Subspace Differential Privacy

Ruobin Gong, Department of Statistics, Rutgers University
Jie Gao, Department of Computer Science, Rutgers University
Fang-Yi Yu, Harvard School of Engineering

Summary

When sanitizing data products, data curators need to respect certain **invariants** (data-dependent constraints). Invariants challenge the **formulation**, **implementation** and **interpretation** of privacy guarantees.

Proposal:

- **Subspace differential privacy**: to honestly characterize the dependence of the sanitized output on confidential aspects of the data;
- Two design frameworks (*projection* and *extension*) to revise existing DP mechanisms to induced-subspace DP ones.

Benefits:

- No need to ``post-process'' to impose invariants;
- Preserves transparency and statistical intelligibility of the output;
- Suitable for distributed privatization implementation.

Invariants

Definition Given a query $A: X^* \to \mathbb{R}^n$, and $C \in \mathbb{R}^{n_c \times n}$ be a $n_c \times n$ matrix with rank $n_c < n$. A (random) mechanism $M: X^* \to Y \subseteq \mathbb{R}^n$ satisfies the *linear equality invariant C with query A*, if for all x,

$$CM(x) = CA(x)$$

with probability one over the randomness of M.

2020 census data

- population totals
- counts of total housing units
- group quarter and facilities

A common method to impose invariants is via "post-processing" using optimization/distance minimization, e.g. Census TopDown.

- Not differentially private anymore
- Systematic bias and obscurity

Invariants and bias

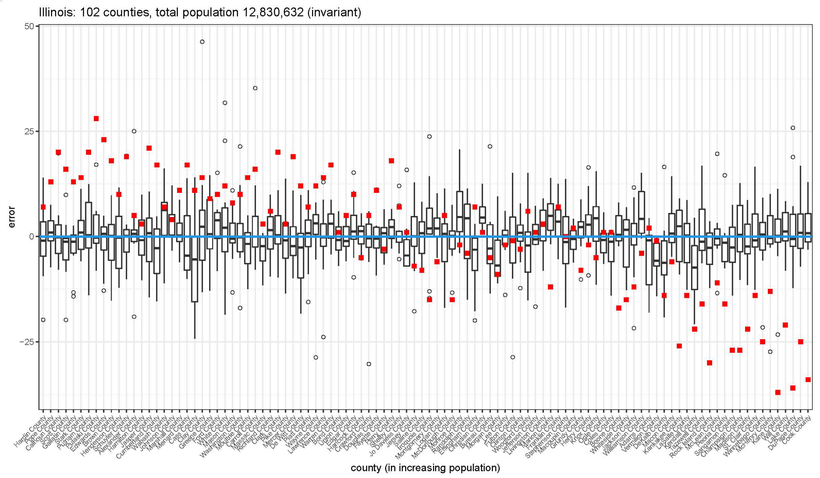

Figure: Census DAS associates positive errors (red squares) with larger counties and negative errors with smaller counties, when the state population is held as invariant. Pictured are counties of Illinois in increasing true population sizes.

Invariants and privacy

Example. Let x be a database with domain $X = \{1,2,3,4\}$, A be the histogram query, and $C = (1,0,1,0)$. That is, the curator should accurately report the total number of individuals with odd features.

Suppose M satisfies the linear equality invariant C with query A. Consider two neighboring databases $x = (1,2,4)$ and $x' = (2,2,4)$. For $S = \{(y_1, y_2, y_3, y_4): y_1 + y_3 = 1\} \subset \mathbb{R}^4$, the probability ratio is unbounded:

$$\frac{\Pr[M(x) \in S]}{\Pr[M(x') \in S]} = \infty,$$

because $\Pr[M(x) \in S] = 1$ but $\Pr[M(x') \in S] = 0$. Thus, M violates (ϵ, δ)-DP for any $\epsilon > 0$ and $\delta < 1$.

Definition Given a query $A: X^* \to \mathbb{R}^n$, and $C \in \mathbb{R}^{n_c \times n}$ be a $n_c \times n$ matrix with rank $n_c < n$ with null space $N := \{v: Cv = 0\}$. A mechanism $M: X^* \to \mathbb{R}^n$ is **(ϵ, δ)-induced subspace differentially private** for query A and an invariant C, if

- $\Pr[\Pi_N M(x) \in S] \leq e^{\epsilon} \Pr[\Pi_N M(x') \in S] + \delta$ for all adjacent databases x, x', and $S \subseteq N$, and
- M satisfies the linear equality invariant C, i.e.

$$\Pr[CM(x) = CA(x)] = 1.$$

Design private mechanisms

Two approaches

Projection framework

- Converting an existing DP mechanism M to ISDP

$$\mathcal{M}(x) := A(x) + \Pi_N(M(x) - A(x))$$

- Project the noise into null space

Extension framework

- Choose a DP mechanism $\widehat{M}$ for query $\Pi_N A(x)$

$$\mathcal{M}(x) := \Pi_R A(x) + \widehat{M}(x)$$

- Augmenting a smaller private query invariant-compatibly

Revising additive DP mechanisms

If the original (ϵ, δ)-DP mechanism M is additive:

$$M(x) = A(x) + e,$$

where the noise e is independent of $A(x)$, e.g., Gaussian mechanism.

- Projected Gaussian

$$A(x) + \Pi_N e$$

the variance of e is of order $\Delta_2(A)$

- Extended Gaussian

$$A(x) + Q_N e$$

  - Q_N is a rotation matrix of N
  - the variance of e is of order $\Delta_2(Q_N^\top A)$

Discussion

- Optimality
  - optimal DP for query $\Pi_N A$ = optimal ISDP for A and invariant C
  - Optimal ISDP from the correlated Gaussian mechanism (Nikolov et al 13)
- Unbiasedness
  - Projected and extended Gaussian/Laplace mechanism are unbiased
- Transparency and statistical intelligibility

Future work

- General invariants
  - Inequality
  - Discrete output space
- Trade off between utility and privacy