Power Mechanism

ponkshekaustubh11

May 2023

1 Problem Statement

Given a dataset with n samples $\mathbf{X}_{k\times n} \in \mathbb{R}^k$, a positive integer $p \in \mathbb{Z}^+$, and an operator $\mathbf{H}_{k\times k} : \mathbf{X_{k\times 1}} \mapsto \mathbf{Z_{k\times 1}}$ such that $\mathbf{Z} = \mathbf{H(X_i)}^p\mathbf{X_i}$; what are the required conditions that need to be satisfied by $\mathbf{H}(X)$ and $\mathbf{p}$ to formally guarantee that $\mathbf{Z}$ is ϵ-Lipschitz private with respect to the dataset $\mathbf{X}$?

To avoid confusion, we restate that $\mathbf{H(X)}$ denotes a matrix whose entries depend on $\mathbf{X}$. In the rest of the paper we use $\mathbf{H(X)}$ and $\mathbf{H}$ interchangeably to mean the same thing without any loss of generality.

[Decorrelating privacy theorem] For $\mathbf{X} \in \mathbb{R}^k$ distributed as $\sim f_X(x)$, applying $\mathbf{Z}_p = \mathbf{H(X)}^p\mathbf{X}$ guarantees ϵ-Lipschitz privacy through $\mathbf{Z}$ when the integer power p satisfies

$$\frac{\epsilon - \left\| \frac{f'(\mathbf{X})}{f(\mathbf{X})} \right\|}{\left\| \mathbf{H(X)}^{-1} \frac{\partial \mathbf{H(X)}}{\partial \mathbf{X}} \right\|} \geq p \geq \frac{\left\| \frac{f'(\mathbf{X})}{f(\mathbf{X})} - \epsilon \right\|}{\left\| \mathbf{H(X)}^{-1} \frac{\partial \mathbf{H(X)}}{\partial \mathbf{X}} \right\| (2 \left\| \frac{f'(\mathbf{X})}{f(\mathbf{X})} \right\| - 1)}$$

Proof. The equation $\mathbf{Z}_{p_i} = [\mathbf{H(X_i)}]^p\mathbf{X}$ can be unrolled as

$$\mathbf{Z_{p_i}} = g_p \circ g_{p-1} \circ \cdots \circ g_1(\mathbf{X_i}) \tag{1}$$

where $g_p \circ g_{p-1}(\cdot) = \mathbf{H(X)}.g_{p-1}(\cdot)$.

If g is a one-to-one function on the support of $\mathbf{X}$ whose pdf is given by $f_X(x)$ where $x \in \mathbb{R}^k$, then the pdf of $\mathbf{Z} = \mathbf{g(X)}$ is

$$h(\mathbf{Z}) = f_{\mathbf{X}}(g^{-1}(\mathbf{Z}))|\det(\mathbf{J}(g^{-1}(\mathbf{Z})))|$$

for $\mathbf{Z}$ in the range of g, where $\mathbf{J(X)}$ is the Jacobian matrix of g that is evaluated at $\mathbf{X}$. This is classically known as the multidimensional change of variable theorem in the context of probability density functions. But since we have $g_p \circ g_{p-1} \circ \cdots \circ g_1(\mathbf{X})$ instead of a single $g(\cdot)$, this can be written as

$$h_p(\mathbf{Z}_p) = h_{p-1}(g_p^{-1}(\mathbf{Z}_p)) \left| \det \frac{dg_p^{-1}}{d\mathbf{Z}_p} \right|$$

We can rearrange the Jacobian of our iteration as follows

$$\frac{\partial \mathbf{H}^p\mathbf{X}}{\partial \mathbf{Z}_{p-1}} = \frac{\partial \mathbf{H}^p X}{\partial \mathbf{H}^{p-1} X} = \frac{\partial \mathbf{Z}^p}{\partial \mathbf{Z}^{p-1}} = \frac{\partial \mathbf{H}\mathbf{H}^{p-1}\mathbf{X}}{\partial \mathbf{H}^{p-1}\mathbf{X}} = \mathbf{J}(\mathbf{Z}_{\mathbf{(p-1)_i}})$$

Let us find this jacobian matrix J. For that consider the equation

$$Z_p = H(Z_{p-1})Z_{p-1}$$

$$\therefore Z_{p_i} = \sum_{j=1}^{k} H(Z_{p-1})_{ij} Z_{p-1_j}$$

Since the Jacobian matrix J is

$$J_{ij} = \frac{\partial Z_{p_i}}{\partial Z_{p-1_j}}$$

$$J_{ij} = \frac{\partial \sum_{l=1}^{k} H(Z_{p-1})_{il} Z_{p-1_l}}{\partial Z_{p-1_j}}$$

$$J_{ij} = \sum_{l=1}^{k} \frac{\partial H(Z_{p-1})_{il}}{\partial Z_{p-1_j}} Z_{p-1_l} + H(Z_{p-1})_{ij}$$

But we have the following: $\left|\det\left(\frac{dg_p}{d\mathbf{Z}_{p-1}}\right)^{-1}\right| = \left|\det\frac{dg_p}{d\mathbf{z}_{p-1}}\right|^{-1}$. Therefore upon applying log to the result of the change of variable theorem in our case, we get

$$= \log h_{p-2}(\mathbf{Z}_{p-2}) - \log\left|\det\frac{dg_{p-1}}{d\mathbf{Z}_{p-2}}\right| - \log\left|\det\frac{dg_p}{d\mathbf{Z}_{p-1}}\right|$$
$$= \dots$$
$$= \log h_0(\mathbf{Z}_0) - \sum_{i=1}^{p}\log\left|\det\frac{dg_i}{d\mathbf{Z}_{i-1}}\right|$$

Therefore we have that the logarithm of the ratio of the probability densities before and after P iterations as

$$\log\left(\frac{h(\mathbf{Z})}{f(\mathbf{X})}\right) = -\sum_{p=1}^{p}\log|\det\mathbf{J}(\mathbf{Z}_{(\mathbf{p-1})_\mathbf{i}}| = -\log(\prod_{i=1}^{p}|\det\mathbf{J}(\mathbf{Z}_{(\mathbf{p-1})_\mathbf{i}}|)$$

Now applying the derivative to the log probability and taking its norm and setting it to be less than ϵ we get the following required condition in order to satisfy Lipschitz privacy

$$\left\|\frac{\partial}{\partial\mathbf{X}}\log h(\mathbf{Z})\right\| = \left\|\frac{f'(\mathbf{X})}{f(\mathbf{X})} - \sum_{p=1}^{p}\frac{\frac{\partial|\det\mathbf{J}(\mathbf{Z}_{(\mathbf{p-1})_\mathbf{i}}|}{\partial X_i}}{|\det\mathbf{J}(\mathbf{Z}_{(\mathbf{p-1})_\mathbf{i}}|}\right\|$$

Now to differentiate the determinant of a matrix, we use Jacobi's formula

$$\left\|\frac{\partial}{\partial\mathbf{X}}\log h(\mathbf{Z})\right\| = \left\|\frac{f'(\mathbf{X})}{f(\mathbf{X})} - \sum_{p=1}^{p}\log(|\det(J))\right\|$$

$$\left\|\frac{\partial}{\partial\mathbf{X}}\log h(\mathbf{Z})\right\| = \left\|\frac{f'(\mathbf{X})}{f(\mathbf{X})} - \sum_{p=1}^{p}\frac{\det(J(Z_{(p-1)_i})tr(J^{-1}\frac{\partial J(Z_{(p-1)_i})}{\partial X_i})}{|\det\mathbf{J}(\mathbf{Z}_{(\mathbf{p-1})_\mathbf{i}}|}\right\|$$

Finally, we need to evaluate the term $J' = \frac{\partial J(Z_{(p-1)_i})}{\partial X_i}$

$$J'_{lm} = \frac{\partial J(Z_{(p-1)_i})_{lm}}{\partial X_i} = \frac{\partial(\sum_{n=1}^{k} \frac{\partial H(Z_{p-1})_{ln}}{\partial Z_{p-1_m}} Z_{p-1_n} + H(Z_{p-1})_{lm})}{\partial X_i}$$

$$J'_{lm} = \sum_{n=1}^{k} (\frac{\partial^2 H(Z_{p-1})_{ln}}{\partial X_i \partial Z_{p-1_m}} Z_{p-1_n} + \frac{\partial H(Z_{p-1})_{ln}}{\partial Z_{p-1_m}} \frac{\partial Z_{p-1_n}}{\partial X_i})$$

Therefore for obtaining $\epsilon-$Lipschitz privacy, we need to have

$$\left\| \frac{\partial}{\partial \mathbf{X}} \log h(\mathbf{Z}) \right\| = \left\| \frac{f'(\mathbf{X})}{f(\mathbf{X})} - \sum_{p=1}^{p} \log(|\det(J))\right\| \leq \epsilon$$

$$\left\| \frac{\partial}{\partial \mathbf{X}} \log h(\mathbf{Z}) \right\| = \left\| \frac{f'(\mathbf{X})}{f(\mathbf{X})} - p\mathbf{J}^{-1} \frac{\partial \mathbf{J}}{\partial \mathbf{X}} \right\| \leq \left\| \frac{f'(\mathbf{X})}{f(\mathbf{X})} \right\| + \left\| p\mathbf{H}^{-1} \frac{\partial \mathbf{J}}{\partial \mathbf{X}} \right\| \leq \epsilon$$

□

2 Estimating sample probability

We use Kernel Density Estimation for estimating the probability density of each sample.

$$\hat{f}(x) = \frac{1}{nh^d} \sum_{i=1}^{n} K(\frac{x - X_i}{h})$$

The Gaussian kernel is given by

$$K(u) = \frac{e^{-||u||^2}}{(2\pi)^{d/2}}$$

However, we need to find confidence intervals for these probability density estimates. The range in which the true probability density lies with $1-\alpha$ probability is given by

$$CI_{1-\alpha} = [\hat{f}(x) - z_{1-\alpha/2}\sqrt{\frac{\mu_K \hat{f}(x)}{nh^d}}, \hat{f}(x) + z_{1-\alpha/2}\sqrt{\frac{\mu_K \hat{f}(x)}{nh^d}}]$$

The term μ_K is given by

$$\mu_K = \int K^2(x) dx$$

For Gaussian kernel, this evaluates to

$$\mu_K = 1/(2^d \pi^{d/2})$$

The confidence bound for the gradient of is given by

$$\frac{\partial f(x)}{\partial x_i} - \frac{\partial \hat{f}(x)}{\partial x_i} = O(h^2) + O_P\left(\sqrt{\frac{1}{nh^{d+2}}}\right)$$

$$f(x) = \hat{f}(x) + \sqrt{K\hat{f}(x)} \mathcal{N}(0,1)$$

$$\frac{\partial f(x)}{\partial x_i} = \frac{\partial \hat{f}(x)}{\partial x_i} + \sqrt{\frac{K}{4\hat{f}(x)}} \hat{f}(x) \mathcal{N}(0,1)$$

3 Bringing it together

The condition for ϵ Lipschitz Privacy is given by

$$\left\| \frac{\partial}{\partial \mathbf{X}} \log h(\mathbf{Z}) \right\| \leq \epsilon$$

For obtaining Lipschitz privacy on estimated probability with $1-\alpha$ confidence,

$$\left\| \frac{\partial}{\partial \mathbf{X}} \log h(\mathbf{Z}) \right\| = \left\| \frac{f'(\mathbf{X})}{f(\mathbf{X})} - \frac{\partial}{\partial \mathbf{X}} \sum_{p=1}^{p} \log(|\det(J))\right\| = \left\| \frac{\hat{f}'(\mathbf{X})}{f(\mathbf{X})} + \frac{f'(\mathbf{X}) - \hat{f}'(\mathbf{X})}{f(\mathbf{X})} - \frac{\partial}{\partial \mathbf{X}} \sum_{p=1}^{p} \log(|\det(J))\right\| \leq \epsilon$$

$$\left\| \frac{\partial}{\partial \mathbf{X}} \log h(\mathbf{Z}) \right\| \leq \left\| \frac{\hat{f}'(\mathbf{X})}{f(\mathbf{X})} - \frac{\partial}{\partial \mathbf{X}} \sum_{p=1}^{p} \log(|\det(J))\right\| + \left\| \frac{f'(\mathbf{X}) - \hat{f}'(\mathbf{X})}{f(\mathbf{X})} \right\| \leq \epsilon$$

Now let's use the confidence interval founds on $f(X)$ to estimate ϵ

$$\text{Let } \epsilon' = \max\left(\left\| \frac{\hat{f}'(\mathbf{X})}{\hat{f}(x) - z_{1-\alpha/2}\sqrt{\frac{\mu_K \hat{f}(x)}{nh^d}}} - \sum_{p=1}^{p} \log(|\det(J))\right\|, \left\| \frac{\hat{f}'(\mathbf{X})}{\hat{f}(x) + z_{1-\alpha/2}\sqrt{\frac{\mu_K \hat{f}(x)}{nh^d}}} - \sum_{p=1}^{p} \log(|\det(J))\right\| \right)$$

$$\therefore \epsilon = \epsilon' + \left\| \frac{f'(\mathbf{X}) - \hat{f}'(\mathbf{X})}{f(\mathbf{X})} \right\|$$