CS291A Introduction to Differential Private **Fall 2021**

Lecture 13: Adaptive DP I: Smoothed Sensitivity and Median

Lecturer: 133 *Scribes: Fuheng Zhao*

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

13.1 Inefficient Global Sensitivity Approach

13.1.1 Median Query

Assume each data point in a dataset of size N (N is odd) is drawn from a bounded universe {0, ..., U} without losing of generality. Then, let dataset x contains $ceil(N/2)$ of zeros and $floor(N/2)$ of Us and its median is 0, i.e, {0, ...0,U,...U}. Using the replace-one neighbouring dataset definition, there exists a x' neighbouring dataset such that it contains $floor(N/2)$ of zeros and $ceil(N/2)$ of Us. The x''s median is U, and hence $|median(x) - median(x')| = U$.

13.1.2 Linear Regression

In linear regression, the goal is to solve $\text{argmin}_\theta ||X\theta - Y||^2 + \lambda||\theta||^2$ and $\theta^* = (X^TX)^{-1}X^TY$ where $D = (X, Y)$ are the dataset. However, using add-one neighbouring dataset definition, the neighbouring dataset with new data point (x, y) is $D' = ((X, x), (Y, y)$. The $\hat{\theta}^* = (X^TX + x^Tx)^{-1}(X^TY + xy)$. The global sensitivity becomes unbounded, and the global sensitivity approach does not exploit the well conditioned dataset. If we know $x^Tx > \alpha nI$, then $||\hat{\theta}^* - \theta^*||_2 \leq \frac{L}{\lambda(x^Tx)}$ which yields better convergence bound.

13.2 Local Sensitivity

Local Sensitivity is definied as $LS_q(x) = max\{q(x) - q(x')|d(x, x') \leq 1\}$ and it measures the stability of a query at a particular dataset. The local sensitivity for median query in a sorted dataset D becomes $max\{D_{\frac{n+1}{2}} - D_{\frac{n-1}{2}}, D_{\frac{n+3}{2}} - D_{\frac{n+1}{2}}\}$; and the local sensitivity for linear regression becomes $\frac{2L}{\lambda_{min}(x^Tx)}$. However, the magnitude of the noise may reveal sensitive information about the dataset itself. Recently, Gadotti et al. [1] presents a noise-exploitation attacks and from the noise, one can infer private information about individuals in the dataset.

13.3 Data-Dependent Differential Privacy

"Data-dependent DP mechanism" aims at more stably calibrating noise to local sensitivity (at least for query releases), and there are many different approaches:

- Smooth sensitivity
- Propose-test-release
- Privately bounding the local-sensitivity
- Stability-based query release (Distance2Stability)

13.3.1 Smooth Sensitivity

Since we can not add noise based on local sensitivity, it is ideal to construct smooth upper bound of local sensitivity, and the noise should satisfies stability under "translation" and "scaling" are admissible

Definition 13.3.1 (Smooth Sensitivity)**.** For β ¿ 0, the β smooth sensitivity of f is $S^*_{f,\beta} = max_{y \in D^n}(LS_f(y) \cdot e^{-\beta d(x,y)})$.

Smooth sensitivity satisfies a smoothing property and it is the optimal bound satisfying this property. The two property that one should satisfy to smooth out the local sensitivity: (i) $\forall x \in D^n : S(x) \geq LS_f(x)$; (ii) $\forall x, y \in D^n, d(x,y) = 1 : S(x) < e^{\beta} S(y)$

Lemma 13.1. $S^{f,\beta}$ is a β-smooth upper bound on LS_f. In addition, $S^{f,\beta}(x) \leq S(x)$ for all $x \in D^n$ for every β-smooth upper bound S on LS_f .

Notation. For a subset S of ^{d}, we write $S + \Delta$ for the set $\{z + \Delta | z \in S\}$, and $e^{\lambda} \cdot S$ for the set $\{e^{\lambda} \cdot z | z \in S\}$. We also write $a \pm b$ for the interval $[a - b, a + b]$.

Definition 13.3.2 (Admissible Noise Distribution)**.** A probability distribution h on R^d is (α, β)-admissible if, for $\alpha = \alpha(\epsilon, \delta)$, $\beta = \beta(\epsilon, \delta)$, the following two conditions hold for all $||\Delta|| \leq \alpha$ and $|\lambda| \leq \beta$, and for all subsets SR^d:
Sliding Property: $Pr_{Z\ h}(Z \in S) \leq e^{\frac{\epsilon}{2}} \cdot Pr_{Z\ h}(Z \in S + \Delta) + \frac{\delta}{2}$.
Dilation Property: $Pr_{Z\ h}(Z \in S) \leq e^{\frac{\epsilon}{2}} \cdot Pr_{Z\ h}(Z \in e^{\lambda} \cdot S) + \frac{\delta}{2}$.

Then, $A(x) = f(x) + \frac{S(x)}{\alpha} \cdot Z$ satisfies (ϵ, δ)-DP.

13.3.2 Privacy Analysis

Similar to group privacy, we know that $P(A(x) \in S) \leq e^{\epsilon} P(A(x') \in S) + \delta'$, and $\Delta = \frac{\alpha(f(x) - f(x'))}{S(x)}$, $|\Delta|_1 \leq \frac{\alpha LS(x)}{S(x)} \leq]\alpha$

$P(A(x) \in S) = P(f(x) + \frac{S(x)}{\alpha}) \cdot Z \in S)$
$= P(Z \in \frac{\alpha(S - f(x))}{S(x)})$
$\leq e^{\frac{\epsilon}{2}} P(Z \in \frac{\alpha(S - f(x'))}{S(x)}) + \frac{\delta}{2}$
$= e^{\frac{\epsilon}{2}} P(Z \in \frac{S(x')}{S(x)} \frac{\alpha(S - f(x')}{S(x')}) + \delta/2$
$\leq e^{\frac{\epsilon}{2}} (e^{\frac{\epsilon}{2}} P(Z \in \frac{\alpha(S - f(x')}{S(x')}) + \delta/2) + \delta/2$
$= e^{\epsilon} P(f(x') \in S) + (e^{\frac{\epsilon}{2}} + 1) \frac{\delta}{2}$

13.3.3 Example Noises

Lemma 13.2. For any $\gamma > 1$, the distribution with density $h(z) \propto \frac{1}{1+|z|^\gamma}$ is $(\frac{\epsilon}{2\gamma+1}, \frac{\epsilon}{2\gamma+1})$-admissible (with $\delta = 0$). Moreover, the d-dimensional product of independent copies of h is $(\frac{\epsilon}{2\gamma+1}, \frac{\epsilon}{2\gamma+1})$-admissible.

Lemma 13.3. For ϵ, $\delta \in$ (0,1), the d-dimensional Laplace distribution, $h(z) = \frac{1}{2^d} \cdot e^{||z||_1}$, is (α, β)-admissible with $\alpha = \frac{\epsilon}{2}$, and $\beta = \frac{\epsilon}{2\rho_{\frac{\delta}{2}}||Z||_1}$, where Z h. In practice, it suffices to use $\alpha = \frac{\epsilon}{2}$ and $\beta = \frac{\epsilon}{4(d+ln(2/\delta))}$. For $d = 1$, it suffices to use $\beta = \frac{\epsilon}{2ln(2/\delta)}$.

Lemma 13.4. Guassian Distribution: For ϵ, $\delta \in$ (0,1), the d-dimensional Laplace distribution, $h(z) = \frac{1}{(2\pi)^{d/2}} \cdot e^{-0.5||z||_2^2}$, is (α, β)-admissible for Euclidean metrics with $\alpha = \frac{\epsilon}{5\rho_{\delta/2}(Z_1)}$ and $\beta = \frac{\epsilon}{2\rho_{\delta/2}(||Z||_2^2)}$, where $Z = (Z_1, ..., Z_d)$ h. In particular, it suffices to take $\alpha = \frac{\epsilon}{5\sqrt{2ln(2/\delta)}}$ and $\beta = \frac{\epsilon}{4(d+ln(2/\delta))}$.

An easier way to solve this optimization is:

$$S^*_{f,\epsilon}(x) = max_{k=1,...,n} 2^{-k\epsilon}(max_{y:d(x,y)=k} LS_f(y))$$

References

[1] Andrea Gadotti et al. "When the signal is in the noise: Exploiting Diffix's Sticky Noise." In: *28th {USENIX} Security Symposium ({USENIX} Security 19)*. 2019, pp. 1081–1098.