PROJECT PROPOSAL

Online Variational Bayesian Inference (SVI)
CS698S - Piyush Rai

Kushal Kumar-14346
Pratik Mishra- 14493

Introduction:

Online variational Bayesian inference, also called the stochastic variational inference (SVI), is an essential tool in Bayesian statistics to approximate posterior distributions for complex statistical models. It uses optimization technique on a family of density functions over latent variables, that finds a member of the family with least KL divergence from the required conditional density, upto a constant (ELBO), which can be further used for Bayesian inferences. Variational Bayesian models can also be used to find *lower bounds* for marginal likelihood, which is helpful in *model selection*. The advantage of using SVI over other approximation techniques like MCMC is that SVI is suited for larger data sets and performs better on mixture models, though the best solution is not guaranteed in this case and the model may propose values beyond the expected range also.

Proposed Goals:

We intend to theoretically examine the concepts in online variational Bayesian inference, the Kullback-Leibler divergence, the Evidence Lower Bound (ELBO), Mean-Field variational inference and then re-implement the concepts proposed in some of the recent research papers, like [3]. In the latter half we may try to apply variational inference to new statistical models, as suggested in the proposed idea itself and understand how SVI is implemented for different models and use the predicted posterior for meaningful inferences. We may also shed our focus on the limitations and suggest some possible ideas for improvement, that were not implemented.

Recent Works:

Jake Stolee and Neill Patterson in [3] use Neural Network Matrix Factorization(NNMF) to apply collaborative filtering to predict movie ratings as a function of user and movie latent variables by extending the point estimate model to use variational Bayesian inference instead of MAP estimates with the help of mean-field variational approximation to the posteriors. By examining the learned posterior distribution, they evaluated with confidence, the interest of a user in a movie, and could compare the interests between users, with higher precision.

References

[1] Variational Inference, David M.Blei. Link: https://www.cs.princeton.edu/courses/archive/fall11/cos597C/lectures/variational-inference-i.pdf

[2] Variational Inference: A Review for Statisticians, David M. Blei, Alp Kucukelbir, Jon D. McAuliffe, November 3, 2016. Link: https://arxiv.org/pdf/1601.00670.pdf

[3] Matrix Factorization with Neural Networks, Jake Stolee, Neill Patterson, Link: https://www.cs.toronto.edu/$\sim jstolee/projects/matrix\_factorization\_neural$.pdf