TODO

Kenny Chiu

March 8, 2022

1   Critical analysis

1.1   Introduction

TODOTitsias and Ruiz [4] introduce Unbiased Implicit Variational Inference (UIVI) as a variational inference method that allows for a flexible variational family and that addresses the issues of the methods that it is built on. In this analysis, we summarize the work of Titsias and Ruiz [4] in the context of the literature and critically examine the strengths and limitations of UIVI. This analysis is organized as follows: Section 1.2 introduces the problem context and previous work; Section 1.3 describes how UIVI works and how it addresses the limitations of previous methods, and discusses its own limitations.

1.2   Context and previous work

Variational inference (VI) [1] is a Bayesian inference method that formulates the problem of finding the posterior distribution $p(\mathbf{z}|\mathbf{x})$ of latent variables $\mathbf{z}$ given data $\mathbf{x}$ as an optimization problem. VI posits a variational family $\mathcal{Q} = \{q_\theta\}$ of distributions indexed by variational parameters θ and aims to approximate the posterior distribution by some simpler variational distribution $q_\theta(\mathbf{z}) \in \mathcal{Q}$. In standard VI, the selected distribution q_θ is the one that minimizes the Kullback-Leibler (KL) divergence of q_θ and $p(\mathbf{z}|\mathbf{x})$ or equivalently, the one that maximizes the evidence lower bound (ELBO) denoted as

$$\mathcal{L}(\theta) = \mathbb{E}_{q_\theta(\mathbf{z})}\left[\log p(\mathbf{x}, \mathbf{z}) - \log q_\theta(\mathbf{z})\right] .$$

To maximize the ELBO, standard VI places strong restrictions on the choice of the model and the variational family in order to allow the use of a coordinate ascent algorithm. These restrictions include (1) a mean-field assumption where the latent variables are marginally independent and the variational distribution factorizes as $q_\theta(\mathbf{z}) = \prod_{i=1}^{d} q_{\theta_i}(\mathbf{z}_i)$ and (2) the model has conjugate conditionals where $p(\mathbf{z}_i)$ and $p(\mathbf{z}_i|\mathbf{x}, \mathbf{z}_{\neg i})$ are from the same distribution family.

A major development from standard VI was black box VI (BBVI) [2] which relaxed the restrictive assumptions by optimizing the ELBO through a different approach. By rewriting the ELBO gradient in terms of an expectation, gradients could be estimated via Monte Carlo approaches. Exchanging the above assumptions for the different assumption that one can sample from the variational distribution $q_\theta(\mathbf{z})$, this expanded the possibilities for the choice of the variational family. In particular, hierarchical variational families [3] of the form $q_\theta(\mathbf{z}) = \int q_\theta(\mathbf{z}|\varepsilon)q_\theta(\varepsilon)d\varepsilon$ were proposed with their main strength being the ability to model marginal dependencies between latent variables (which the mean-field family could not) through the mixing distribution $q_\theta(\varepsilon)$.
 TODOSIVI
 TODOblack box vi and hierarchical methods + limitations

1.3   Current work

TODOhow UIVI works, main advantages, other contributions of paper (empirical vs sivi), limitations

2   Project report

TODOtitle

Abstract

TODO

2.1   Introduction

References

[1] Michael I. Jordan, Zoubin Ghahramani, and et al. An introduction to variational methods for graphical models. In *Machine Learning*, pages 183–233. MIT Press, 1999.

[2] Rajesh Ranganath, Sean Gerrish, and David Blei. Black box variational inference. In *Artificial Intelligence and Statistics*, pages 814–822. PMLR, 2014.

[3] Rajesh Ranganath, Dustin Tran, and David Blei. Hierarchical variational models. In *International Conference on Machine Learning*, pages 324–333. PMLR, 2016.

[4] Michalis K Titsias and Francisco Ruiz. Unbiased implicit variational inference. In *The 22nd International Conference on Artificial Intelligence and Statistics*, pages 167–176. PMLR, 2019.