Project Overview

In this project, I explore different universal boosting

Mean field variational inference is a widely used approach in Bayesian inference, where the posterior distribution is approximated with a variational distribution that factorizes across the parameters of the model. However, this assumption restricts the approximating distribution to be a unimodal distribution, and therefore cannot capture the multimodality. Moreover, MFVI tends to underestimate the posterior covariance, which can lead to false certainty about the parameter estimates.

A natural idea is to use a mixture distribution as the approximating variational distribution. Using simple components such as Gaussian's, we are able to approximate any continuous probability density arbitrarily well.

Proof

What have you done so far? Report any preliminary findings (positive or negative)

From my original proposal, I have

1. (Done) Implement Automatic Differentiation Variational Inference (ADVI) and Black Box Variational Inference (BBVI).

   I have implemented BBVI with mean-field Gaussian families and variational inference from scratch in JAX. I used the re-paramaterization trick for mean-field Gaussian families.

2. (Done) For both debugging and educational purposes, create a visualization program to show refinement of ADVI and BBVI to the posterior using simple 1-D and 2-D distributions.

   I have created some simple visualizations to visualize that my BBVI on simple 1D and 2D distributions, such as a Gamma distribution, and a two dimensional Gaussian distribution with correlation.

3. (Done) Implement Universal BVI (UBVI) from Campbell et al. (2019) and add to visualization program.

   I have completed a literature review for UBVI and explored the literature of variational inference that do not use the KL-divergence as the variational objective. I have tested that Universal Boosting Variational Inference can be used to approximate a Cauchy

4. (Partially) Implement Black Box Boosting Variational Inference

I have completed a literature review for UBVI and explored the literature of variational inference that do not use the KL-divergence as the variational objective. I have tested that Universal Boosting Variational Inference can be used to approximate a Cauchy

4. (04/06, pushed) Implement BVI from Miller et al. (2016)

5. (04/13) Implement BVI from Guo et al. (2016) and compare differences and analyze issues with joint weight and mixture optimization from step 3 using visualization program.

6. (Optional) Implement Stein Variational Gradient Descent (SVGD) and add to visualization program. Use ideas from BVI to formulate a boosting or sequential SVGD where particles

are iteratively added.

7. (Optional) Add logistic regression experiments using the different implemented VI algorithms.

What do you plan