Master's thesis

Master's Programme in Data Science

Differentially Private Markov Chain Monte Carlo

Ossi Räisä

September 2, 2020

Supervisor(s): Professor Antti Honkela

Examiner(s): Professor Antti Honkela
Dr. Antti Koskela

University of Helsinki
Faculty of Science

P. O. Box 68 (Pietari Kalmin katu 5)
00014 University of Helsinki

HELSINGIN YLIOPISTO — HELSINGFORS UNIVERSITET — UNIVERSITY OF HELSINKI

Tiedekunta — Fakultet — Faculty	Koulutusohjelma — Utbildningsprogram — Degree programme
Faculty of Science	Master's Programme in Data Science

Tekijä — Författare — Author

Ossi Räisä

Työn nimi — Arbetets titel — Title

Differentially Private Markov Chain Monte Carlo

Työn laji — Arbetets art — Level	Aika — Datum — Month and year	Sivumäärä — Sidantal — Number of pages
Master's thesis	September 2, 2020	13

Tiivistelmä — Referat — Abstract

ACM Computing Classification System (CCS):

Avainsanat — Nyckelord — Keywords

Differential Privacy, Markov Chain Monte Carlo

Säilytyspaikka — Förvaringsställe — Where deposited

Muita tietoja — Övriga uppgifter — Additional information

Contents

1 Introduction **1**

2 Background **3**

2.1 Differential Privacy 3

2.2 Bayesian Inference and MCMC 5

3 Differentially Private MCMC **7**

4 Variations of the Penalty Algorithm **9**

5 Conclusions **11**

Bibliography **13**

1. Introduction

2. Background

2.1 Differential Privacy

Differential privacy [2] is a property of an algorithm that quantifies the amount of information about private data an adversary can gain from the publication of the algorithm's output. The most commonly used definition uses two real numbers, ϵ and δ the quantify the information gain, or, from the perspective of the data subjects, the privacy loss of the algorithm.

The most common definition is called (ϵ, δ)-DP, approximate DP or ADP [2]. The case where $\delta = 0$ is commonly called ϵ-DP or pure DP.

Definition 1. *An algorithm $\mathcal{M}\colon \mathcal{X} \to \mathcal{U}$ is (ϵ, δ)-ADP if for all neighbouring inputs $x \in \mathcal{X}$ and $x' \in \mathcal{X}$ and all measurable sets $S \subset \mathcal{U}$*

$$P(\mathcal{M}(x) \in S) \leq e^{\epsilon} P(\mathcal{M}(x') \in S) + \delta$$

The neighbourhood relation in the definition is domain specific. With tabular data the most common definitions are the add/remove neighbourhood and substitute neighbourhood.

Definition 2. *Two tabular datasets are said to be add/remove neighbours if they are equal after adding or removing at most one row to or from one of them. The datasets are said to be in substitute neighbours if they are equal after changing at most one row in one of them.*

The neighbourhood relation is denoted by $\sim$. The definitions and theorems of this section are valid for all neighbourhood relations.

There many other definitions of differential privacy that are mostly used to compute (ϵ, δ)-bounds for ADP. This thesis uses two of them: Rényi-DP (RDP) [3] and zero-concentrated differential privacy (zCDP) [1]. Both are based on Rényi divergence [3], which is a particular way of measuring the difference between random variables.

Definition 3. *For random variables with density or probability mass functions P and*

Q the Rényi divergence of order $1 < \alpha < \infty$ is

$$D_\alpha(P \,||\, Q) = \frac{1}{\alpha - 1} \ln E_{x \sim Q} \left(\frac{P(x)^\alpha}{Q(y)^\alpha} \right)$$

Orders $\alpha = 1$ and $\alpha = \infty$ are defined by continuity:

$$D_1(P \,||\, Q) = \lim_{\alpha \to 1-} D_\alpha(P \,||\, Q)$$

$$D_\infty(P \,||\, Q) = \lim_{\alpha \to \infty} D_\alpha(P \,||\, Q)$$

Both Rényi-DP and zCDP can be expressed as bounds on the Rényi divergence between the outputs of an algorithm with neighbouring inputs:

Definition 4. *An algorithm $\mathcal{M}$ is (α, ϵ)-Rényi DP if for all $x \sim x'$*

$$D_\alpha(\mathcal{M}(x) \,||\, \mathcal{M}(x')) \leq \epsilon$$

$\mathcal{M}$ is ρ-zCDP if for all $\alpha > 1$ and all $x \sim x'$

$$D_\alpha(\mathcal{M}(x) \,||\, \mathcal{M}(x')) \leq \rho\alpha$$

A very useful property of all of these definitions is composition [2]: if algorithms $\mathcal{M}$ and $\mathcal{M}'$ are DP, the algorithm first computing $\mathcal{M}$ and then $\mathcal{M}'$, outputting both results, is also DP, although with worse bounds. More precisely

Definition 5. *Let $\mathcal{M}\colon \mathcal{X} \to \mathcal{U}$ and $\mathcal{M}'\colon \mathcal{X} \times \mathcal{U} \to \mathcal{U}'$ be algorithms. Their composition is the algorithm outputting $(\mathcal{M}(x), \mathcal{M}'(x, \mathcal{M}(x)))$ for input x.*

Theorem 1. *Let $\mathcal{M}\colon \mathcal{X} \to \mathcal{U}$ and $\mathcal{M}\colon \mathcal{X} \times \mathcal{U} \to \mathcal{U}'$ be algorithms. Then*

1. *If $\mathcal{M}$ is (ϵ, δ)-ADP and $\mathcal{M}'$ is (ϵ', δ')-ADP, then their composition is $(\epsilon+\epsilon', \delta+\delta')$-ADP [2]*

2. *If $\mathcal{M}$ is (α, ϵ)-RDP and $\mathcal{M}'$ is (α, ϵ')-RDP, then their composition is $(\alpha, \epsilon + \epsilon')$-RDP [3]*

3. *If $\mathcal{M}$ is ρ-zCDP and $\mathcal{M}'$ is ρ'-zCDP, then their composition is $(\rho + \rho')$-zCDP [1]*

All of the composition results can be extended to any number of compositions by induction. Note that any step of the composition can depend on the results of the previous steps, not only on the private data.

As any algorithm that does not use private data in any way is $(0, 0)$-ADP, 0-zCDP and $(\alpha, 0)$-RDP with all α, theorem 1 has the following corollary, called post-processing immunity:

Theorem 2. *Let $\mathcal{M}\colon \mathcal{X} \to \mathcal{U}$ be an ADP, RDP or zCDP algorithm with some privacy parameters. Let $f\colon \mathcal{U} \to \mathcal{U}'$ be any algorithm not using the private data. Then the composition of $\mathcal{M}$ and f is ADP, RDP or zCDP with the same privacy parameters.*

There are many different DP algorithms that are commonly used, which are also called mechanisms [2]. This thesis only requires one of the most commonly used ones: the Gaussian mechanism [2].

Definition 6. *The Gaussian mechanism with parameter σ^2 is an algorithm that, with input x, outputs a sample from $\mathcal{N}(x, \sigma^2)$, where $\mathcal{N}$ denotes the normal distribution.*

The RDP and zCDP bounds for the Gaussian mechanism are quite simple. The ADP bound is more complicated:

Theorem 3. *If for all inputs x and x', $||x - x'||_2 \leq \Delta$, the Gaussian mechanism is*

1. $(\alpha, \frac{\alpha\Delta^2}{2\sigma^2})$*-RDP [3]*
2. $\frac{\Delta^2}{2\sigma^2}$*-zCDP [1]*
3. *n compositions of the Gaussian mechanism are $(\epsilon, \delta(\epsilon))$-ADP [4] with*

$$\delta(\epsilon) = \frac{1}{2}\left(\mathrm{erfc}\left(\frac{\sigma(\epsilon - n\mu)}{\sqrt{2n}\Delta}\right) - e^{\epsilon}\,\mathrm{erfc}\left(\frac{\sigma(\epsilon + n\mu)}{\sqrt{2n}\Delta}\right)\right)$$

where $\mu = \frac{\Delta^2}{2\sigma^2}$ and erfc *is the complementary error function.*

The most common use case for the Gaussian mechanism is computing a function $f\colon \mathcal{X} \to \mathbb{R}$ of private data and feeding the result into the Gaussian mechanism to privately release the function value. The condition that the inputs of the Gaussian mechanism cannot vary too much leads into the concept of sensitivity of a function

Definition 7. *The l_p-sensitivity Δ_p, with neighbourhood relation $\sim$, of a function $f\colon \mathcal{X} \to \mathbb{R}^n$ is*

$$\Delta_p f = \sup_{x \sim x'} ||f(x) - f(x')||_p$$

Theorem 3 implies that the value of any function with finite l_2-sensitivity can be privately released using the Gaussian mechanism with appropriate noise variance σ^2. Of course, the usefulness of the released value depends on the magnitude of σ^2 compared to the actual value.

2.2 Bayesian Inference and MCMC

3. Differentially Private MCMC

4. Variations of the Penalty Algorithm

5. Conclusions

Bibliography

[1] M. Bun and T. Steinke. Concentrated differential privacy: Simplifications, extensions, and lower bounds. In *Theory of Cryptography - 14th International Conference, TCC 2016-B, Beijing, China, October 31 - November 3, 2016, Proceedings, Part I*, pages 635–658, 2016.

[2] C. Dwork and A. Roth. The algorithmic foundations of differential privacy. *Foundations and Trends in Theoretical Computer Science*, 9(3-4):211–407, 2014.

[3] I. Mironov. Rényi differential privacy. In *30th IEEE Computer Security Foundations Symposium, CSF 2017, Santa Barbara, CA, USA, August 21-25, 2017*, pages 263–275, 2017.

[4] D. M. Sommer, S. Meiser, and E. Mohammadi. Privacy loss classes: The central limit theorem in differential privacy. *PoPETs*, 2019(2):245–269, 2019.