Today

- FriendlyCore for Gaussian Mean estimation
- Beyond Gaussians

FriendlyCore: "general-purpose" aggregator
Can use within subsample & aggregate

Today: apply to Gaussian mean estimation

Theorem There exists an (ε, δ)-DP algorithm that, given n iid samples from $N(\mu, \Pi_d)$, with probability at least 0.99 returns $\tilde{\mu}$ with $\|\tilde{\mu} - \mu\|_2 \leq \alpha$, as long as

$$n \gtrsim \frac{d}{\alpha^2} + \frac{d\sqrt{\log 1/\delta}}{\alpha\varepsilon} + \frac{\log(1/\delta)}{\varepsilon}\Big)$$

where $\gtrsim$ hides universal constant.

Remarks

- lower bound $n \gtrsim \frac{d}{\alpha^2} + \frac{d}{\alpha\varepsilon} + \frac{\log(1/\delta)}{\varepsilon}$
- No dependence on n
- Efficient, nearly linear time
- Some other algorithms with similar (or better) guarantees.

Alg 1 Basic Filter

Input: data $x_1, \dots, x_n$, predicate $f: X \times X \to \{0,1\}$

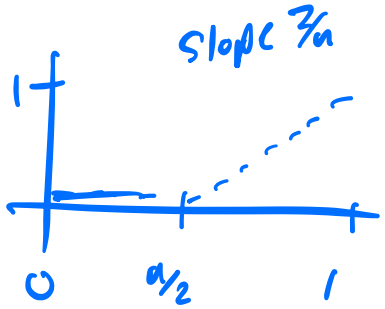
Output: weights $\vec{w} = w_1, \dots, w_n \in [0,1]$

(1) For $i = 1, \dots, n$

(2) $c_i \leftarrow \sum_{j=1}^{n} f(x_i, x_j)$

(3) $w_i \leftarrow \begin{cases} 1 & \text{if } c_i = n \\ 0 & \text{if } c_i \le n/2 \\ \frac{c_i - n/2}{n} & \text{otherwise} \end{cases}$

(4) Return $\vec{w}$



Alg 2 Friendly Avg

Input: data $x_1, \dots, x_n \in \mathbb{R}^d$, $\varepsilon, \delta \in (0,1)$

Output: $\hat{\mu} \in \mathbb{R}^d$ or $\perp$

(1) Predicate $dist(x,y) \triangleq \mathbb{1}\{\|x-y\|_2 \le 10\sqrt{d}\}$

(3) $\vec{w} \leftarrow \text{BasicFilter}(x, dist)$

(4) If $\|\vec{w}\|_1 + Lap(1/\varepsilon) \le n - \frac{2\log 1/\delta}{\varepsilon}$

(5) Return $\perp$

(6) Else

(7) Return $\frac{1}{n}\sum_{i=1}^{n} w_i x_i + \mathcal{N}\left(0, \frac{1000d}{n^2} \cdot \sigma_{\varepsilon,\delta}^2 \cdot \mathbb{I}_d\right)$

Talk through accuracy proof

Talk through privacy proof

- What do we need to show Gaussian is DP?
- Laplace noise?

Claim 1 Suppose distributions P & Q are supported in an ℓ_2 ball of radius R. Then

$$\|\mu(P) - \mu(Q)\|_2 \le 2R \cdot TV(P,Q)$$

Proof Rescale so $supp(P), supp(Q) \subseteq B(0,R)$.

Then $$\left\| \mathop{\mathbb{E}}_{x\sim P}[x] - \mathop{\mathbb{E}}_{y\sim Q}[y] \right\|_2 = \left\| \sum_x (xp(x) - xq(x)) \right\|_2$$

$$\le \sum_x \|x\|_2 \cdot |p(x) - q(x)|$$

$$\le R \cdot \|p - Q\|_1 = 2R \cdot TV(P,Q)$$ ☒

Claim 2 Let $w, w' \in [0,1]^n$ satisfy

i) $\|w - w'\|_1 \le a$

ii) $\|w\|_1, \|w'\|_1 \ge n - b$.

Then $$TV\left(\frac{w}{\|w\|_1}, \frac{w'}{\|w'\|_1}\right) \le \frac{a}{n-b}$$

Proof as exercise.

Abstract three claims about filter.

Completeness if all x_i, x_j satisfy $f(x_i, x_j) = 1$, then $\vec{w} = \vec{1}$

Soundness if $w_i, w_j > 0$, then $\|x_i - x_j\|_2 \leq 20\sqrt{d}$.

Stability If $x \sim x'$, then $\|w - w'\|_1 \leq 3$

Proof Suppose $x_1 \neq x_1'$. Then for any $i \neq 1$,

$$|c_i - c_i'| = \left|\sum_{j=1}^{n} f(x_i, x_j) - f(x_i', x_j')\right|$$

$$= |f(x_i, x_1) - f(x_i, x_i')| \leq 1$$

Then

$$\|w - w'\|_1 = \sum_{i=1}^{n} |w_i - w_i'|$$

$$\leq 1 + \sum_{i=2}^{n} |w_i - w_i'|$$

$$\leq 1 + \sum_{i=1}^{n} \frac{2}{n} = 3$$

$\boxtimes$

Then adding Laplace noise & releasing

$\|w\|_1 + \text{Lap}(3/\varepsilon)$ is ε-DP, as

$$\left| \|w\|_1 - \|w'\|_1 \right| = \left| \sum_i w_i - w_i' \right|$$

$$\leq \sum_i |w_i - w_i'| = \|w - w'\|_1$$

Now to Gaussian noise. Want to show, wp $1-\delta$, that

$$\left\| \frac{1}{n} \sum_i w_i x_i - \frac{1}{n} \sum_i w_i' x_i' \right\|_2 \leq \frac{\sqrt{d}}{n}$$

Then the guarantees of Gaussian mechanism will apply.

Similar to Claim 1.

Claim 3 Suppose $w \neq \vec{0}$, $w' \neq \vec{0}$.

Then $\left\| \frac{1}{n} \sum_i w_i x_i - \frac{1}{n} \sum_i w_i' x_i' \right\|_2 \leq \frac{\sqrt{d}}{n}$.

Proof Can shift all x_i's so that they all lie in ball around $\vec{0}$ of radius $20\sqrt{d}$, by soundness property.

Then, suppose $x_1 \neq x_1'$, and we have

$$\left\| \frac{1}{n} \sum_{i=1}^{n} w_i x_i - \frac{1}{n} \sum_{i=1}^{n} w_i' x_i' \right\|_2$$

$$= \frac{1}{n} \left\| w_1 x_1 - w_1' x_1' + \sum_{i=2}^{n} (w_i - w_i') x_i \right\|_2$$

$$\leq \frac{1}{n} \left(40\sqrt{d} + \left\| \sum_{i=2}^{n} (w_i - w_i') x_i \right\|_2 \right)$$

And, by the triangle inequality

$$\left\| \sum_{i=2}^{n} (w_i - w_i') x_i \right\|_2 \leq \sum_{i=2}^{n} |w_i - w_i'| \cdot \|x_i\|_2$$

$$\leq 20\sqrt{d} \cdot \|w - w'\|_1 .$$

$$= O(\sqrt{d})$$

□

Finally, an important issue:

Claim 3 relies on the soundness property of the weights. This only gives guarantees for points which recieve weight. Thus, if our test of $\|w\|_1$ passes but the weights are actually zero, we could be in big trouble.

However, this happens only with probability at most $1-\delta$, so it

gets wrapped up in the privacy parameters.