delta = 1/n$ in this case). But the fraction

[2] This number includes county equivalents, and was drawn from the Wikipedia article "List of United States counties and county equivalents" in February 2021.

can only take values between 0 and 1—the noise is thus of larger magnitude than the "'signal" one is trying to release. This feature is inherent. As we will see next lecture, differentially private algorithms cannot yield any useful information when $\varepsilon < 1/n$.

Exercise 1.7. Prove Lemma 1.4. To bound $\mathbb{E}(M)$ in the final statement, it may be useful to recall that for any nonnegative random variable M, we have $\mathbb{E}(M) = \int_{x=0}^{\infty} \Pr(M > x)$. This inequality allows us to compute expectations in terms of "tail bounds".

Summary

1.1 Key Points

- To reason about information leakage and confidentiality we must look at the algorithms that process the data, not only the form of the output.
- Differential privacy is one way to quantify how much an algorithm leaks about individual inputs. It is parametrized by a positive real number ε, which bounds the amount of leakage.
- Randomized response and the Laplace mechanism satisfy ε-DP. For the same privacy budget, the Laplace mechanism adds far less error.

Additional Reading and Watching

- The paper that defined differnetial privacy [DMNS06, DMNS16]
- A nontechnical introduction to DP [WAB+18]
- Dwork and Roth book, Chapter 2 [DR14]
- Videos from the MinutePhysics Youtube channel: "Protecting Privacy with MATH" and "When It's OK to Violate Privacy", 2019.
- Tutorial talks by K. Ligett (NeurIPS 2016 Tutorial) and A. Smith (NASIT 2019 Tutorial).
- For further discussion of composition attacks, see [GKS08]

References

[DMNS06] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating noise to sensitivity in private data analysis. In *Conference on Theory of Cryptography*, TCC '06, 2006.

[DMNS16] Cynthia Dwork, Frank McSherry, Kobbi Nissim, and Adam Smith. Calibrating noise to sensitivity in private data analysis. *Journal of Privacy and Confidentiality*, 7(3), 2016.

[DR14] Cynthia Dwork and Aaron Roth. *The Algorithmic Foundations of Differential Privacy*. NOW Publishers, 2014.

[GKS08] Srivatsava Ranjit Ganta, Shiva Prasad Kasiviswanathan, and Adam Smith. Composition attacks and auxiliary information in data privacy. In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '08, 2008. ACM.

[WAB+18] Alexandra Wood, Micah Altman, Aaron Bembenek, Mark Bun, Marco Gaboardi, James Honaker, Kobbi Nissim, David O'Brien, Thomas Steinke, and Salil Vadhan. Differential privacy: A primer for a non-technical audience. *Vanderbilt Journal of Entertainment & Technology Law*, 21(1):209, 2018.