The Journal of Open Source Software

SGMCMCJax: a lightweight JAX library for stochastic gradient Markov chain Monte Carlo algorithms

Jeremie Coullon[1] **and Christopher Nemeth**[2]

1 Cervest, London, UK **2** Lancaster University, UK

DOI: 10.21105/joss.04113

Software

- Review
- Repository
- Archive

Editor: Dan Foreman-Mackey

Reviewers:

- @canyon289
- @ColCarroll

Submitted: 17 January 2022
Published: 18 April 2022

Summary

In Bayesian inference, the *posterior distribution* is the probability distribution over the model parameters resulting from the prior distribution and the likelihood. One can compute integrals over this distribution to obtain quantities of interest, such as the posterior mean and variance, or credible uncertainty regions. However, as these integrals are often intractable for problems of interest they require numerical methods to approximate them.

Markov Chain Monte Carlo (MCMC) is currently the gold standard for approximating integrals needed in Bayesian inference. However, as these algorithms become prohibitively expensive for large datasets, stochastic gradient MCMC (SGMCMC) (Ma et al., 2015; Nemeth & Fearnhead, 2021) is a popular approach to approximate these integrals in these cases. This class of scalable algorithms uses data subsampling techniques to approximate gradient based sampling algorithms, and are regularly used to fit statistical models or Bayesian neural networks (BNNs). The SGMCMC literature develops new algorithms by finding novel gradient estimation techniques, designing more efficient diffusions, and finding more stable numerical discretisations to these diffusions. SGMCMCJax is a lightweight library that is designed to allow the user to innovate along these lines or use one of the existing gradient-based SGMCMC algorithms already included in the library. This makes SGMCMCJax very well suited for both research purposes and practical applications.

Statement of need

SGMCMCJax is a Python package written in the popular JAX library (Bradbury et al., 2018). Although there are libraries for SGMCMC algorithms in other languages and automatic differentiation frameworks (Abadi et al., 2015; Baker et al., 2019), there is no mature library for the JAX ecosystem. However, as this has become a popular framework for machine learning and scientific computing, this gap has become more noticeable. As SGMCMC algorithms are a standard tool to train Bayesian neural networks as well as statistical models with large datasets, we have written this library of samplers to fill this gap.

SGMCMCJax uses JAX to perform automatic differentiation and compilation to XLA. The use of JAX allows the SGMCMCJax library to effortlessly run on GPUs and TPUs, which is essential for large models such as BNNs. As a result, the library has an easy-to-use interface and provides very competitive speed performance. SGMCMCJax is designed in a modular framework allowing users to simply run one of its many algorithms or to create new algorithms for research purposes by using the existing algorithms as building blocks. Furthermore, SGMCMCJax can integrate easily with other codebases within the JAX ecosystem such as Flax, a neural network library for JAX. As SGMCMC algorithms are often used to train BNNs, which are usually written using frameworks such as Flax, having a library that can interact with these packages is essential to support research in the growing area of Bayesian neural network modelling.

There also exists many probabilistic programming languagues (PPLs) such as NumPyro, Stan, or PyMC, which allow users to define models and sample from them using advanced MCMC algorithms. However, these PPLs do not usually include SGMCMC algorithms in their inference toolkit which are useful for dealing with large datasets. In contrast, SGMCMCJax has a wide range of standard and state-of-the-art SGMCMC algorithms, does not include a modelling language, and is designed in a modular way that allows users to develop new sampling algorithms. As a result, this lightweight library fills the need for these scalable algorithms, and can be used to sample from models defined using these PPLs.

Software requirements and external usage

SGMCMCJax is written using JAX (Bradbury et al., 2018) as its main requirement. Although SGMCMCJax is a recent library it has already been used in a research paper (Coullon et al., 2021) as well as used in the code to accompany the book *"Machine learning: a probabilistic perspective"* (Murphy, 2023).

Acknowledgements

The design of the codebase was inspired by JAX's optimizers module as well as the Blackjax library of MCMC samplers. We give special thanks to Kevin Murphy, Remi Louf, Colin Carroll, Charles Matthews, and Sharad Vikram for code contributions and insightful discussions.

The authors were supported by the Engineering and Physical Sciences Research Council grantEP/V022636/1.

References

Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., … Zheng, X. (2015). *TensorFlow: Large-scale machine learning on heterogeneous systems*. https://www.tensorflow.org/

Baker, J., Fearnhead, P., Fox, E. B., & Nemeth, C. (2019). sgmcmc: An R package for stochastic gradient Markov chain Monte Carlo. *Journal of Statistical Software*, *91*(3), 1–27. https://doi.org/10.18637/jss.v091.i03

Bradbury, J., Frostig, R., Hawkins, P., Johnson, M. J., Leary, C., Maclaurin, D., Necula, G., Paszke, A., VanderPlas, J., Wanderman-Milne, S., & Zhang, Q. (2018). *JAX: Composable transformations of Python+NumPy programs* (Version 0.2.5) [Computer software]. http://github.com/google/jax

Coullon, J., South, L., & Nemeth, C. (2021). *Efficient and generalizable tuning strategies for stochastic gradient MCMC*. https://arxiv.org/abs/2105.13059

Ma, Y.-A., Chen, T., & Fox, E. (2015). A complete recipe for stochastic gradient MCMC. In C. Cortes, N. Lawrence, D. Lee, M. Sugiyama, & R. Garnett (Eds.), *Advances in neural information processing systems* (Vol. 28). Curran Associates, Inc. https://proceedings.neurips.cc/paper/2015/file/9a4400501febb2a95e79248486a5f6d3-Paper.pdf

Murphy, K. (2023). *Probabilistic machine learning: Advanced topics*. https://probml.github.io/pml-book/book2.html

Nemeth, C., & Fearnhead, P. (2021). Stochastic gradient Markov chain Monte Carlo. *Journal of the American Statistical Association*, *116*(533), 433–450.