On the Geometry of Differential Privacy

Moritz Hardt
Princeton University

Kunal Talwar
MSR SVC

Problem Definition

- Input Database $x \in \mathbb{R}^n$
  - Thought of as Histogram of N people
- Want to compute linear function $Fx \in \mathbb{R}^d$
  - Each entry of F in $\{-1,1\}$
- Output approximation $M(x)$ such that
  - M is differentially private (w.r.t. to l_1 norm)
  - $M(x)$ is close to Fx

Differential Privacy

[DworkMcSherryNissimSmith06]

A mechanism M provides ε−differential privacy if for all $x_1, x_2 \in \mathbb{Z}_+^n$, for any $S \subseteq \mathbb{R}^d$

$$\frac{\Pr[M(x_1) \in S]}{\Pr[M(x_2) \in S]} \leq \exp(\varepsilon |x_1 - x_2|_1)$$

Neighboring Databases: $|x_1 - x_2| \leq 1$. One person changes type from i to j. Output distribution nearly unchanged.

Differential Privacy

[DworkMcSherryNissimSmith06]

A mechanism M provides ε−differential privacy if for all $x_1, x_2 \in \cancel{\mathbb{Z}_+^n} \mathbb{R}^n$, for any $S \subseteq \mathbb{R}^d$

$$\frac{\Pr[M(x_1) \in S]}{\Pr[M(x_2) \in S]} \leq \exp(\varepsilon |x_1 - x_2|_1)$$

Neighboring Databases: $|x_1 - x_2| \leq 1$. One person changes type from i to j. Output distribution nearly unchanged.

Problem Definition

- Input Database $x \in \mathbb{R}^n$
  - Thought of as Histogram of N people
- Want to compute linear function $Fx \in \mathbb{R}^d$
  - Each entry of F in $\{-1,1\}$
- Output distribution $M(x)$ such that
  - M is differentially private (w.r.t. to l_1 norm)
  - $E_M[\,\|Fx - M(x)\|_2\,]$ is as small as possible.

Problem Definition

- Input Database $x \in \mathbb{R}^n$
- Want $Fx \in \mathbb{R}^d$ with $F \in \{-1, +1\}^{d \times n}$
- Define distribution $M(x)$ for every $x \in \mathbb{R}^n$
- Minimize $\max\limits_{x \in \mathbb{R}^n} E_M[\|Fx - M(x)\|_2]$

Let $err(F, M)$ denote the above error

Let $err(F)$ denote $\min\limits_{M \text{ is } \varepsilon\text{DP}} err(F, M)$

Questions

- How big can $err(F)$ be?
  - Universal upper bounds?
  - Lower bounds?

- Given F, what is $err(F)$?

Known results [1 of 2]

[DworkMcSherryNissimSmith06]

Laplacian mechanism gives εDP for any F

Thus $err(F) \leq O\left(d\sqrt{d}\right)$

- i.e. error at most d per coordinate

[BlumLigettRoth08]

Can do with error $\sim O\left(N^{\frac{2}{3}}\, d^{\frac{1}{3}}\right)$ per coordinate

- Better than Laplacian when N is small
- Result more general, d is VC dimension of concept class.

Known results [2 of 2]

[DinurNissim03]

For random F, $err(F) \geq \Omega(d)$

- i.e. need error at least $\sqrt{d}$ per coordinate
- Lower bound applies to essentially any privacy definition

Various extensions [DMT07,DY08,KRY09]

[GhoshRoughgardenSundararajan09]

Laplace noise is optimal for d=1

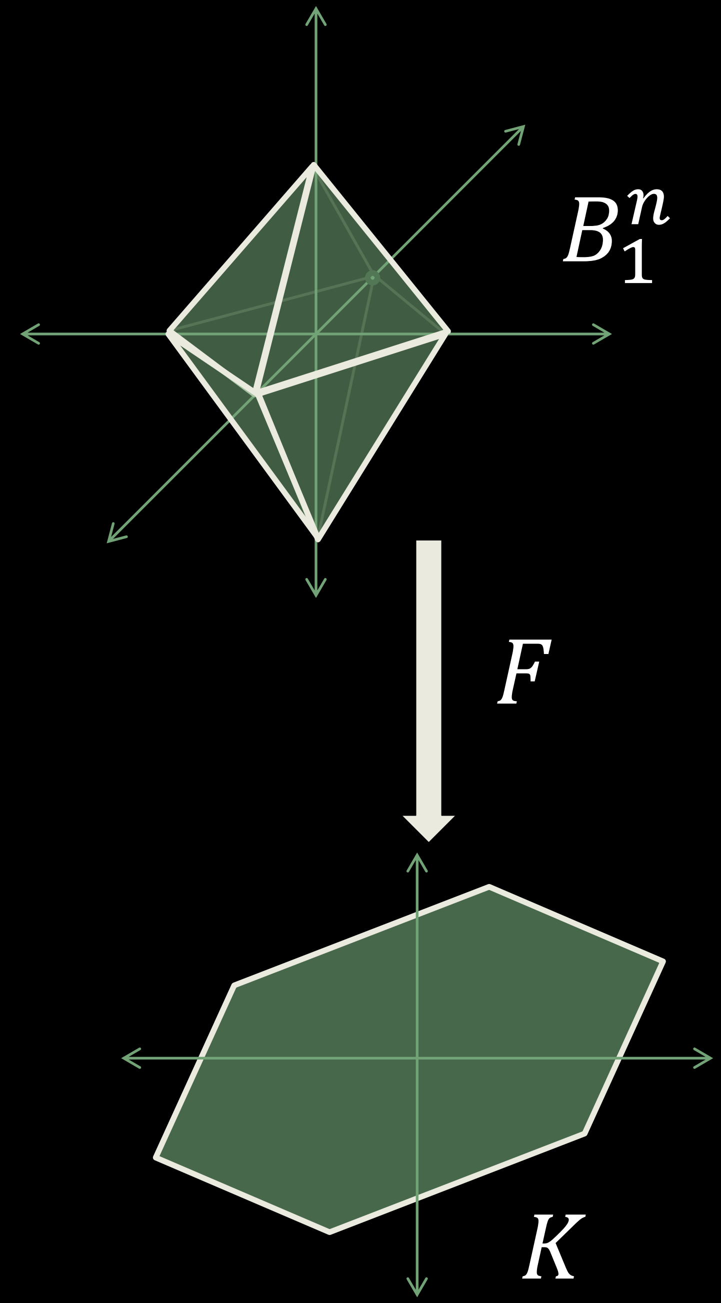
The body K

Given $F \in \{-1,1\}^{d\times n}$

Let $K = FB_1^n$ be the image of the unit l_1 ball under F.

K is symmetric convex hull of columns of F.

We relate $err(F)$ to parameters of K



Results

Lower bound: $err(F) \geq \Omega(d\sqrt{d}\, Vol(K)^{\frac{1}{d}})$

Upper bound: $err(F) \leq O(d\, E_{z \in K}[\|z\|_2])$

For Random $\boldsymbol{F}$ (using [KlartagKozma09]):

$$err(F) \text{ is } \Theta(d) \cdot \min\left(\sqrt{d}, \sqrt{\log \frac{n}{d}}\right)$$

Results

Lower bound: $err(F) \geq \Omega(d\sqrt{d}\, Vol(K)^{\frac{1}{d}})$

Upper bound: $err(F) \leq O(d\, E_{z \in K}[\|z\|_2])$

For Random F:

$$err(F) \text{ is } \Theta(d) \cdot \min\left(\sqrt{d}, \sqrt{\log \frac{n}{d}}\right)$$

– For $d < \log n$, Laplace is optimal for random F

– For $d > \log n$, can do better.

Results

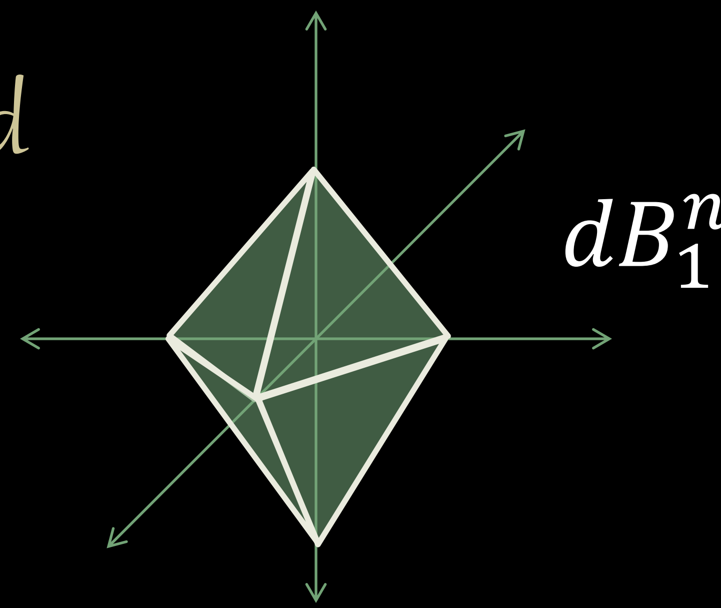
Old conjecture from convex geometry

Assume the Hyperplane conjecture.

Then for any F, we give an εDP mechanism M such that $err(F, M) \leq O\left(\log^{\frac{3}{2}} d\right) \cdot err(F)$

- I.e. we give a $O\left(\log^{\frac{3}{2}} d\right)$ approximation to the best εDP mechanism.
- For specific F, error can be much smaller than lower bounds for random F.

Lower Bound



Basic idea:

Suppose $Vol(K)$ is large, error small.

Then can find $\exp(d)$ points in dK that are mutually far (distance $2r$ from each other).

Let $y_1, \ldots, y_S$ be such a code.

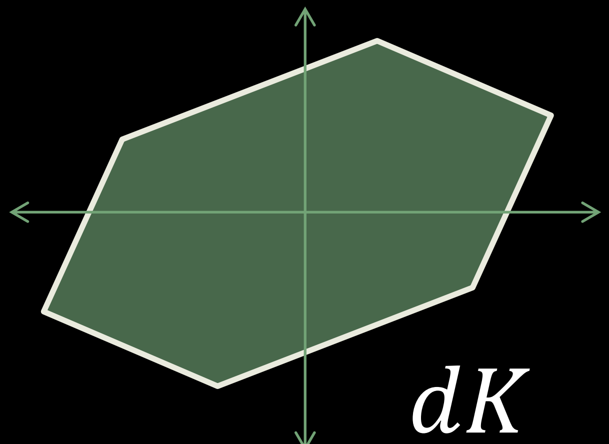
Let $x_1, \ldots, x_S$ be preimages.

By low error $\Pr[M(x_i) \in B(y_i, r)] \geq \frac{1}{2}$

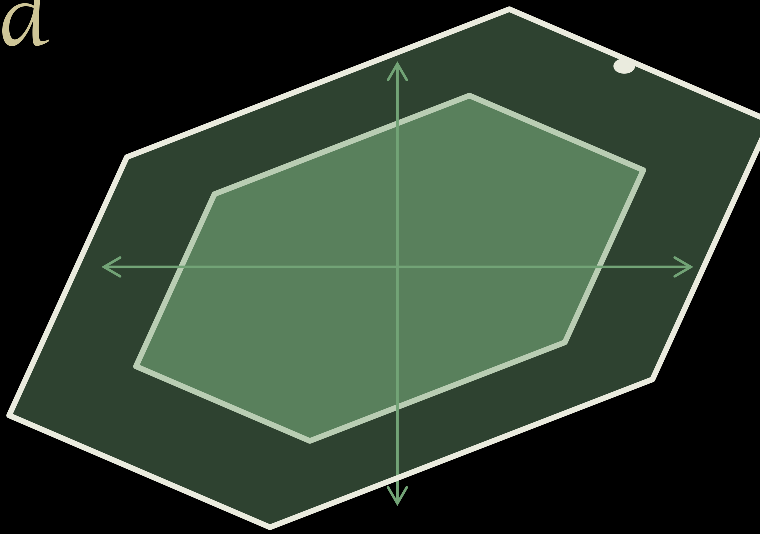
By Privacy:$\Pr\left[M\left(x_j\right) \in B(y_i, r)\right] \geq \frac{\exp(-\varepsilon d)}{2}$

y_i' s far: $\Pr\left[M\left(x_j\right) \in \cup B(y_i, r)\right] \geq \frac{\exp((1-\varepsilon)d)}{2}$

Contradiction!



Upper Bound



Basic Idea: Tailor noise to K

Consider norm $\|\cdot\|_K$

$$\|y\|_K = \min\{\lambda : y \in \lambda K\}$$

By definition $\|Fx_1 - Fx_2\|_K \le |x_1 - x_2|_1$

Use Exponential mechanism [McSherryT.07] M_K:

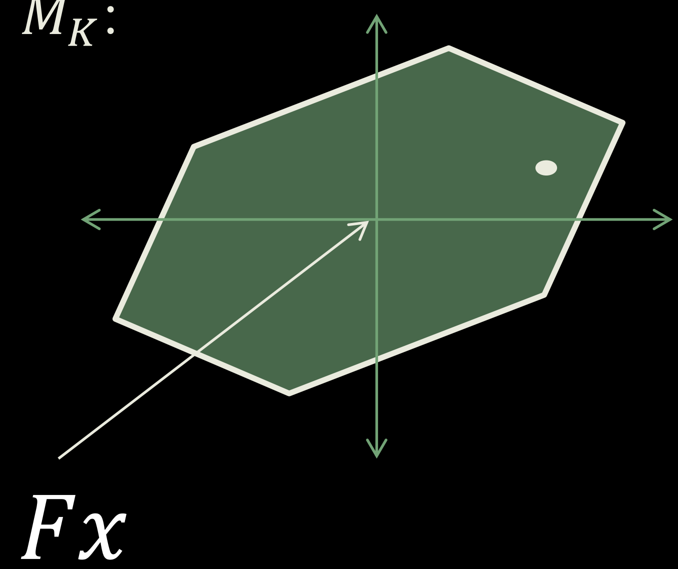
On input $x \in \mathbb{R}^n$

Sample y with prob. $\propto \exp(\varepsilon\|y - Fx\|_K)$

Same as:

Pick r from appropriate distribution

Sample y from $Fx + rK$



Upper Bound

Calculation: $err(F, M_K) \leq O(d\, E_{z \in K}[\|z\|_2])$

Recall lower bound: $err(F) \geq \Omega(d\sqrt{d}\, Vol(K)^{\frac{1}{d}})$

Value for the ball in $\mathbb{R}^d$

Hyperplane conjecture:

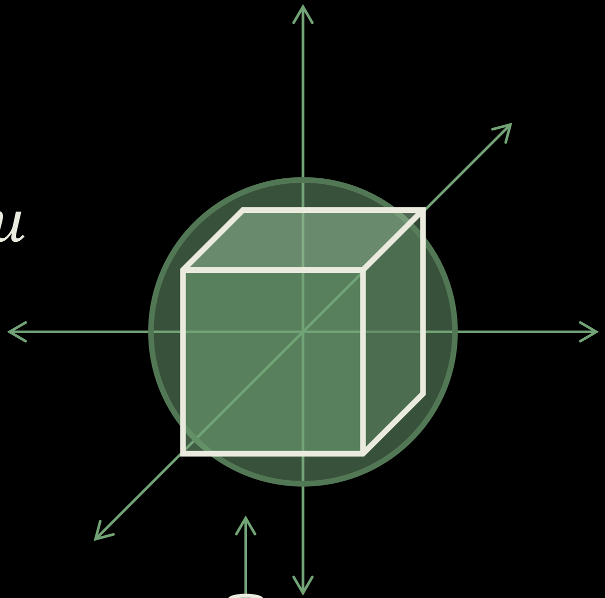
For any isotropic K, $\frac{E_{z \in K}[\|z\|_2]}{Vol(K)^{\frac{1}{d}}}$ is $O(\sqrt{d})$

i.e. $\frac{err(F, M_K)}{err(F)}$ is $O(1)$

Upper Bound

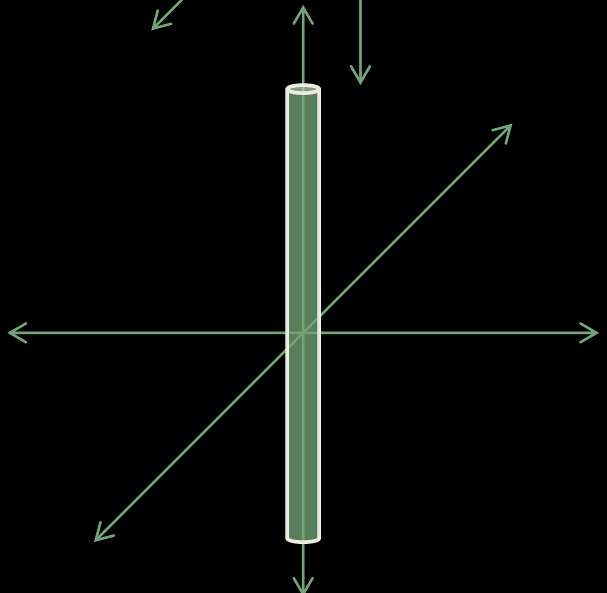
Isotropic: same expected projection in all directions

$E_{z \in K}[\langle z, u \rangle^2]$is the same for all unit vectors u



When K is not isotropic, we can decompose along directions with approximately equal expected projection.

Apply M_K to K restricted to those directions.



In fact very general

- Let K be the (symmetric convex hull of the) set of changes to any function $G\colon D \to \mathbb{R}^d$ that one person can cause.
  $K = symconv\,\{\, G(x) - G(x')\colon x, x'\ neighbours\}$

- Mechanism M_K works for arbitrary G
- Can actually use any $K' \supseteq K$

Efficient Implementation

As defined mechanism M_K requires sampling uniformly from K.

Can get arbitrarily close to the uniform distribution using geometric random walks.

Leads to polynomial time algorithm with the same guarantees.

Polynomial not awesome.

Caveats

- Lower bound applies for small ε and large N. (E.g. $N \approx n^2$ suffices)
- Better mechanisms do exist when N is small.

Linear Program

$$Minimize \quad \mathrm{E}$$

$$\sum_a \mu(x,a) = 1 \qquad \forall x \in \mathbb{R}^n$$

$$\mu(x,a) \geq 0 \qquad \forall x \in \mathbb{R}^n, \forall a \in \mathbb{R}^d$$

$$\mu(x,a) \leq \exp(\varepsilon)\, \mu(x',a) \qquad \forall x, x' \text{ neighbours} \in \mathbb{R}^n, \forall a \in \mathbb{R}^d$$

$$\sum_a \mu(x,a \| a - Fx \|) \leq E \qquad \forall x \in \mathbb{R}^n$$

Conclusions

- Gave new mechanisms and lower bounds for differentially private mechanisms

- Better polynomial running times?
- Better lower bounds/mechanisms for small N?
- Online mechanisms?
- Relaxations of εDP?
- Compute $err(F)$ for specific functions F.