APPROXIMATE DP:

(ε, δ)-DP if $M : \mathcal{X}^n \to \mathcal{Y}$

$\forall$ neighboring databases x, x' and all $t \in \mathcal{Y}$,

$$\Pr[M(x) = t] \leq e^{\varepsilon} \cdot \Pr[M(x') = t] + \delta$$

L_2- SENSITIVITY:

$$f : \mathcal{X}^n \to \mathbb{R}^k$$

$$S_2(f) = \max_{\underbrace{x, x'}_{\text{neighboring databases}}} \| f(x) - f(x') \|_2$$

GAUSSIAN MECHANISM:

→ x

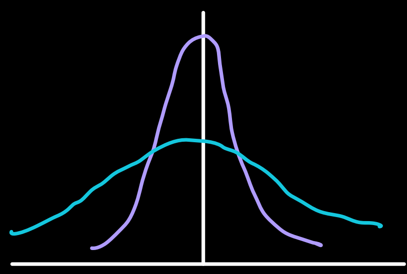
→ $f : \mathcal{X}^n \to \mathbb{R}^k$

→ Sample a noise vector $z_1, \ldots, z_k$ that are i.i.d $N(0, \sigma^2)$

→ Return $f(x) + (z_1, z_2, \ldots, z_k)$

Recall : Gaussian

$$\Pr[z = y] = \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-y^2/2\sigma^2}$$



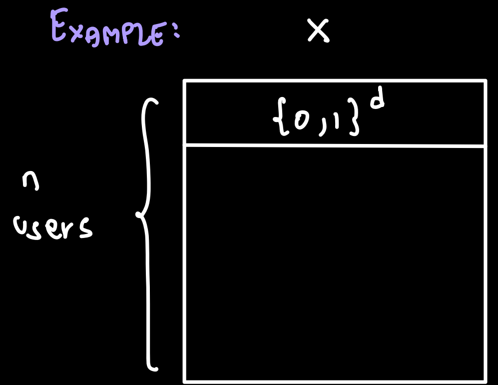
THEOREM:

Let $f: X^n \to \mathbb{R}^k$. Then Gaussian Mechanism with

$$\sigma = \underbrace{\sqrt{2\ln(1.25/\delta)}}\cdot \frac{S_2(f)}{\varepsilon}$$ ensures (ε, δ) - DP.

Really think of this as $\sigma = c \cdot \sqrt{\ln(1/\delta)} \cdot \dfrac{S_2(f)}{\varepsilon}$

EXAMPLE:



$$f(x) = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad (\in \mathbb{R}^d)$$

$$S_1(f) = \max_{\underbrace{x, x'}_{\text{neighboring}}} \| f(x) - f(x') \|_1 = \frac{d}{n}$$

$$S_2(f) = \max_{x, x'} \| f(x) - f(x') \|_2 = \frac{\sqrt{d}}{n}$$

To use Laplacian Mechanism, we would add much more noise than to use Gaussian mechanism.

"Mean Release"

LAPLACIAN

→ Noise $\simeq d/n$

→ Error : d/n per coordinate

GAUSSIAN

Noise $\simeq \sqrt{d}/n$

Error $\approx \sqrt{d}/n$ per coordinate.

GROUP PRIVACY OF APPROXIMATE DP:

Suppose x and x' are two databases that differ in k rows.

If M is (ε, δ)-DP,

$$\Pr[M(x) = y] \leq e^{k\varepsilon} \Pr[M(x') = y] + k \cdot e^{k\varepsilon} \cdot \delta$$

(Recall for pure DP, we had

$$\Pr[M(x) = y] \leq e^{k\varepsilon} \Pr[M(x') = y]$$

Recall:

Suppose $M_1, \cdots, M_k$ are (ε- Differentially Private). Then, their composition $\underbrace{M_1 \circ M_2 \circ \cdots \circ M_k}_{\text{"Adaptive Queries"}}$ is $k \cdot \varepsilon$-DP.

Basic Composition Theorem for Approximate DP:

If $M_1, M_2, \ldots, M_k$ are (ε, δ)-DP adaptive queries, then the composition is $(k\varepsilon, k\delta)$-DP.

Advanced Composition Theorem:

If $M_1, M_2, \ldots, M_k$ are (ε, δ') DP mechanisms for answering adaptive queries. Then, the composition is

$$\left(\varepsilon\sqrt{2k\ln(1/\delta)} + k\varepsilon^2, k\delta' + \delta\right)\text{-DP.}$$

(as long as $\varepsilon < 1$)

Remark:

Main win for advanced composition is that you get non-trivial privacy for the composition when $\varepsilon \ll 1/\sqrt{k}$.

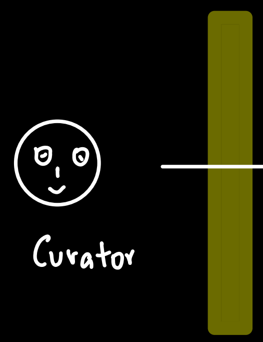
Remark:

The above was actually used to make Kaggle "leaderboard" better.

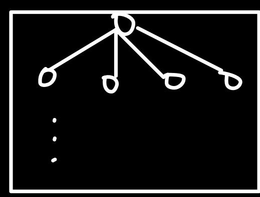
"A Reliable Leaderboard for Machine Learning Competitions".

PRIVATE ML:

User 1
User 2
⋮
User n



$\longrightarrow$ A Model / Predictor ...



Weights can "memorize" some information.

PRIVATE EMPERICAL RISK MINIMIZATION:

x_1	y_1
x_2	y_2
⋮	
x_n	y_n

Parametric family of predictors

$$h_\theta : X \to Y$$

$$L(\theta) = \frac{1}{n} \sum_{i=1}^{n} \ell(\theta; x_i, y_i)$$

The loss that h_θ gets on predicting y_i from x_i.

$$\theta^* = \arg\min_{\theta} L(\theta) = f(\bar{x})$$

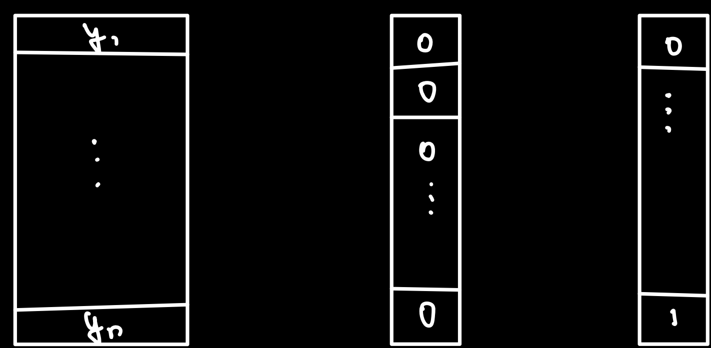
We have to answer query f privately.

Approach 1:

Add input noise or output noise.

PROBLEM: Sensitivity of θ^* is very high.

Example:



Mean is 0 Mean is non-zero.

Predictor family is just constants. Loss is squared loss:

$$L(\theta) = \frac{1}{n} \sum_{i=1}^{n} (y_i - \theta)^2$$

$$\theta^* = \arg\min L(\theta) = \frac{1}{n}(y_1 + \dots + y_n)$$

Recall that for DP, we need to have closeness in distributional outcomes.

THREE APPROACHES FOR PRIVATE ERM:

a. Output Perturbation : * Compute θ^*

* Add Noise

b. "Objective" Perturbation : * $\tilde{L}(\theta) = L(\theta) + \underbrace{\langle b, \theta \rangle}_{\text{"Some noise vector"}}$

* Output

$\arg\min \tilde{L}(\theta)$

"Privacy guarantees depend on exact solution".

c. "Gradient" perturbation

Intuitively Add additional noise to each gradient step for solving ERM.

PRIVATE STOCHASTIC GRADIENT DESCENT:

$$L(\theta) = \frac{1}{n} \sum_{i=1}^{n} \ell(\theta; x_i, y_i)$$

SGD:

→ Pick a start θ_0

→ For $t = 1, \ldots, T$:

a. Pick index $i \in [n]$ uniformly at random

$$\theta_{t+1} = \theta_t - \underbrace{\eta_t}_{\text{"Step-size"}} \nabla_\theta \ell(\theta; x_i, y_i)$$

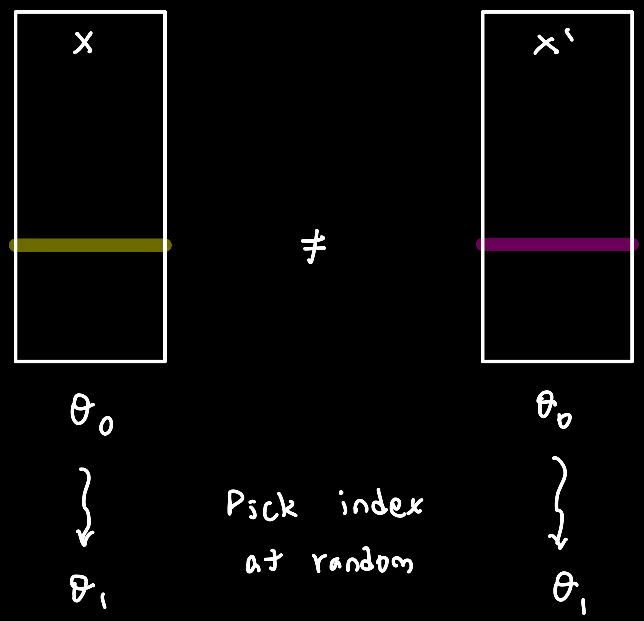
DP-SGD:

→ Pick a start θ_0

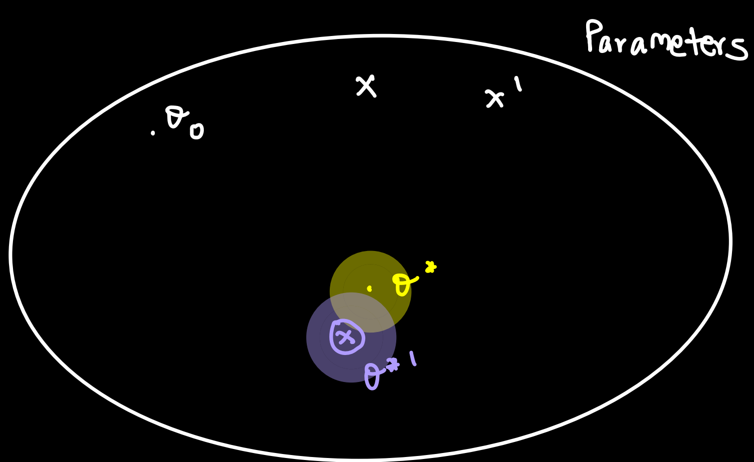
→ For $t = 1, \ldots, T$:

a. Pick index $i \in [n]$ uniformly at random

$$\theta_{t+1} = \theta_t - \eta_t \left(\nabla_\theta \ell(\theta; x_i, y_i) + \underbrace{\text{Noise}_t}_{\text{Typically Gaussian}}\right)$$



as long as picked index is not the one entry where the two differ, we have same state.



Idea: Since SGD has "randomized" to begin with adding a bit of noise does not hurt too much.

1. How much noise needs to be added to ensure (ε, δ)-DP.

Depends on how many iterations of SGD you are going to use?

Depends on "sensitivity" of gradient function".

Advanced Composition is really helpful.

Suppose we were looking to solve

$$\min_{\theta} L(\theta) \quad ; \quad L(\theta) = \frac{1}{n} \sum_{i=1}^{n} \ell(\theta; x_i, y_i)$$

SGD for L convex, 1-Lipschitz and $\|\theta_0 - \theta^*\| \leq 1$

→ Then running "true SGD" for $T = O(1/\alpha^2)$ iterations

gives $L() - L(\theta^*) \leq \alpha$

accuracy

THEOREM:

We can use Gaussian Mechanism + DP-SGD to get (ε, δ)-DP and accuracy α if $n > \dfrac{\sqrt{d}\sqrt{\log(1/\delta)}}{\varepsilon \cdot \alpha^2}$.

n: size of the database

REMARKS:

1. Privacy is important and ad-hoc solutions don't work.
2. We need to quantify privacy: DP is one of the best ways to do so.
3. Companies now use DP in aggregating data.
4. Many topics we did not cover...