Variational Bayesian Monte Carlo & Variational Bayesian Monte Carlo with Noisy Likelihoods

Gagesh Madaan - 20111406 (gagesh20@iitk.ac.in)
Shruti Sharma - 20111061 (shruti20@iitk.ac.in)
Musale Krushna Pavan - 20111268 (krushna20@iitk.ac.in)
Pinaki Chakraborty - 20211265 (pinakic20@iitk.ac.in)

Feb 2021

1 Motivation

The motivation is to explore the feasibility of combining variational inference technique of optimization with the sampling based ideas to come up with a novel approach which would be more scalable in higher dimensions as claimed in the original paper. We are excited to observe and understand the underlying intricacies and evaluate it with the models presented in the papers as well as some unlisted models/datasets.

2 Problem Statement

Without any doubt, Bayesian Inference is a neat and systematic way that helps us to account for model prediction and parameter uncertainty by computing quantities like marginal likelihood, posterior distribution over parameters and posterior predictive distribution over parameters. But Bayesian Inference, in general, is analytically intractable and in order to overcome this hurdle, various approximate inference techniques like Markov Chain Monte Carlo and Variational Inference are used.
Although widely used, a major drawback of such techniques are the prerequisites i.e. access to the gradients or large number of model evaluations, that cannot be met by more realistic and practically used black-box probabilistic models with computationally expensive likelihoods.

A more recent alternative that has shown promising result in this scenario is building a probabilistic-model based approximation of the function of interest, for example via Gaussian processes (GP). This approach has been extensively utilised in Bayesian Optimization (BO) and in Bayesian quadrature for the computation of intractable integrals. GP-based Bayesian Quadrature has also been recently used in approximating posteriors and marginal likelihoods but no-one has tried to simultaneously approximate both. Moreover, analysis of realistic requirements like medium dimensionality (up to 10), mild multi-modality, heavy tails, and parameters that exhibit strong correlations remain unexplored.

We aim to study Variational Bayesian Monte Carlo[14,15] (VBMC) that combines variational inference and active-sampling Bayesian quadrature via GP, and claims to give solutions to the above discussed issues showing indications of being used as a tool for approximating intractable posteriors. Our final goal would be to bring into light some improvements and suggestions for the given approach.

3 Prior Work

GP-based Bayesian quadrature sampling approach has been used for the estimation of marginal likelihood [8, 9, 10, 11] and GP surrogates to approximate posteriors [12, 13]. Apart from this, GP based approach has been successfully applied in Bayesian Optimisation [2, 3, 4, 5, 6] and computation of intractable integrals [7, 8]. We are planning to read handful of these to understand the intuitiveness of the approximations that we build using this approach.

The previous work in this domain doesn't offer methods to work with both posterior and marginal likelihood at the same time. They also do not specify anything regarding noisy and expensive likelihood like medium dimensionality, heavy tails etc.

4 Timeline

1. Build background on Probabilistic Inference
2. Learn about various inference methods.
3. Learn methods that solve intractable coupling, like MCMC etc.
4. Why shift to VBMC from MCMC? Speed? Computation power? Accuracy?
5. Practical applications to employ VBMC in.
6. Optimisations for Variational Bayes Algorithm.
7. Discuss variants of the approach

Time Period	**Proposed Work**
Feb - Before Midsems	Go through and completely understand the original paper. Grasp the concepts and build a strong mathematical base to understand the derivations involved. Create a high-level report, in simple language, to help fellow-students understand the topic well before midsem exams.
March - After Midsems	Study more literature dealing with our research problem and understand the paper's implementation. We will try to add more efficient implementation and evaluate it. Then, compare those results with the ones mentioned in the original paper. Keep a record of whatever advancements are made in the concerned topic. Seek feedback of the instructor.
April - Before Final Evaluation	We will conclude our research experience by creating a detailed report of all our work, evaluations, findings, improvements. If we are able to come up with another state-of-the-art approach, we wish to continue our work and publish a research paper.

Table 1: Plan of work

5 References

[1] Rasmussen, C. Williams, C. K. I. (2006) Gaussian Processes for Machine Learning. (MIT Press).

[2] Jones, D. R., Schonlau, M., Welch, W. J. (1998) Efficient global optimization of expensive black-box functions. Journal of Global Optimization 13, 455–492.

[3] Brochu, E., Cora, V. M., De Freitas, N. (2010) A tutorial on Bayesian optimization of expensive cost functions, with application to active user modeling and hierarchical reinforcement learning. arXiv preprint arXiv:1012.2599.

[4] Snoek, J., Larochelle, H., Adams, R. P. (2012) Practical Bayesian optimization of machine learning algorithms. Advances in Neural Information Processing Systems 25, 2951–2959.

[5] Shahriari, B., Swersky, K., Wang, Z., Adams, R. P., de Freitas, N. (2016) Taking the human out of the loop: A review of Bayesian optimization. Proceedings of the IEEE 104, 148–175.

[6] Acerbi, L. Ma, W. J. (2017) Practical Bayesian optimization for model fitting with Bayesian adaptive direct search. Advances in Neural Information Processing Systems 30, 1834–1844.

[7] O'Hagan, A. (1991) Bayes–Hermite quadrature. Journal of Statistical Planning and Inference 29, 245–260.

[8] Ghahramani, Z. Rasmussen, C. E. (2002) Bayesian Monte Carlo. Advances in Neural Information Processing Systems 15, 505–512.

[9] Osborne, M., Duvenaud, D. K., Garnett, R., Rasmussen, C. E., Roberts, S. J., Ghahramani, Z. (2012) Active learning of model evidence using Bayesian quadrature. Advances in Neural Information Processing Systems 25, 46–54.

[10] Gunter, T., Osborne, M. A., Garnett, R., Hennig, P., Roberts, S. J. (2014) Sampling for inference in probabilistic models with fast Bayesian quadrature. Advances in Neural Information Processing Systems 27, 2789–2797.

[11] Briol, F.-X., Oates, C., Girolami, M., Osborne, M. A. (2015) Frank-Wolfe Bayesian quadrature: Probabilistic integration with theoretical guarantees. Advances in Neural Information Processing Systems 28, 1162–1170.

[12] Kandasamy, K., Schneider, J., Póczos, B. (2015) Bayesian active learning for posterior estimation. Twenty-Fourth International Joint Conference on Artificial Intelligence.

[13] Wang, H. Li, J. (2018) Adaptive Gaussian process approximation for Bayesian inference with expensive likelihood functions. Neural Computation pp. 1–23.

[14] Luigi Ascerbi, Variational Bayesian Monte Carlo with Noisy Likelihoods, 2020

[15] Luigi Ascerbi, Variational Bayesian Monte Carlo, 2018