Lecture 20: Robustness-to-Privacy

CS 839

March 27, 2026

Plan for Today

- ☐ Review: (Smoothed) Inverse Sensitivity Mechanism
- ☐ Robust Estimation
- ☐ Privacy Induces Robustness
- ☐ Robustness Implies Privacy
- ☐ Discussion/Questions

Inverse Sensitivity Mechanism

Definition

For a function $f : \mathcal{X}^n \to \mathcal{Y}$, the *inverse sensitivity mechanism* A_{inv} samples from

$$p(y) \propto \exp\left(-\frac{\varepsilon}{2} \cdot \text{len}_f(y; x)\right)$$

where

$$\text{len}_f(y; x) = \inf_{x'} \left\{ d_H(x, x') \mid f(x') = y \right\}$$

Here $d_H(x, x')$ is Hamming distance.

Important questions we will ignore:

- How can we compute $\text{len}_f(y; x)$?
- How can we sample from $p(y)$?

Local Modulus of Continuity

We've seen *local sensitivity*:

$$\mathrm{LS}_f(x) = \sup_{x' \sim x} \left\| f(x) - f(x') \right\|$$

$$\mathrm{LS}_f^k(x) = \sup_{\substack{x' \\ d_H(x,x') \leq k}} \left\| f(x) - f(x') \right\|.$$

This has another name.

Definition

The *local modulus of continuity* of f at x is, for $k \in \mathbb{N}$,

$$\omega_f(x; k) = \sup_{\substack{x' \\ d_H(x,x') \leq k}} \left\| f(x) - f(x') \right\|.$$

Utility for Inverse Sensitivity, Discrete Case

Let $\mathrm{diam}(\mathcal{Y}) = \sup_{u,v \in \mathcal{Y}} \|u - v\|$.

Theorem

For $f : \mathcal{X}^n \to \mathcal{Y}$, on any input x, with probability $1 - \beta$ A_{inv} returns $\tilde{y}$ such that

$$\|\tilde{y} - f(x)\| \leq \omega_f\left(x; \frac{2}{\varepsilon} \cdot \log\left(\frac{2|\mathcal{Y}|\mathrm{diam}(\mathcal{Y})}{\beta\varepsilon}\right)\right).$$

Vacuous guarantee when $|\mathcal{Y}| = +\infty$, which is what we usually care about. Two options to proceed:

- Can assume more about the distribution/estimator.
- Can modify the mechanism.

Smoothed Inverse Sensitivity Mechanism

Definition

For a function $f : \mathcal{X}^n \to \mathcal{Y}$ and $\rho \geq 0$, the *smoothed inverse sensitivity mechanism* A^{ρ}_{inv} samples from

$$p(y) \propto \exp\left(-\frac{\varepsilon}{2} \cdot \text{len}^{\rho}_{f}(y; x)\right)$$

where

$$\text{len}^{\rho}_{f}(y; x) = \inf_{x'} \left\{ d_H(x, x') \mid \left\| f(x') - y \right\| \leq \rho \right\}$$

Utility for Smoothed Inverse Sensitivity

Theorem (Utility for Smoothed Inverse)

For $f : \mathcal{X}^n \to \mathcal{Y} \subseteq B(0, R)$, on any input x, with probability $1 - \beta$ A^{ρ}_{inv} returns $\tilde{y}$ such that

$$\|\tilde{y} - f(x)\| \leq \omega_f\left(x; \frac{2}{\varepsilon} \cdot (d \log(R/\rho + 1) + \log(1/\beta))\right).$$

Proof is very similar to analysis on Homework 2.

Plan for Today

- ✓ Review: (Smoothed) Inverse Sensitivity Mechanism
- ☐ Robust Estimation
- ☐ Privacy Induces Robustness
- ☐ Robustness Implies Privacy
- ☐ Discussion/Questions

Robust Estimation

Well-studied definition that captures insensitivity to model misspecification or data corruptions, even adversarial ones.

Definition

Fix a distribution P and population quantity $\mu(P)$. Algorithm A is a (τ, β, α)-*robust estimator* for $\mu(P)$ if with probability $1 - \beta$ over the draw of $x \sim P^{\otimes n}$ for all x' that differ from x in at most $n\tau$ entries we have

$$\|A(x') - \mu\| \leq \alpha.$$

Today, will assume the robust estimator is deterministic.

Prior Work on Robustness and DP

Almost immediately, privacy researchers knew that:

1. DP algorithms are automatically robust to low levels of adversarial corruptions, and
2. Robust algorithms tend to be useful tools for designing DP algorithms.

A formal version of (1) is folklore, but its full implications are still being explored.

Until recently, we lacked general tools to operationalize (2).

Plan for Today

- ✓ Review: (Smoothed) Inverse Sensitivity Mechanism
- ✓ Robust Estimation
- ☐ Privacy Induces Robustness
- ☐ Robustness Implies Privacy
- ☐ Discussion/Questions

Privacy to Robustness

The output of DP algorithms cannot change much if you change a single point.
This applies to changing many points.

Lemma (Group Privacy)

Suppose A is ε-DP. If datasets x and x' satisfy $d_H(x, x') \leq k$, then

$$A(x) \approx_{k\varepsilon} A(x').$$

I.e., for any event E, $\Pr[A(x) \in E] \leq e^{k\varepsilon} \Pr[A(x') \in E]$.

Privacy to Robustness

Theorem

Suppose A is ε-DP and satisfies: for $x \sim P^{\otimes n}$, with probability at least $1-\beta$ we have $\|A(x) - \mu(P)\| \leq \alpha$.
For any $\gamma \in (0,1)$, let $\tau = \frac{\log(1/\gamma)}{n\varepsilon}$. Then A is $(\tau, \beta/\gamma, \alpha)$-robust.

Proof.

Let $W = \{y \mid \|y - u\| > \alpha\}$. For $x \sim P^{\otimes n}$ and x' with $d_H(x, x') \leq \tau n$:

$$\begin{aligned}\Pr[A(x') \in W] &\leq e^{\tau n\varepsilon} \Pr[A(x) \in W] \\ &= \exp\left(\frac{\log(1/\gamma)}{n\varepsilon} \cdot n\varepsilon\right) \cdot \Pr[A(x) \in W].\end{aligned}$$

Finally, $\Pr[A(x) \in W] \leq \beta$ by assumption. □

Plan for Today

- ✓ Review: (Smoothed) Inverse Sensitivity Mechanism
- ✓ Robust Estimation
- ✓ Privacy Induces Robustness
- ☐ Robustness Implies Privacy
- ☐ Discussion/Questions

Robust-to-Private

Concurrent works of Hopkins, Kamath, Majid, Narayanan (2023) and Asi, Ullman, and Zakynthinou (2023).

Theorem

Let $A : \mathcal{X}^n \to \mathcal{Y} \subseteq B(0, R)$ *be a* (τ, β, α)*-robust estimator for* $\mu(P)$*. If*

$$n \geq \frac{2}{\tau\varepsilon}(d \log(R/\alpha + 1) + \log(1/\beta))$$

With probability at least $1 - 2\beta$*,* A^{ρ}_{inv} *with* $\rho = \alpha$ *returns* $\tilde{y}$ *such that,*

$$\|\tilde{y} - \mu\| \leq 4\alpha.$$

Proof on Board

Theorem (Utility for Smoothed Inverse)

For $f : \mathcal{X}^n \to \mathcal{Y}$, on any input x, with probability $1 - \beta$ A^{ρ}_{inv} returns $\tilde{y}$ such that

$$\|\tilde{y} - f(x)\| \leq \omega_f\left(x; \frac{2}{\varepsilon} \cdot (d \log(R/\rho + 1) + \log(1/\beta))\right).$$

Theorem

Let $A : \mathcal{X}^n \to \mathcal{Y} \subseteq B(0, R)$ be a (τ, β, α)-robust estimator for $\mu(P)$. If

$$n \geq \frac{2}{\tau\varepsilon}(d \log(R/\alpha_0 + 1) + \log(1/\beta))$$

With probability at least $1 - 2\beta$, A^{ρ}_{inv} with $\rho = \alpha$ returns $\tilde{y}$ such that,

$$\|\tilde{y} - \mu\| \leq 4\alpha.$$

Plan for Today

- ✓ Review: (Smoothed) Inverse Sensitivity Mechanism
- ✓ Robust Estimation
- ✓ Privacy Induces Robustness
- ✓ Robustness Implies Privacy
- ☐ Discussion/Questions

Today's Discussion

We saw, informally,

- DP algorithms are automatically somewhat robust
- Robust algorithms imply the existence of DP algorithms

Remember: the details matter! This topic is far from settled.

Our discussion was restricted to

- Pure DP
- Information-theoretic (no talk of computation)
- Parameter estimation in a fixed norm

Some Open Directions

Surprising observation: for many tasks, the best-known DP algorithm is *also* robust.

Consider the following simpler "robustness-to-privacy" transformation: for robust A, run $A(x)$ and add Gaussian noise. This doesn't work, but it's not obviously a dead-end algorithmically, either.

Many questions about sample and computational efficiency for key tasks.

Further Reading

- Hopkins, S. B., Kamath, G., Majid, M., & Narayanan, S. Robustness implies privacy in statistical estimation. STOC 2023.
- Asi, H., Ullman, J., & Zakynthinou, L. From Robustness to Privacy and Back. ICML 2023.
- Georgiev, K., & Hopkins, S. Privacy induces robustness: Information-computation gaps and sparse mean estimation. NeurIPS 2022.
- Dwork, C., & Lei, J. Differential privacy and robust statistics. STOC 2009
- Asi, H., & Duchi, J. C. Instance-optimality in differential privacy via approximate inverse sensitivity mechanisms. NeurIPS 2020.
- Diakonikolas, I., & Kane, D. M. Algorithmic high-dimensional robust statistics. Cambridge university press, 2023.
- Kamath, Gautam. The broader landscape of robustness in algorithmic statistics. IEEE BITS the Information Theory Magazine (2025).