Differentially Private Markov Chain Monte Carlo

Mikko Heikkilä*[1], Joonas Jälkö*[2], Onur Dikmen[3] and Antti Honkela[1]

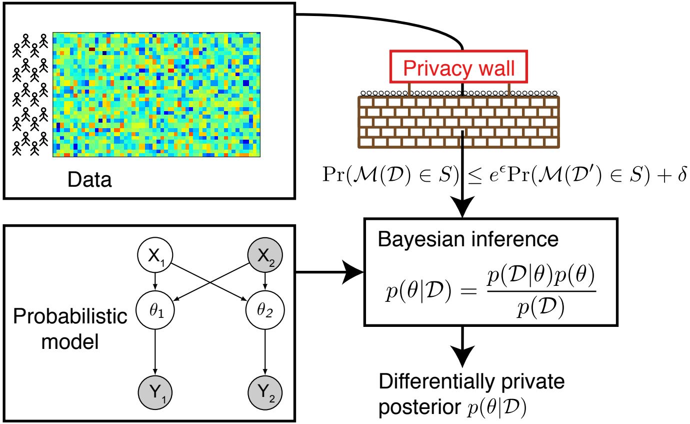
* Equal contribution
[1] University of Helsinki
[2] Aalto University
[3] Halmstad University

NeurIPS, 12 December 2019

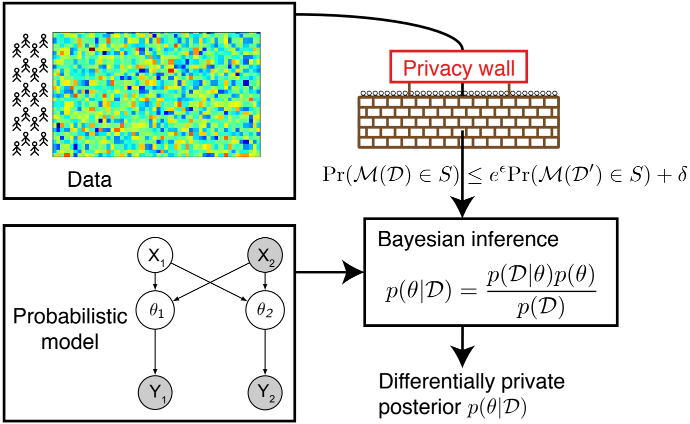
FCAI Finnish Center for Artificial Intelligence

UNIVERSITY OF HELSINKI Aalto University HALMSTAD UNIVERSITY

Motivation


Data
Privacy wall
$\Pr(\mathcal{M}(\mathcal{D}) \in S) \leq e^{\epsilon}\Pr(\mathcal{M}(\mathcal{D}') \in S) + \delta$
Probabilistic model
X_1
X_2
θ_1
θ_2
Y_1
Y_2
Bayesian inference
$p(\theta|\mathcal{D}) = \frac{p(\mathcal{D}|\theta)p(\theta)}{p(\mathcal{D})}$
Differentially private posterior $p(\theta|\mathcal{D})$

Motivation



We propose a method for sampling from posterior distribution under DP guarantees.

DP mechanisms for Bayesian inference

Three general purpose approaches for DP Bayesian inference:

1. Drawing single samples from the posterior with the exponential mechanism (Dimitrakakis *et al.*, ALT 2014; Wang *et al.*, ICML 2015; Geumlek *et al.*, NIPS 2017)

   Privacy is conditional to sampling from the true posterior.

DP mechanisms for Bayesian inference

Three general purpose approaches for DP Bayesian inference:

1. Drawing single samples from the posterior with the exponential mechanism (Dimitrakakis *et al.*, ALT 2014; Wang *et al.*, ICML 2015; Geumlek *et al.*, NIPS 2017)

   Privacy is conditional to sampling from the true posterior.

2. Perturbation of gradients in SG-MCMC (Wang *et al.*, ICML 2015, Li *et al.*, AISTATS 2019) or variational inference (Jälkö *et al.*, UAI 2017) with Gaussian mechanism, similar to DP stochastic gradient descent

   No guarantees where the algorithm converges, requires differentiability

DP mechanisms for Bayesian inference

Three general purpose approaches for DP Bayesian inference:

1. Drawing single samples from the posterior with the exponential mechanism (Dimitrakakis *et al.*, ALT 2014; Wang *et al.*, ICML 2015; Geumlek *et al.*, NIPS 2017)

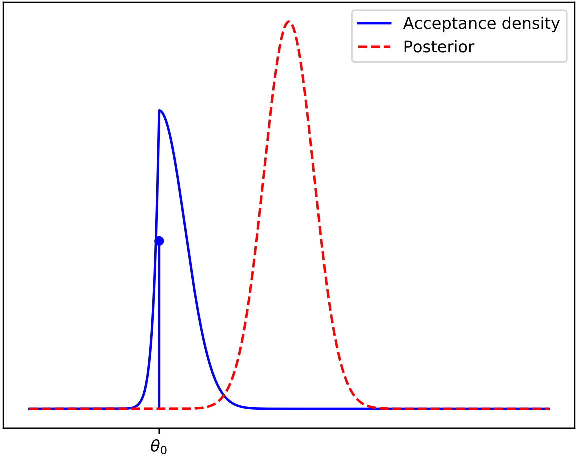
   Privacy is conditional to sampling from the true posterior.

2. Perturbation of gradients in SG-MCMC (Wang *et al.*, ICML 2015, Li *et al.*, AISTATS 2019) or variational inference (Jälkö *et al.*, UAI 2017) with Gaussian mechanism, similar to DP stochastic gradient descent

   No guarantees where the algorithm converges, requires differentiability

3. Computing the privacy cost of Metropolis–Hastings acceptances for the entire MCMC chain (Heikkilä *et al.*, NeurIPS 2019; Yıldırım & Ermiș, Stat Comput 2019)

Intuition



We employ the stochasticity of this decision to assure privacy

Outline of the method

Acceptance test (Barker et al. 1965)

Accept θ' from proposal q if $\Delta(\theta'; \mathcal{D}) + V_{logistic} > 0$

Outline of the method

Acceptance test (Barker et al. 1965)

Accept θ' from proposal q if $\Delta(\theta'; \mathcal{D}) + V_{logistic} > 0$

Subsampled MCMC (Seita et al. 2017)

Instead of using full data, evaluate above using $S \subset \mathcal{D}$

Decompose the logistic noise : $V_{logistic} = V_{normal} + V_{correction}$

$\Rightarrow$ Accept θ' from proposal q if $\Delta(\theta'; S) + \tilde{V}_{normal}(\sigma_{\Delta}^2) + V_{correction} > 0$

Outline of the method

Acceptance test (Barker et al. 1965)

Accept θ' from proposal q if $\Delta(\theta'; \mathcal{D}) + V_{logistic} > 0$

Subsampled MCMC (Seita et al. 2017)

Instead of using full data, evaluate above using $S \subset \mathcal{D}$

Decompose the logistic noise : $V_{logistic} = V_{normal} + V_{correction}$

$\Rightarrow$ Accept θ' from proposal q if $\Delta(\theta'; S) + \tilde{V}_{normal}(\sigma^2_\Delta) + V_{correction} > 0$

Analyse the privacy implications (**This work**)

We use Rényi DP to compute the privacy guarantees of the acceptance condition

Subsampling allows us to benefit from privacy amplification (Wang *et al.*, AISTATS 2019)

Conclusions

- We have formulated a DP MCMC method for which privacy guarantees do not rely on the convergence of the chain.

Come see us at our poster **#158** in **East Exhibition Hall (B + C)**

Mikko

Joonas

Onur

Antti