E Chen[1][0009−0006−3013−8790] Yang Cao[2][1111−2222−3333−4444]
Gennian Ge[1,3]

2023 年 5 月 4 日

Abstract. The abstract should briefly summarize the contents of the paper in 15–250 words.

Keywords: First keyword · Second keyword · Another keyword.

1 Introduction

Definition 1. *For a domain $\mathcal{D}$, let $\mathcal{R}^{(i)} : \mathcal{S}^{(1)} \times \mathcal{S}^{(2)} \times \cdots \times \mathcal{S}^{(i-1)} \times \mathcal{D} \to \mathcal{S}^{(i)}$ for $i \in [n]$, where $\mathcal{S}^{(i)}$ is the range space of $\mathcal{R}^{(i)}$ be a sequence of algorithms such that $\mathcal{R}^{(i)}(z_{1:i-1}, \cdot)$ is an (ϵ_i, δ_i)-DP local randomizer for all values of auxiliary inputs $z_{1:i-1} \in \mathcal{S}^{(1)} \times \mathcal{S}^{(2)} \times \cdots \mathcal{S}^{(i-1)}$. Let $\mathcal{A}_R : \mathcal{D} \to \mathcal{S}^{(1)} \times \mathcal{S}^{(2)} \times \cdots \times \mathcal{S}^{(n)}$ be the algorithm that given a dataset $x_{1:n} \in \mathcal{D}^n$, then sequentially computes $z_i = \mathcal{R}^{(i)}(z_{1:i-1}, x_i)$ for $i \in [n]$ and outputs $z_{1:n}$. We say $\mathcal{A}_R(\mathcal{D})$ is an (ϵ_i, δ_i)-DP adaptive process. Similarly, if we first sample a permutation π uniformly at random, then sequentially computes $z_i = \mathcal{R}^{(i)}(z_{1:i-1}, x_{\pi_i})$ for $i \in [n]$ and outputs $z_{1:n}$, we say this process is shuffled (ϵ_0, δ_0)-DP adaptive and denote it by $\mathcal{A}_{R,S}(\mathcal{D})$.*

Lemma 1 (KOV15). *Let $R^{(i)} : D \to S$ be an (ϵ_i, δ_i)-DP algorithm, and $x_0, x_1 \in D$, then there exists two discrete random variables $\tilde{X}_0$ and $\tilde{X}_1$, where*

$$P(\tilde{X}_0 = x) = \begin{cases} \delta_i & \text{if } x = A, \\ \frac{(1-\delta_i)e^{\epsilon_i}}{1+e^{\epsilon_i}} & \text{if } x = 0, \\ \frac{1-\delta_i}{1+e^{\epsilon_i}} & \text{if } x = 1, \\ 0 & \text{if } x = B, \end{cases}$$

⋆Corresponding author

and

$$P(\tilde{X_1} = x) = \begin{cases} 0 & \text{if } x = A, \\ \frac{1-\delta_i}{1+e^{\epsilon_i}} & \text{if } x = 0, \\ \frac{(1-\delta_i)e^{\epsilon_i}}{1+e^{\epsilon_i}} & \text{if } x = 1, \\ \delta_i & \text{if } x = B. \end{cases}$$

The (ϵ_i, δ_i)-differentially private mechanism

Lemma 2. *Given an (ϵ_i, δ_i)-DP adaptive process, then in the i-th step, local randomizer $\mathcal{R}^{(i)}$: $\mathcal{D} \to \mathcal{S}$ and for any $n+1$ inputs $x_1^0, x_1^1, x_2, \cdots, x_n \in \mathcal{D}$, there exists distributions $\mathcal{Q}_1^0, \mathcal{Q}_1^1, \mathcal{Q}_1, \mathcal{Q}_2, \cdots, \mathcal{Q}_n$ such that*

$$\mathcal{R}^{(i)}(x_1^0) = \frac{(1-\delta_i)e^{\epsilon_i}}{1+e^{\epsilon_i}}\mathcal{W}_1^0 + \frac{1-\delta_i}{1+e^{\epsilon_i}}\mathcal{W}_1^1 + \delta_i\mathcal{W}_1,$$

$$\mathcal{R}^{(i)}(x_1^1) = \frac{(1-\delta_i)e^{\epsilon_i}}{1+e^{\epsilon_i}}\mathcal{W}_1^0 + \frac{1-\delta_i}{1+e^{\epsilon_i}}\mathcal{W}_1^1 + \delta_i\mathcal{W}_1.$$

$$\forall x_i \in \{x_2, \cdots, x_n\}, \mathcal{R}(x_i) = \frac{1-\delta_i}{1+e^{\epsilon_i}}\mathcal{Q}_1^0 + \frac{1-\delta_i}{1+e^{\epsilon_i}}\mathcal{Q}_1^1 + (1 - \frac{2(1-\delta_i)}{1+e^{\epsilon_i}})\mathcal{Q}^i$$

Proof. For inputs $X_0 = \{x_1^0, x_2, \ldots, x_n\}$ and $X_1 = \{x_1^1, x_2, \ldots, x_n\}$, $\mathcal{R}^{(i)}$ satisfies the constraints of Lemma **??**, so there exists an (ϵ_i, δ_i)-DP local randomizer $\mathcal{R}' : \mathcal{D} \to \mathcal{Z}$ for the i-th output, and post-processing function $proc(\cdots)$ such that $proc(\mathcal{R}'^{(i)}(x)) = \mathcal{R}^{(i)}(x)$, and

$$P(\mathcal{R}'^{(i)}(x_1^0) = z) = \begin{cases} 0 & \text{if } z = A, \\ \frac{1-\delta_i}{1+e^{\epsilon_i}} & \text{if } z = 0, \\ \frac{(1-\delta_i)e^{\epsilon_i}}{1+e^{\epsilon_i}} & \text{if } z = 1, \\ \delta_i & \text{if } z = B. \end{cases}$$

$$P(\mathcal{R}'^{(i)}(x_1^1) = z) = \begin{cases} \delta_i & \text{if } z = A, \\ \frac{1-\delta_i}{1+e^{\epsilon_i}} & \text{if } z = 0, \\ \frac{(1-\delta_i)e^{\epsilon_i}}{1+e^{\epsilon_i}} & \text{if } z = 1, \\ 0 & \text{if } z = B. \end{cases}$$

Let $L = \{z \in \mathcal{Z} | \mathbb{P}(\mathcal{R}'(x_1^0 = z)) = \frac{(1-\delta_i)e^{\epsilon_i}}{1+e^{\epsilon_i}}$ and $\mathbb{P}(\mathcal{R}'(x_1^1 = z)) = \frac{1-\delta_i}{1+e^{\epsilon_i}}\}$, $U = \{z \in \mathcal{Z} | \mathbb{P}(\mathcal{R}'(x_1^1 = z)) = \frac{1-\delta_i}{1+e^{\epsilon_i}}$ and $\mathbb{P}(\mathcal{R}'(x_1^1 = z)) = \frac{(1-\delta_i)e^{\epsilon_i}}{1+e^{\epsilon_i}}\}$. Let $M = \mathcal{Z}\{L \bigcup U\}$ and $p = \sum_{z \in \mathcal{L}} p_z = \sum_{z \in \mathcal{U}} p_z$. Since conditioned on the output

lying in $\mathcal{L}$, the distribution of $\mathcal{R}'(x_1^0)$ and $\mathcal{R}'(x_1^1)$ are same. Let $\mathcal{W}_1^0 = \mathcal{R}'(x_1^0)|L = \mathcal{R}'(x_1^1)|L$, $\mathcal{W}_1^1 = \mathcal{R}'(x_1^0)|U = \mathcal{R}'(x_1^1)|U$ and $\mathcal{W}_1 = \mathcal{R}'(x_1^0)|M = \mathcal{R}'(x_1^1)|M$. Then

$$\mathcal{R}'(x_1^0) = \frac{(1-\delta_i)e^{\epsilon_i}}{1+e^{\epsilon_i}}\mathcal{W}_1^0 + \frac{1-\delta_i}{1+e^{\epsilon_i}}\mathcal{W}_1^1 + \delta_i\mathcal{W}_1,$$

$$\mathcal{R}'(x_1^1) = \frac{(1-\delta_i)e^{\epsilon_i}}{1+e^{\epsilon_i}}\mathcal{W}_1^0 + \frac{1-\delta_i}{1+e^{\epsilon_i}}\mathcal{W}_1^1 + \delta_i\mathcal{W}_1.$$

Further, for all $x_i \in \{x_2, \cdots, x_n\}$,

$$\mathcal{R}'(x_i) \geq \frac{1-\delta_i}{1+e^{\epsilon_i}}\mathcal{W}_1^0 + \frac{1-\delta_i}{1+e^{\epsilon_i}}\mathcal{W}_1^1 + (1 - \frac{2(1-\delta_i)}{1+e^{\epsilon_i}})\mathcal{W}_i.$$

Letting $\mathcal{Q}_1^0 = proc(\mathcal{W}_1^0)$, $\mathcal{Q}_1^1 = proc(\mathcal{W}_1^1)$, $\mathcal{Q}_1 = proc(\mathcal{W}_1)$ and for all $i \in \{2, \cdots, n\}$, $\mathcal{Q}_i = proc(\mathcal{W}_i)$.

Theorem 1. *For a domain D, if $\mathcal{A}_R(\mathcal{D})$ is an shuffled (ϵ_i, δ_i)-DP adaptive process, then there exists a post-processing function $proc(\cdot)$: $(0,1,2) \to \mathcal{S}^{(1)} \times \mathcal{S}^{(2)} \times \cdots \times \mathcal{S}^{(n)}$, such that*

$$T(\mathcal{A}_{R,S}(X_0), \mathcal{A}_{R,S}(X_1)) = T(proc(P_0), proc(P_1)),$$

where $(X_0, X_1, X_2) \sim P_0$, $X_0 = \sum_{i=1}^{n-1} Bern(\frac{1-\delta_i}{1+e_i^\epsilon})$, $X_1 = \sum_{i=1}^{n-1} Bern(\frac{1-\delta_i}{1+e_i^\epsilon})$, $X_2 = \sum_{i=1}^{n-1} Bern(1 - \frac{2(1-\delta_i)}{1+e_i^\epsilon})$.

Proof. According to Lemma **??**, for arbitrary

注意到多项分布趋近于正态分布，但是趋近速率使用Berry-Esseen bound只能达到$O(1/\sqrt{n})$, 这相对于洗牌模型, 误差会很大. 因此, 这边使用连续修正的方法给出$O(1/n)$的收敛速率.对于二项分布和正态分布的逼近速率, 见[DRS22].

$$\mathbb{P}(X = k) = \int_{k-0.5}^{k+0.5} \frac{1}{\sqrt{2\pi np(1-p)}} e^{-\frac{(x-np)^2}{2np(1-p)}} dx$$

对于多项分布$Multinom(n; p/2, p/2, 1-p)$和对应的正态分布$N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, 有

$$\boldsymbol{\mu} = n(p/2, p/2, 1-p)$$

,

$$\tilde{\boldsymbol{\Sigma}} = n \begin{pmatrix} \frac{p}{2}(1-\frac{p}{2}) & -\frac{p^2}{4} & -\frac{p(1-p)}{2} \\ -\frac{p^2}{4} & \frac{p}{2}(1-\frac{p}{2}) & -\frac{p(1-p)}{2} \\ -\frac{p(1-p)}{2} & -\frac{p(1-p)}{2} & p(1-p) \end{pmatrix}$$

又因为根据post-processing的相关性质知, 考虑前两个分量的分布即可. 因此下面考虑(X_1, X_2)同正态分布的全变差即可, 根据公式可以得到对应的界.

$$\alpha + \beta \geq 1 - TV(P, Q) \tag{1}$$

数值结果已经表明, 多项分布以$O(1/n)$的速率趋近正态, 下面给出理论证明.

Lemma 3. *Let* $P \sim Multinom(n; p/2, p/2, 1-p)$, $Q \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, 那么

$$\|P - Q\|_{TV} = O(1/n)$$

, 其中$\|P(A) - Q(A)\|_{TV}$ 表示*supremum of* $|P(A) - Q(A)|$ *over all measurable sets* A.

Lemma 4 (参照Lemma E.2, 67-73). *Let* $F_n(x) = P(\frac{X - E(X)}{\sqrt{Var(X)}} \leq x)$, $X \sim Bin(n, p)$, 则有

$$|F_n(x) - \Phi(x)| \leq \frac{C}{n},$$

for $n \geq 2$.

Proof.

References

1. On the dependence of the Berry ‐ Esseen bound on dimension.
2. A multivariate Berry ‐ Esseen theorem with explicit constants.