Topics in Probability (STAT 547C)
Fall 2019
https://ben-br.github.io/stat-547c-fall-2019/

Vittorio Romaniello
vittorio.romaniello@stat.ubc.ca

Stein's method in machine learning

Project Outline

1 Title

The working title of my project is *Stein's method in machine learning.*

2 Background

Stein's method [Ste+72] is a technique that can quantify the error in the approximation of one distribution by another in a variety of metrics [Ros+11]. The method is mainly used to derive bounds on approximations as well as convergence results in a variety of disciplines (see e.g.[Rei11] for some areas of application). In recent years, Stein's method has been introduced in the machine learning literature. [GM15; LLJ16; CSG16; OGC17] used the method to design goodness of fit and sample quality tests for sampling methods and [LW16; Zhu+17] achieved better performance and speed ups in variational inference using a version of Stein's discrepancy [LLJ16]. I am interested in gaining an understanding of Stein's method and in conducting a more in-depth analysis of Stein's variational inference.

3 Technical aspects

The project will draw on technical aspects of the following areas: convergence, expectations, kernels and measures, variational inference.

4 Literature

The key references for this project are:

- [Ros+11], provides a review fo Stein's method and examples for different metrics and distributions.
- [LW16] the paper introduces Stein's variational inference.
- [Zhu+17] improves the method of Liu and Wang [LW16].
- [LLJ16; CSG16] are foundational works for [LW16] as they introduce Stein's discrepancy and the derivation of the optimal result minimizing the discrepancy. The results are then adopted in the variational inference context.

5 Plan

I will carry out this project with the following sequence of steps:

1. I will introduce Stein's method and derive some of the theoretical results needed as background for the understanding of Stein's variational inference.
2. I will provide a more theoretical introduction to Reproducing Kernel Hilbert Space (RKHS).
3. I will reformulate Stein's variational inference of Liu and Wang [LW16] in a more rigorous probabilistic way.
4. If time allows, I will also discuss Zhuo et al. [Zhu+17] in similar terms.

6 Why I'm interested in this topic

I am interested in variational inference and Markov Chain Monte Carlo methods and it seems like Stein's method can provide improvements to existing methods in the literature. I hope a deeper understanding of the method can help me find new areas of research.

References

[Ste+72] C. Stein et al. "A bound for the error in the normal approximation to the distribution of a sum of dependent random variables". In: *Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Volume 2: Probability Theory*. The Regents of the University of California. 1972.

[Ros+11] N. Ross et al. "Fundamentals of Stein's method". In: *Probability Surveys* 8 (2011), pp. 210–293.

[Rei11] G. Reinert. "A short introduction to Stein's method". In: *Lecture Notes* (2011).

[GM15] J. Gorham and L. Mackey. "Measuring sample quality with Stein's method". In: *Advances in Neural Information Processing Systems*. 2015, pp. 226–234.

[LLJ16] Q. Liu, J. Lee, and M. Jordan. "A kernelized Stein discrepancy for goodness-of-fit tests". In: *International conference on machine learning*. 2016, pp. 276–284.

[CSG16] K. Chwialkowski, H. Strathmann, and A. Gretton. "A kernel test of goodness of fit". In: JMLR: Workshop and Conference Proceedings. 2016.

[OGC17] C. J. Oates, M. Girolami, and N. Chopin. "Control functionals for Monte Carlo integration". In: *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 79.3 (2017), pp. 695–718.

[LW16] Q. Liu and D. Wang. "Stein variational gradient descent: A general purpose bayesian inference algorithm". In: *Advances in neural information processing systems*. 2016, pp. 2378–2386.

[Zhu+17] J. Zhuo et al. "Message passing stein variational gradient descent". In: *arXiv preprint arXiv:1711.04425* (2017).