Differentially Private Empirical Risk Minimization: Efficient Algorithms and Tight Error Bounds

Presented by Ronak Mehta and Sathya Ravi

R. Bassily A. Smith A. Tahkurta

Outline

Cryptography

Differential Privacy

Motivation

Cryptography

Medical Records

Differential Privacy

Motivation

Cryptography

Medical Records

Trade Secrets

An Intuitive Example

The Median

- $X = \{88, 43, 23, 90, 80, 44, 12, 38, 46, 56, 43\}$

An Intuitive Example

The Median

- $X = \{88, 43, 23, 90, 80, 44, 12, 38, 46, 56, 43\}$
- For symmetric distributions, the sample median is an unbiased estimator of the population mean.

An Intuitive Example

The Median

- $X = \{88, 43, 23, 90, 80, 44, 12, 38, 46, 56, 43\}$
- For symmetric distributions, the sample median is an unbiased estimator of the population mean.
- $Median(X) = 44$.

An Intuitive Example

The Median

- $X = \{88, 43, 23, 90, 80, 44, 12, 38, 46, 56, 43\}$
- For symmetric distributions, the sample median is an unbiased estimator of the population mean.
- $Median(X) = 44$. **What's wrong with this?**

An Intuitive Example

The Median

- $X = \{88, 43, 23, 90, 80, 44, 12, 38, 46, 56, 43\}$
- For symmetric distributions, the sample median is an unbiased estimator of the population mean.
- $\mathit{Median}(X) = 44$. **What's wrong with this?**
- We have exposed an exact value of one of the inputs!

An Intuitive Example

The Median

- $X = \{88, 43, 23, 90, 80, 44, 12, 38, 46, 56, 43\}$
- For symmetric distributions, the sample median is an unbiased estimator of the population mean.
- $Median(X) = 44$. **What's wrong with this?**
- We have exposed an exact value of one of the inputs!
- **Goal**: We want to collect some useful information about our data while ensuring some level of "privacy."

Differential Privacy

More formally...

Differential Privacy

An algorithm $\mathcal{A}$ is (ϵ, δ)-differentially private if for all sample sets D, D' differing by one sample, for all events in the output space $\mathcal{O}$ of $\mathcal{A}$,

$$\mathcal{P}(\mathcal{A}(D) \in \mathcal{O}) \leq e^{\epsilon}\mathcal{P}(\mathcal{A}(D') \in \mathcal{O}) + \delta \tag{1}$$

Differential Privacy

More formally...

Differential Privacy

An algorithm $\mathcal{A}$ is (ϵ, δ)-differentially private if for all sample sets D, D' differing by one sample, for all events in the output space $\mathcal{O}$ of $\mathcal{A}$,

$$\mathcal{P}(\mathcal{A}(D) \in \mathcal{O}) \leq e^{\epsilon}\mathcal{P}(\mathcal{A}(D') \in \mathcal{O}) + \delta \tag{1}$$

- For $\epsilon < 1, \delta \ll 1/n$, we have meaningful privacy guarantees.

Differential Privacy

More formally...

Differential Privacy

An algorithm $\mathcal{A}$ is (ϵ, δ)-differentially private if for all sample sets D, D' differing by one sample, for all events in the output space $\mathcal{O}$ of $\mathcal{A}$,

$$\mathcal{P}(\mathcal{A}(D) \in \mathcal{O}) \leq e^{\epsilon}\mathcal{P}(\mathcal{A}(D') \in \mathcal{O}) + \delta \tag{1}$$

- For $\epsilon < 1, \delta \ll 1/n$, we have meaningful privacy guarantees.
- Randomization is almost always essential for any non-trivial privacy guarantee. (Deterministic algorithms differing in output on D and D' can be used to learn the difference in input.)
- $\delta \neq 0$ implies a weaker privacy, and fewer guarantees.
- We will focus on the $(\epsilon, 0)$ setting.

- The *privacy loss* is defined as
- $$\mathcal{L}^{(\xi)}_{\mathcal{M}(x)||\mathcal{M}(y)} = \ln\left(\frac{\mathbb{P}[\mathcal{M}(x)=\xi]}{\mathbb{P}[\mathcal{M}(y)=\xi]}\right) \tag{2}$$
- May be positive or negative.

- The *privacy loss* is defined as
- $$\mathcal{L}^{(\xi)}_{\mathcal{M}(x)||\mathcal{M}(y)} = \ln\left(\frac{\mathbb{P}[\mathcal{M}(x)=\xi]}{\mathbb{P}[\mathcal{M}(y)=\xi]}\right) \tag{2}$$
- May be positive or negative.
- **Key Result:** (ϵ, δ)-differential privacy ensures that for all adjacent x, y, the absolute value of the privacy loss will be bounded by ϵ with probability at least $1-\delta$.

- The *privacy loss* is defined as
- $$\mathcal{L}^{(\xi)}_{\mathcal{M}(x)||\mathcal{M}(y)} = \ln\left(\frac{\mathbb{P}[\mathcal{M}(x)=\xi]}{\mathbb{P}[\mathcal{M}(y)=\xi]}\right) \tag{2}$$
- May be positive or negative.
- **Key Result:** (ϵ, δ)-differential privacy ensures that for all adjacent x, y, the absolute value of the privacy loss will be bounded by ϵ with probability at least $1-\delta$.
- Related to stability? PAC Learnability?

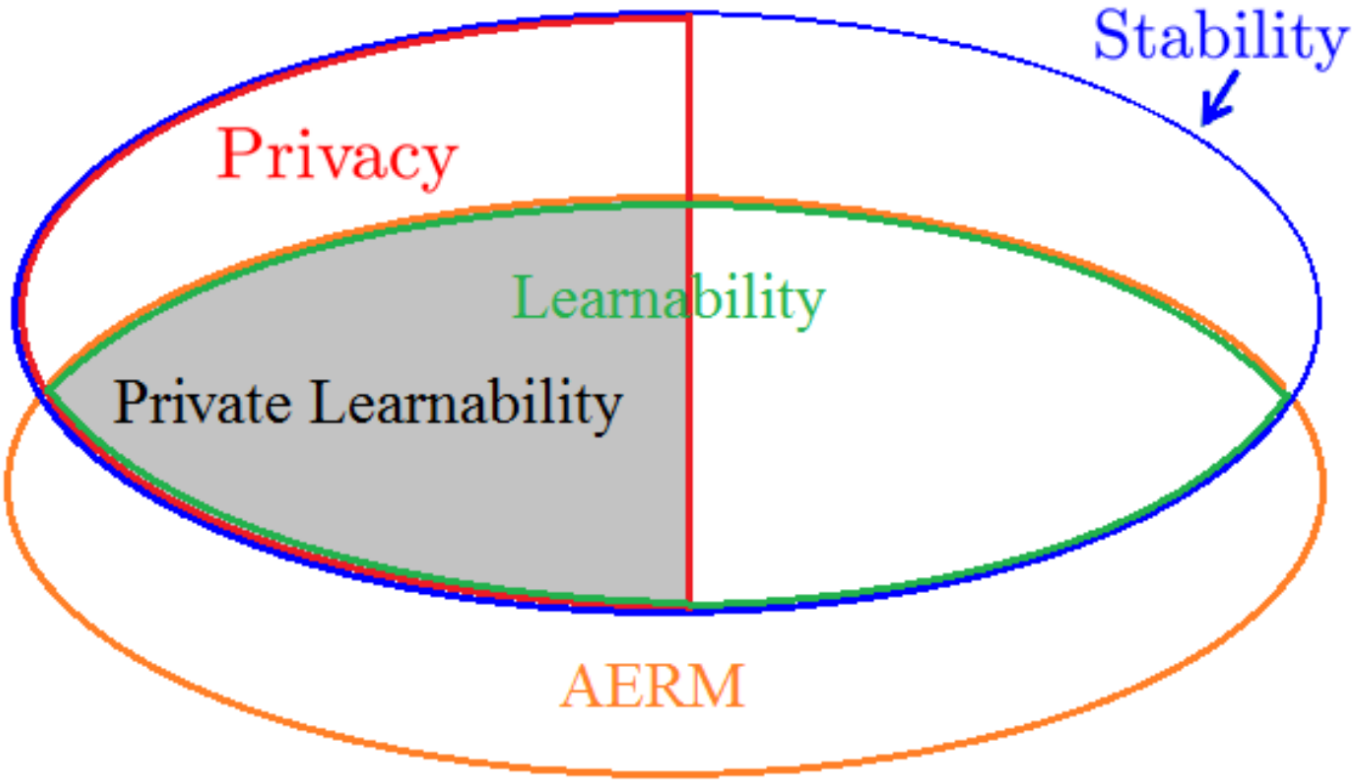

Figure: Some relationships among learning paradigms [1].

- For each input, randomly choose between it and a random other sample.

- For each input, randomly choose between it and a random other sample.
- For some binary property P,
  1. Flip a coin.
  2. If **tails**, take the binary input as given.
  3. If **heads**, flip a second coin and respond "Yes" if heads, "No" if tails.

- For each input, randomly choose between it and a random other sample.
- For some binary property P,
  1. Flip a coin.
  2. If **tails**, take the binary input as given.
  3. If **heads**, flip a second coin and respond "Yes" if heads, "No" if tails.
- "Privacy" comes from plausible deniability of any outcome.
- Accuracy comes from understanding of noise generation process.

- With a fair coin, this is $(\ln 3, 0)$-differentially private.
- $\mathbb{P}[\text{Resp.} = \text{Yes} | \text{Truth} = \text{Yes}] = \mathbb{P}[\text{Resp.} = \text{No} | \text{Truth} = \text{No}] = \frac{3}{4}$.
- $\mathbb{P}[\text{Resp.} = \text{Yes} | \text{Truth} = \text{No}] = \mathbb{P}[\text{Resp.} = \text{No} | \text{Truth} = \text{Yes}] = \frac{1}{4}$.
-

$$\frac{\mathbb{P}[\text{Resp.} = \text{Yes} | \text{Truth} = \text{Yes}]}{\mathbb{P}[\text{Resp.} = \text{Yes} | \text{Truth} = \text{No}]} = \tag{3}$$

$$\frac{\mathbb{P}[\text{Resp.} = \text{No} | \text{Truth} = \text{No}]}{\mathbb{P}[\text{Resp.} = \text{No} | \text{Truth} = \text{Yes}]} = \frac{3/4}{1/4} = 3 \tag{4}$$

1. Let $f : D \to \mathbb{R}$ (imagine counting queries).

General Methods

Laplace Mechanism

1. Let $f : D \to \mathbb{R}$ (imagine counting queries).
2. Sensitivity of a function: $S_f := \max_{D,D'} |f(D) - f(D')|$.
3. Algorithm: On querying f, output $f + \text{lap}\left(\frac{S_f}{\epsilon}\right)$.
4. Next analysis...

Let $y \in \mathcal{O}$.

$$\frac{\mathcal{P}(f(D) + \text{lap}(S_f/\epsilon) = y)}{\mathcal{P}(f(D') + \text{lap}(S_f/\epsilon) = y)} = \frac{\exp(-\frac{|y-f(D)|\epsilon}{S_f})}{\exp(-\frac{|y-f(D')|\epsilon}{S_f})}$$

$$= \exp\left(\frac{\epsilon}{S_f}(|y - f(D')| - |y - f(D)|)\right)$$

$$\leq \exp\left(\frac{\epsilon}{S_f}(|f(D) - f(D')|)\right) \leq e^{\epsilon}$$

General Methods

Laplace Mechanism: d dimensions

1. Define $S_f := \max_{D,D'} \|f(D) - f(D')\|_1$

General Methods

Laplace Mechanism: d dimensions

1. Define $S_f := \max_{D,D'} \|f(D) - f(D')\|_1$
2. Use laplacian noise, can use the same proof as before.

General Methods

Laplace Mechanism: d dimensions

1. Define $S_f := \max_{D,D'} \|f(D) - f(D')\|_1$
2. Use laplacian noise, can use the same proof as before.
3. Specifically, on query f, to achieve $\epsilon-$differential privacy, add $[\text{lap}(S_f/\epsilon)]^d$

1. Nonnumeric queries: Common hair color in this room?
2. "High sensitivity": Price selection.

1. Nonnumeric queries: Common hair color in this room?
2. "High sensitivity": Price selection.
3. Define a mechanism $M : \mathcal{D} \to R$, where R is some abstract range.
4. Define quality score $q : \mathcal{D} \times R \to \mathbb{R}$. Intuitively, $q(D, r)$ represents how good r is for database D.

1. Nonnumeric queries: Common hair color in this room?
2. "High sensitivity": Price selection.
3. Define a mechanism $M : \mathcal{D} \to R$, where R is some abstract range.
4. Define quality score $q : \mathcal{D} \times R \to \mathbb{R}$. Intuitively, $q(D, r)$ represents how good r is for database D.
5. Sensitivitiy $S_q := \max_{D,D',r \in R} |q(D, r) - q(D', r)|$.

1. Output $r \in R$ with probability proportional to

$$\mathcal{P}(r) \propto \exp\left(\frac{\epsilon q(D, r)}{2S_q}\right) \tag{5}$$

1. Output $r \in R$ with probability proportional to

$$\mathcal{P}(r) \propto \exp\left(\frac{\epsilon q(D, r)}{2S_q}\right) \tag{5}$$

2. Essentially make high quality outputs exponentially more likely.

Theorem

Exponential mechanism preserves $(\epsilon, 0)$ *differential privacy.*

Theorem

Exponential mechanism preserves $(\epsilon, 0)$ *differential privacy.*

Proof.

Similar to laplacian noise, we consider,

$$\frac{\mathcal{P}[\mathit{Exponential}(D, R, q, \epsilon)]}{\mathcal{P}[\mathit{Exponential}(D', R, q, \epsilon)]} = \frac{\frac{\exp(\frac{\epsilon q(D,r)}{2S_q})}{\sum_{r'} \exp(\frac{\epsilon q(D,r')}{2S_1})}}{\frac{\exp(\frac{\epsilon q(D',r)}{2S_q})}{\sum_{r'} \exp(\frac{\epsilon q(D',r')}{2S_1})}} = \frac{\exp(\frac{\epsilon q(D,r)}{2S_q})}{\exp(\frac{\epsilon q(D',r)}{2S_q})} \frac{\sum_{r'} \exp(\frac{\epsilon q(D',r')}{2S_q})}{\sum_{r'} \exp(\frac{\epsilon q(D,r')}{2S_q})} \tag{6}$$

Bound the two terms separately. □

Theorem

Exponential mechanism preserves $(\epsilon, 0)$ differential privacy.

Proof.

Similar to laplacian noise, we consider,

$$\frac{\mathcal{P}[Exponential(D, R, q, \epsilon)]}{\mathcal{P}[Exponential(D', R, q, \epsilon)]} = \frac{\frac{\exp(\frac{\epsilon q(D,r)}{2S_q})}{\sum_{r'} \exp(\frac{\epsilon q(D,r')}{2S_1})}}{\frac{\exp(\frac{\epsilon q(D',r)}{2S_q})}{\sum_{r'} \exp(\frac{\epsilon q(D',r')}{2S_1})}} = \frac{\exp(\frac{\epsilon q(D,r)}{2S_q})}{\exp(\frac{\epsilon q(D',r)}{2S_q})} \frac{\sum_{r'} \exp(\frac{\epsilon q(D',r')}{2S_q})}{\sum_{r'} \exp(\frac{\epsilon q(D,r')}{2S_q})} \tag{6}$$

Bound the two terms separately. □

- *Gaussian noise can also be used, similar to laplacian mechanism. It uses sensitivity in ℓ_2 norm.*

General Methods

Composition theorems

1. Suppose we have k algorithms $\mathcal{A}_i(D, z_i)$ where z_i is some auxiliary input and $\mathcal{A}_i$ are $\epsilon-$ differentially private for any auxiliary input z_i.
2. Consider the sequence $\{z_i = \mathcal{A}_i(D, \oplus_{j=1}^{i-1} z_j)\}$. Let the output be $\mathcal{A}(D) = z_k$.

1. Suppose we have k algorithms $\mathcal{A}_i(D, z_i)$ where z_i is some auxiliary input and $\mathcal{A}_i$ are $\epsilon-$ differentially private for any auxiliary input z_i.
2. Consider the sequence $\{z_i = \mathcal{A}_i(D, \oplus_{j=1}^{i-1} z_j)\}$. Let the output be $\mathcal{A}(D) = z_k$.

Theorem

$\mathcal{A}(D)$ is $k\epsilon-$differentially private.

Proof.

$$\begin{aligned}
&\mathcal{P}[\mathcal{A}(D) = z_k] \\
&= \mathcal{P}[\mathcal{A}_1(D) = z_1]\mathcal{P}[\mathcal{A}_1(D; z_1) = z_2] \cdots \mathcal{P}[\mathcal{A}_k(D; z_1, \cdots, z_{k-1}) = z_k] \\
&\leq \exp(k\epsilon) \prod_{i=1}^{k} \mathcal{P}[\mathcal{A}_i(D'; z_1, \cdots, z_{i-1}) = z_i] = \exp(k\epsilon)\mathcal{P}[\mathcal{A}(D') = z_k]
\end{aligned}$$

1. Suppose we have k algorithms $\mathcal{A}_i(D, z_i)$ where z_i is some auxiliary input and $\mathcal{A}_i$ are $\epsilon-$ differentially private for any auxiliary input z_i.
2. Consider the sequence $\{z_i = \mathcal{A}_i(D, \oplus_{j=1}^{i-1} z_j)\}$. Let the output be $\mathcal{A}(D) = z_k$.

Theorem

$\mathcal{A}(D)$ is $k\epsilon-$differentially private.

Proof.

$$\begin{aligned}
&\mathcal{P}[\mathcal{A}(D) = z_k] \\
&= \mathcal{P}[\mathcal{A}_1(D) = z_1]\mathcal{P}[\mathcal{A}_1(D; z_1) = z_2] \cdots \mathcal{P}[\mathcal{A}_k(D; z_1, \cdots, z_{k-1}) = z_k] \\
&\leq \exp(k\epsilon) \prod_{i=1}^{k} \mathcal{P}[\mathcal{A}_i(D'; z_1, \cdots, z_{i-1}) = z_i] = \exp(k\epsilon)\mathcal{P}[\mathcal{A}(D') = z_k]
\end{aligned}$$

- *Next: To the paper!* □

Convex (ϵ, δ) privacy: Phew, finally!

GD

$\mathcal{A}_{\text{NOISE-GD}}$: Differentially Private Gradient Descent

---

Input: Data set $\mathcal{D}$ of size n, loss function l, privacy parameters ϵ, δ, learning rate function $\eta : [n^2] \to \mathbb{R}$ and convex set $\mathcal{C}$.

1. Set noise variance $\sigma \leftarrow c(L, n, \delta, \epsilon)$.
2. Choose $\tilde{\theta}_1 \in \mathcal{C}$
3. **for** $t = 1 : n^2 - 1$ **do**
4. Pick $d \propto_u \mathcal{D}$ with replacement.
5. $\tilde{\theta}_{t+1} = \Pi_{\mathcal{C}}(\tilde{\theta}_t - \eta(t)[n g_d(\tilde{\theta}_t) + b_t])$ where $b_t \propto \mathcal{N}(0, I\sigma^2)$.
6. Output $\theta^{priv} = \tilde{\theta}_{n^2}$.

Convex (ϵ, δ) privacy

Analysis

Theorem

$\mathcal{A}_{\text{NOISE-GD}}$ is $(\epsilon, \delta)-$differentially private.

Proof.

Sketch. Let $X_t(\mathcal{D}) = ng_d(\tilde{\theta}_t) + b_t$ be a random variable over b_t and conditioned on $\tilde{\theta}_t$. Privacy loss (at iteration t) can be written as,

$$W_t = \left| \log \frac{\mu_{X_t(\mathcal{D})}(X_t(\mathcal{D}))}{\mu_{X_t(\mathcal{D}')}X_t(\mathcal{D}')} \right| \tag{7}$$

where $\mu_{X_t(\mathcal{D})}(y)$ is the measure of the r.v. $X_t(\mathcal{D})$ induced on $y \in \mathbb{R}$.

1. *Use gaussian noise privacy (with amplification) to get that w.p. at least $1 - \delta$, $W_t \leq \frac{\epsilon}{n\sqrt{\log(1/\delta)}}$.*

Convex (ϵ, δ) privacy

Analysis

Theorem

$\mathcal{A}_{\text{NOISE-GD}}$ is $(\epsilon, \delta)-$differentially private.

Proof.

Sketch. Let $X_t(\mathcal{D}) = ng_d(\tilde{\theta}_t) + b_t$ be a random variable over b_t and conditioned on $\tilde{\theta}_t$. Privacy loss (at iteration t) can be written as,

$$W_t = \left| \log \frac{\mu_{X_t(\mathcal{D})}(X_t(\mathcal{D}))}{\mu_{X_t(\mathcal{D}')}X_t(\mathcal{D}')} \right| \tag{7}$$

where $\mu_{X_t(\mathcal{D})}(y)$ is the measure of the r.v. $X_t(\mathcal{D})$ induced on $y \in \mathbb{R}$.

1. *Use gaussian noise privacy (with amplification) to get that w.p. at least $1 - \delta$, $W_t \leq \frac{\epsilon}{n\sqrt{\log(1/\delta)}}$.*
2. *Use* **strong composition** *to get the final result for n^2 iterations.*

Convex (ϵ, δ) privacy

Analysis

Theorem

$\mathcal{A}_{\text{NOISE-GD}}$ is $(\epsilon, \delta)-$differentially private.

Proof.

Sketch. Let $X_t(\mathcal{D}) = ng_d(\tilde{\theta}_t) + b_t$ be a random variable over b_t and conditioned on $\tilde{\theta}_t$. Privacy loss (at iteration t) can be written as,

$$W_t = \left| \log \frac{\mu_{X_t(\mathcal{D})}(X_t(\mathcal{D}))}{\mu_{X_t(\mathcal{D}')}X_t(\mathcal{D}')} \right| \tag{7}$$

where $\mu_{X_t(\mathcal{D})}(y)$ is the measure of the r.v. $X_t(\mathcal{D})$ induced on $y \in \mathbb{R}$.

1. *Use gaussian noise privacy (with amplification) to get that w.p. at least $1 - \delta$, $W_t \leq \frac{\epsilon}{n\sqrt{\log(1/\delta)}}$.*
2. *Use* **strong composition** *to get the final result for n^2 iterations.*
3. *Next: Pure privacy.*

Convex $(\epsilon, 0)$ privacy

Algorithm/Analysis

$\mathcal{A}_{exp-samp}$: Exponential sampling based convex optimization

Input: Data set $\mathcal{D}$ of size n, loss function l, privacy parameter ϵ, and convex set $\mathcal{C}$.

1. $\mathcal{L}(\theta; \mathcal{D}) = \sum_{i=1}^{n} l(\theta, d_i)$
2. Sample a point θ^{priv} from the convex set $\mathcal{C}$ with probability proportional to $\exp\left(-\frac{\epsilon}{2L||\mathcal{C}||_2}\mathcal{L}(\theta; \mathcal{D})\right)$ and output.

Convex $(\epsilon, 0)$ privacy

Algorithm/Analysis

$\mathcal{A}_{exp-samp}$: Exponential sampling based convex optimization

Input: Data set $\mathcal{D}$ of size n, loss function l, privacy parameter ϵ, and convex set $\mathcal{C}$.

1. $\mathcal{L}(\theta; \mathcal{D}) = \sum_{i=1}^{n} l(\theta, d_i)$
2. Sample a point θ^{priv} from the convex set $\mathcal{C}$ with probability proportional to $\exp\left(-\frac{\epsilon}{2L\|\mathcal{C}\|_2}\mathcal{L}(\theta; \mathcal{D})\right)$ and output.

Theorem

$\mathcal{A}_{exp-samp}$ is $\epsilon-$differentially private.

Proof.

Note that step 2. is the same as using $\exp(-\frac{\epsilon}{L\|\mathcal{C}\|_2}(\mathcal{L}(\theta; \mathcal{D}) - \mathcal{L}(\theta_0; \mathcal{D}))$ *for some* $\theta_0 \in \mathcal{C}$. *Sensitivity of* $\mathcal{L}(\theta; \mathcal{D}) - \mathcal{L}(\theta_0; \mathcal{D}) \leq L\|\mathcal{C}\|_2$. *Now use the exponential mechanism proof.* □

Theorem

There is an efficient version of the $\mathcal{A}_{exp-samp}$ algorithm that has the following guarantees.

1. **Privacy:** *The algorithm is ϵ-differentially private*
2. **Utility:** *The output $\theta^{priv} \in \mathcal{C}$ of the algorithm satisfies*

$$\mathcal{E}[\mathcal{L}(\theta^{priv}; \mathcal{D}) - \mathcal{L}(\theta^*; \mathcal{D})] = O\left(\frac{pL||\mathcal{C}||_2}{\epsilon}\right) \tag{8}$$

3. **Running time:** *Assuming $\mathcal{C}$ is in isotropic position, the algorithm runs in time*

$$O(||\mathcal{C}||_2^2 p^3 n^3 \max\{p \log(||\mathcal{C}||_2 pn), \epsilon||\mathcal{C}||_2 n\}) \tag{9}$$

The Rest of the Paper

Key Theoretical Ideas

- For Lipschitz convex, strongly convex, and (ϵ, δ) and $(\epsilon, 0)$, all bounds on the risk $\mathcal{L}(\theta; \mathcal{D}) - \mathcal{L}(\theta^*; \mathcal{D})$ are, with probability at least $1/2, 1/3$, $\Omega(\min(n, f(p, \epsilon, n))$.

The Rest of the Paper

Key Theoretical Ideas

- For Lipschitz convex, strongly convex, and (ϵ, δ) and $(\epsilon, 0)$, all bounds on the risk $\mathcal{L}(\theta; \mathcal{D}) - \mathcal{L}(\theta^*; \mathcal{D})$ are, with probability at least $1/2, 1/3$, $\Omega(\min(n, f(p, \epsilon, n))$.
- Efficient sampling from logconcave distributions over convex sets.

The Rest of the Paper

Key Theoretical Ideas

- For Lipschitz convex, strongly convex, and (ϵ, δ) and $(\epsilon, 0)$, all bounds on the risk $\mathcal{L}(\theta; \mathcal{D}) - \mathcal{L}(\theta^*; \mathcal{D})$ are, with probability at least $1/2, 1/3$, $\Omega(\min(n, f(p, \epsilon, n))$.
- Efficient sampling from logconcave distributions over convex sets.
- Proof of crazy runtime efficiency result.

Summary

- Differential privacy is another way of looking at similar problems related to stability and learning.

Summary

- Differential privacy is another way of looking at similar problems related to stability and learning.
- It implies a certain bound on the risk of an algorithm.

Summary

- Differential privacy is another way of looking at similar problems related to stability and learning.
- It implies a certain bound on the risk of an algorithm.
- Adding random noise not only provides stability as we've discussed, but also provides privacy guarantees if done right.


Yu-Xiang Wang, Jing Lei, and Stephen E Fienberg.
Learning with differential privacy: Stability, learnability and the sufficiency and necessity of erm principle.
arXiv preprint arXiv:1502.06309, 2015.