Final Project Proposal: Scaling Gaussian Processes

Philip Pham

May 3, 2020

I'm mainly interested in kernel methods, specfically Gaussian processes due to their Bayesian interpretation. I'm primarily concerned with practical aspects of working on large datasets: namely, the standard formulation of requires inverting an $N \times N$ matrix, which is an $O(N^3)$ operation [Rasmussen and Williams, 2005]. I'd like to survey approximation methods that can scale Gaussian processes to larger datasets.

To this end, I plan to give a brief survey of Gaussian processes, some intuition, and their connection with methods discussed in class [Görtler et al., 2019]. Next, I plan to introduce some yet to be determined methods. Examples may be reduced-rank approximations, greedy methods, random Fourier features, and alternative kernels [Jacot et al., 2018]. For some of the easier to implement methods, I'd like to present some emprical results comparing how well the approximations perform.

References

Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning)*. The MIT Press, 2005. ISBN 026218253X.

Arthur Jacot, Franck Gabriel, and Clément Hongler. Neural tangent kernel: Convergence and generalization in neural networks. *CoRR*, abs/1806.07572, 2018. URL `http://arxiv.org/abs/1806.07572`.

Jochen Görtler, Rebecca Kehlbeck, and Oliver Deussen. A visual exploration of gaussian processes. *Distill*, 2019. doi: 10.23915/distill.00017. https://distill.pub/2019/visual-exploration-gaussian-processes.