CS291A Introduction to Differential Privacy **Fall 2021**

Lecture 6: Advanced Composition (Part II), Gaussian mechanism (October 13)

Lecturer: Yu-Xiang Wang *Scribes: Kaiqi Zhang*

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

6.1 Recap

6.1.1 Exponential mechanism

Utility of Exponential mechanism: utility is defined as

$$u(x,y) = -\max_{q\in Q} \Big|\frac{1}{n}q^T x - \frac{1}{\|y\|_1}q^T y\Big|$$

It satisfy with probability $1-\beta$,

$$u(x,y^*) - u(x,\mathcal{M}(x)) \le \frac{2\Delta u}{\epsilon}\log\frac{|R|}{\beta}$$

Approximation error of a SmallDB: Let $m = \|\tilde{x}\|_1 \ge \frac{\log Q}{\alpha^2}$, $\exists \tilde{x}, s.t. error(\tilde{x}) \le \alpha$, and with probability at least $1-\delta$,

$$error(\mathcal{M}(x)) \le \alpha + \frac{2}{n\epsilon}\Big(\frac{\log|x|\log|Q|}{\alpha^2} + \log\frac{1}{\beta}\Big).$$

6.1.2 Advanced composition

Theorem 1. *The adaptive composition of k (ϵ,δ)-DP mechanisms satisfies $(\tilde{\epsilon},\tilde{\delta})$-DP where*

$$\tilde{\epsilon} = \epsilon\sqrt{2k\log(1/\delta'))} + 2k\epsilon^2, \tilde{\delta} = k\delta + \delta'$$

for any $\epsilon,\delta \ge 0, \delta' \ge 0$.

Application: Laplace mechanism, AboveThresh.

6.1.3 Privacy loss random variable

Definition:

$$\epsilon_{\mathcal{M}}^{x,x'} = \log\frac{p(y)}{p(y')}$$

where $y \sim \mathcal{M}(x)$.

Lemma 2. *(Tail bound to (ϵ, δ)-DP conversion). Let $\epsilon_{\mathcal{M}_i}^{x,x'}$ be the PLRV of a mechanism. If*

$$P(\epsilon_{\mathcal{M}_i}^{x,x'} > \epsilon) \leq \delta$$

for all pairs of neighboring x, x', then $\mathcal{M}$ satisfies (ϵ, δ)-DP.

6.2 Proof of advanced composition for pure DP mechanisms

Claim 3. *The privacy loss random variable (PLRV) of adaptive composition is the sum of each mechanism.*

Proof. Fix two neighboring datasets, consider a sequence of adaptively chosen pure-DP mechanisms $\mathcal{M}_1, \ldots, \mathcal{M}_k$, outputing $y_1, \ldots, y_k$ respectively.

$$\begin{aligned}
\epsilon_{\mathcal{M}_1 \ldots \mathcal{M}_k}^{x,x'} &= \log \frac{p(y_1 \ldots y_k)}{p'(y_1 \ldots y_k)} \\
&= \log \frac{p(y_1)p(y_2|y_1)p(y_k|y_1 \ldots y_{k-1})}{p'(p(y_1)p'(y_2|y_1)p'(y_k|y_1 \ldots y_{k-1})} \\
&= \sum_{i=1}^{k} \epsilon_{\mathcal{M}_i(\cdot, y_1 \ldots y_{i-1})}^{x,x'} \leq k\epsilon
\end{aligned}$$

□

Proof Idea of Advanced Composition:

- Observation 1: sometimes PLRV is positive, other times negative. They cancel with each other.
- Observation 2: as k gets larger, the sum of PLRV concentrates around its mean. One can calculate their mean and bound the deviation from the mean.
- Observation 3: the adaptivity means that the PLRVwill depend on the past.

To prove advanced composition, we introduce Martingale and apply Azuma-Hoeffding's inequality.

Definition 1. *Martingale: a sequence of r.v. $S_1, \ldots, S_n, \ldots$ is a Martingale if for any n,*

$$\begin{aligned}
&\mathbb{E}[|S_n|] < \infty, \\
&\mathbb{E}[S_{n+1}|S_1, \ldots, S_n] = S_n
\end{aligned}$$

Lemma 4. *Azuma-Hoeffding's inequality: Assume $X_1, \ldots, X_n$ are Martingale differences*

$$S_n = X_1 + \cdots + X_n$$

then S_n can be bounded w.h.p:

$$P[S_n > \epsilon] \leq \exp\Big(-\frac{2\epsilon^2}{\sum_{i=1}^{n}(b_i - a_i)^2}\Big)$$

Let the martingale differences be

$$X_i = \epsilon_{\cdot|y_1,\ldots,y_{i-1}}^{x,x'} - \mathbb{E}[\epsilon_{\cdot|y_1,\ldots,y_{i-1}}^{x,x'}|y_1, \ldots, y_{n-1}].$$

and they are bounded by

$$-\epsilon - \mathbb{E}[\epsilon^{x,x'}_{\cdot|y_1,\ldots,y_{i-1}}|y_1,\ldots,y_{n-1}] \le X_i \le \epsilon - \mathbb{E}[\epsilon^{x,x'}_{\cdot|y_1,\ldots,y_{i-1}}|y_1,\ldots,y_{n-1}].$$

Fix x, x', apply Azuma-Hoeffding's inequality

$$P[\sum_{i=1}^{k} \epsilon^{x,x'}_{\mathcal{M}_i} - \mathbb{E}\sum_{i=1}^{k} \epsilon^{x,x'}_{\mathcal{M}_i} \ge t] \le \exp\Big(-\frac{2t^2}{4k\epsilon^2}\Big)$$

which shows that $(\mathcal{M}_1,\ldots,\mathcal{M}_k)$ satisfy $(\tilde{\epsilon},\tilde{\delta})$-DP with

$$\begin{aligned}\tilde{\epsilon} &= \mathbb{E}[\sum_{i=1}^{k} \epsilon^{x,x'}_{\mathcal{M}_i}] + \epsilon\sqrt{2k\log(1/\tilde{\delta})} \\ &\le 2k\epsilon^2 + \epsilon\sqrt{2k\log(1/\tilde{\delta})}\end{aligned}$$

Here we omit the condition $(\cdot|y_1,\ldots,y_{i-1})$. To prove the last inequality above, observe the expectation of PLRV is the KL divergence:

$$\mathbb{E}[\epsilon^{x,x'}_{\mathcal{M}_i}] = \int p(y)\log\frac{p(y)}{q(y)} = D_{KL}(P||Q)$$

KL-divergence is always nonnegative, and satisfy Pinsker's inequality:

Lemma 5. *Pinsker's inequality:*

$$\|P-Q\|_1 \le \sqrt{2D_{KL}(P||Q)}$$

so it can be bounded by

$$\begin{aligned}D_{KL}(P||Q) &= D_{KL}(P||Q) + D_{KL}(Q||P) - D_{KL}(Q||P) \\ &\le \int p(x)\log\frac{p(x)}{q(x)}dx + \int q(x)\log\frac{q(x)}{p(x)}dx \\ &= \int (p(x)-q(x))\frac{p(x)}{q(x)}dx \\ &\le \epsilon\|P-Q\|_1 \\ &\le \epsilon\sqrt{2D_{KL}(P||Q)}\end{aligned}$$

which indicates that

$$D_{KL}(P||Q) \le 2\epsilon^2.$$

There are improved bounds of the KL-divergence and tighter version of Advanced composition:

- Bound from Dwork and Roth book $D_{KL}(P||Q) \le \epsilon(e^\epsilon - 1)$.
- Bound from Bun and Steinke: $D_{KL}(P||Q) \le \epsilon^2/2$.
- Tight bound from Adam Smith (also in the proof of Bun and Steinke): $D_{KL}(P||Q) \le \epsilon\frac{e^\epsilon-1}{e^\epsilon+1} = \epsilon\tanh(\epsilon/2)$.

6.3 Gaussian mechanism

Mechanism: given $f : \mathbb{N}^{|x|} \to \mathbb{R}^d$, output $f(x) + \mathcal{N}(0, \epsilon^2 I_d)$.

Advantages:

- Gaussian noise is more concentrated than Laplace noise.
- L2 sensitivities are often lower than L1 sensitivities.

6.3.1 PLRV of the Gaussian mechanism

Privacy Loss Random Variable of the Gaussian mechanism is Gaussian:

$$\epsilon_{\mathcal{M}}^{x,x'} \sim \mathcal{N}(\eta, 2\eta)$$

where $\eta = D^2/2\sigma^2$, $D = \|f(x) - f(x')\|$.

Proof.

$$\begin{aligned}\log \frac{\exp(-\frac{\|f(x)-y\|^2}{2\sigma^2})}{\exp(-\frac{\|f'(x)-y\|^2}{2\sigma^2})} &= \frac{-\|f(x)-y\|^2 + \|f(x')-y\|^2}{2\sigma^2} \\ &= \frac{1}{2\sigma^2}(\|f(x)-f(x')\|^2 + 2(f(x')-f(x))^T(f(x)-y)) \sim \mathcal{N}(\eta, 2\eta).\end{aligned}$$

□

6.3.2 Privacy analysis

Recall Lemma 2, using the tail bound of Gaussian mechanism, one can prove Gaussian mechanism is DP:

Lemma 6. *Gaussian tail bound: let $X \sim \mathcal{N}(\mu, \sigma^2)$, we have*

$$P(X - \mu \geq u) \leq \exp(-u^2/(2\sigma^2))$$

Combining the above two lemmas, one can find that the Gaussian mechanism with variance σ^2 for a query with L2-sensitivity Δ satisfies (ϵ, δ)-DP with

$$\epsilon = \frac{\Delta^2}{2\sigma^2} + \frac{\Delta^2}{\sigma}\sqrt{2\log(1/\delta)}$$

For $0 < \epsilon, \delta \leq 1$, the mechanism observe (ϵ, δ)-DP if we choose

$$\sigma = \frac{\Delta}{\epsilon}\sqrt{2\log(1.25/\delta)}.$$

6.4 Concentrated Differential Privacy

6.4.1 Centration inequalities

Markov's inequality: For any non-negative r.v. $X : P(X \geq t) \leq \frac{\mathbb{E}[X]}{t}$.

Chebychev's inequality: For any r.v. with variable $\sigma^2, P(|X - E[X]| \geq t\sigma) \leq \frac{1}{t^2}$.
(proof: taking $(X - E[X])^2$ as the r.v. and apply Markov's inequality.)

Generalizing Chebychev inequality: $P(|X - E[X]| \geq t) \leq \frac{\mathbb{E}[|X-E[X]|^k]}{t^k}$.

Chebychev's method: Define $\mu = \mathbb{E}[X]$. For any $t > 0$, we have that

$$P((X - \mu) \geq u) = P(\exp(t(X - \mu)) \geq \exp(tu)) \leq \frac{\mathbb{E}[\exp(t(X - \mu))]}{\exp(tu)}$$

which leads to Chebychev's bound:

$$P((X - \mu) \geq u) \leq \inf_{0 \leq t \leq b} \exp(-t(u + \mu))\mathbb{E}[\exp(tX)].$$

6.4.2 Subgaussian random variables

We say a random variable with mean μ is σ-subgaussian if

$$\mathbb{E}[\exp(t(X - \mu))] \leq \exp(\sigma^2 t^2/2)$$

for all $t \in \mathbb{R}$. We say that X is subgaussian if there exists constants σ. Gaussian random variables and bounded random variables are subgaussian. The tail of subgaussian random variables can be bounded by

$$P(X - \mu > u) \leq \exp(-u^2/(2\sigma^2))$$

The proof uses Chernoff's method and set $t = \frac{u}{\sigma^2}$.

Claim 7. *Average of n independent σ-subgaussian RVs is $\frac{\sigma}{\sqrt{n}}$ subgaussian.*

Proof. Define $\hat{\mu} = \frac{1}{n}\sum_{i=1}^n X_i$. Obvoulsy , $\mathbb{E}[\hat{\mu}] = \frac{1}{n}\sum_{i=1}^n \mu = \mu$

$$\begin{aligned}
\mathbb{E}[\exp(t(\hat{\mu} - \mu))] &= \mathbb{E}[\exp(t/n \sum_{i=1}^n (X_i - \mu))] \\
&= \prod_{i=1}^n \mathbb{E}[\exp(t(X_i - \mu)/n)] \\
&\leq \exp(t^2\sigma^2/(2n))
\end{aligned}$$

□

This implies that

$$P(|\hat{\mu} - \mu| > k\sigma/\sqrt{n}) \leq 2\exp(-k^2/2).$$

To handle mechanisms that are Gaussian-mechanism-like, one can prove that for any neighboring datasets, the PLRV is σ-subgaussian. Then the sum of k PLRVs is $\sigma\sqrt{k}$-subgaussian, which proves the composition.

6.4.3 Renyi Differential Privacy

The Moment Generating function of PLRV is

$$\begin{aligned}\mathbb{E}_{x\sim P}[\exp(t\log\frac{p(x)}{q(x)})] &= \mathbb{E}_{x\sim P}[(\frac{p(x)}{q(x)})^t] \\ &= \int p(x)(\frac{p(x)}{q(x)})^t dx \\ &= \int q(x)((\frac{p(x)}{q(x)})^{t+1})dx \\ &= \mathbb{E}_{x\sim Q}((\frac{p(x)}{q(x)})^{t+1})\end{aligned}$$

Definition 2. *Renyi divergence: for* $\alpha \in (0,1)\cup(1,\infty)$,

$$D_\alpha(P||Q) = \frac{1}{\alpha-1}\ln\int p^\alpha q^{1-\alpha}du$$

Special cases:

- $\alpha = 1$: KL divergence.
- $\alpha = \infty$: $D_\infty(P||Q) = \in (\operatorname{ess\,sup}_P \frac{p}{q})$.
- $\alpha = 2$: χ^2 divergence.
- $\alpha = 1/2$: Hellinger distance.

We say that a mechanism satisfies (α,ϵ)-Renyi DP, if

$$D_\alpha(\mathcal{M}(x)||\mathcal{M}(x')) \le \epsilon.$$

We say a mechanism satisfies ρ-zCDP, if

$$D_\alpha(\mathcal{M}(x)||\mathcal{M}(x')) \le \rho\alpha, \forall\alpha > 1.$$

If a mechanism is ρ-zCDP, it's PLRV is $O(\rho)$ subgaussian.

Properties:

- Adaptive composition: if $\mathcal{M}_1$ is (α,ϵ_1)-Renyi DP, $\mathcal{M}_2$ is (α,ϵ_2)-Renyi DP, then $(\mathcal{M}_1,\mathcal{M}_2)$ is $(\alpha,\epsilon_1+\epsilon_2)$-Renyi DP.
- Conversion to approximate DP: (α,ϵ_1)-Renyi DP implies $(\epsilon(\alpha)+\frac{\log(1/\delta)}{\alpha-1},\delta)-DP$. ρ-zCDP implies $(\rho+2\sqrt{\rho\log(1/\delta)},\delta)$-DP.
- Other properties: Postprocessing, risk multiplier, group privacy (see Mironov, 2017).

zCDP provides tighter composition:

1. pure-DP mechanism: $\epsilon = \frac{k}{2}\epsilon^2 + \epsilon\sqrt{2\log(1/\delta)}$.
2. Gaussian mechanism: $\epsilon = \frac{k\Delta^2}{2\sigma^2} + \frac{\Delta}{\sigma}\sqrt{2k\log(1/\delta)}$.