Venues (/) / UAI 2017 (/group?id=auai.org/UAI/2017)

An Efficient Minibatch Acceptance Test for Metropolis-Hastings

(/pdf?id=SkBOlc53x)

Blinded names

30 Mar 2017 UAI 2017 readers: UAI 2017, UAI 2017 Program Co-Chairs, UAI 2017 Senior Program Committee, UAI 2017 Program Committee, UAI 2017 Paper54 Authors Original (/forum?id=BJNdeqq2x)

Student paper: Yes

Abstract: We present a novel Metropolis-Hastings method for large datasets that uses small expected-size minibatches of data. Previous work on reducing the cost of Metropolis- Hastings tests yield variable data consumed per sample, with only constant factor reductions versus using the full dataset for each sample. Here we present a method that can be tuned to provide arbitrarily small batch sizes, by adjusting either proposal step size or temperature. Our test uses the noise-tolerant Barker acceptance test with a novel additive correction variable. The resulting test has similar cost to a normal SGD update. Our experiments demonstrate several order-of-magnitude speedups over previous work.

TL;DR: We present a Metropolis-Hastings method for minibatch MCMC on large datasets and show significant speed-ups over prior work.

Paperhash: names|an_efficient_minibatch_acceptance_test_for_metropolishastings

Authorids: auai.org/UAI/2017/Paper54/Authors

Keywords: metropolis-hastings, scalability, minibatch methods, MCMC

Subject areas: Algorithms: MCMC methods, Learning: Scalability

4 Replies

Add **Open Comment**

An Efficient Minibatch Acceptance Test for Metropolis-Hastings

UAI 2017 Paper54 AnonReviewer1

16 May 2017 UAI 2017 Paper54 Submit Review readers: UAI 2017 Program Co-Chairs, UAI 2017 Senior Program Committee, UAI 2017 Program Committee, UAI 2017 Paper54 Authors

Rating: 6: Marginally above acceptance threshold

Review: The paper proposes a new method for accept-rejection with minibatches in the Metropolis-Hastings test. It improves upon the sample efficiency of similar MH tests with minibatches in that the test is not approximate, and the test does not feature worst-case bounds that scan through the full data set in a MH test.

The authors are familiar with related work, and tie very nicely to many approaches as, e.g., surveyed in Bardenet et al (2016). The problem is very significant and is key to many MCMC methods beyond vanilla MH as they highlight.

The logistic function seems arbitrary. Have you explored alternative functions that might give more tractable correction distributions or more sample efficiency? Within this section, is the test still exact if approximations are made to find an (approximate) correction distribution?

One issue I have with these methods which use CLT arguments is that in such limiting cases, one has to wonder whether the posterior is any interesting beyond concentration at a point, and if it is interesting, whether these methods are practical. This is also reflected in the error analysis, which focuses on the simple setting of Gaussian densities. Can you describe the key take away from the analysis section?

As another example, how does this handle models with local latent variables, e.g., random effects models or uncollapsed mixture models? Do the CLT arguments still apply, and if not, how do you propose to handle them?

The experiments focus on Gaussian distributed models or logistic regression with uniformly a priori distributed parameters. Do these methods have any hope of successfully training, e.g., a Bayesian neural network? In addition, why not also compare to SGMCMC methods that don't use a MH ratio? It seems a major criticism is that the proposal step sizes are limited because they ultimately require an acceptance of 1 at all times. It would be useful to see if this is also true in practice.

Confidence: 4: The reviewer is confident but not absolutely certain that the evaluation is correct

Add **Open Comment**

This paper proposed a novel acceptance test approach for sampling methods and justified it well.

UAI 2017 Paper54 AnonReviewer3

16 May 2017 UAI 2017 Paper54 Submit Review readers: UAI 2017 Program Co-Chairs, UAI 2017 Senior Program Committee, UAI 2017 Program Committee, UAI 2017 Paper54 Authors

Rating: 6: Marginally above acceptance threshold

Review: This paper studies posterior sampling problems in Bayesian inference. To overcome the difficulty in sampling when the data size is large, the paper proposed a mini-batch method to decrease the computer complexity while maintain a high accuracy. It is proved that the distribution of the sample converges to the desired distribution asymptotically. Numerical simulation on mixture of Gaussian distribution and logistic regression verifies the efficacy of the proposed approach.

Overall, I think the paper is well presented and the related works are investigated comprehensively. I recommend the authors provide a comparison on the sample complexity between the proposed algorithm and the state-of-the-art sampling algorithms to highlight the significance of the current work.

A question about the choice of function g(): Lemma 2 claims that as long as g() satisfies g(s) = exp(s) g(-s), then g() is suitable to generate the acceptance function. I think a simple choice of g(s) is g(s) = exp(s/2), which seems to be simpler than Barker logistic function. So why don't we use the simpler one? Is there any specific reason to use the more complex Barker logistic function?

Confidence: 3: The reviewer is fairly confident that the evaluation is correct

Add **Open Comment**

Important topic with clearly presented contribution

UAI 2017 Paper54 AnonReviewer4

15 May 2017 UAI 2017 Paper54 Submit Review readers: UAI 2017 Program Co-Chairs, UAI 2017 Senior Program Committee, UAI 2017 Program Committee, UAI 2017 Paper54 Authors Revisions (/revisions?id=BJQ10vweW)

Rating: 10: Top 5% of accepted papers, seminal paper

Review: I enjoyed reading this paper a lot. MCMC analysis presents a large number of issues as the sample size gets large; so I might find the results of this paper particularly applicable. The approach presented here, using an alternative to the Metropolis ratio and auxiliary variables seems interesting and usable. The results, both theoretical and empirical, are impressive.

I also thought that the paper was exceedingly well written. In particular, I could follow clearly each step of the argument all the way through.

Confidence: 3: The reviewer is fairly confident that the evaluation is correct

Add **Open Comment**

This paper presents a new method for Metropolis-Hastings sampling which utilises much less data while still achieving good results. It appears to me that the derivations are technically sound (didn’t go through carefully) and the results (speed up) are promising.

UAI 2017 Paper54 AnonReviewer2

05 May 2017 UAI 2017 Paper54 Submit Review readers: UAI 2017 Program Co-Chairs, UAI 2017 Senior Program Committee, UAI 2017 Program Committee, UAI 2017 Paper54 Authors

Rating: 7: Good paper, accept

Review: 1. This paper’s contributions are buried under the technical details and thus not easy to see the main idea of the paper. Since the maths for the derivations are rather involved, I suggest adding more explanation/intuition on why a particular method is useful. For example, why the Barker logistic acceptance function in Sec 3.2 is used?

2. The concept of temperature is not explained. What is temperature and what does it do? If this is a well known thing then perhaps adding a line or two to briefly explain the key words will help.

3. It would be interesting to see the performance of the various MCMC samplers in how accurate they can obtain the posterior distribution of the parameters. As I understand, there are some approximations involved and thus the algorithms might not actually converge to the true posteriors. It would be nice to, say, evaluate the posterior means of the parameters and see how close they are to the ground truth parameters.

4. By using citation as nouns, it makes reading a sentence harder, eg in “under the assumptions of [Korattikara et al., 2014]”. It can be changed to “under the assumptions of Korattikara et al. [2014]” which reads better. This can be done by using \citep, \citet, etc (in natbib package).

5. Minor spelling mistakes in the reference, the proper nouns are not capitalised. e.g. markov chain monte carlo, gibbs sampling. This can easily be fixed by enclosing the proper nouns by curly brackets, i.e. markov chain monte carlo -> {Markov} chain {Monte Carlo}.

Confidence: 1: The reviewer's evaluation is an educated guess

Add **Open Comment**