Project proposal
Improving Explorability in Variational Inference with Annealed Variational Objectives

Mikhail Kurenkov, Timur Chikichev, Aleksei Pronkin

September 2020

Github repository: https://github.com/alexey-pronkin/annealed

Introduction

Variational Inference is widely used for solving a Bayesian inference problem. It is different from Markov Chain Monte Carlo(MCMC) methods, which rely on the Markov chain to reach equilibrium; in VI one can easily draw i.i.d. samples from the variational distribution, and enjoy lower variance in inference. However, vanilla VI has two major problems: overconfidence in prediction distribution and bad local optima with the unimodal posterior distribution. Paper [1] claims that the optimization process could limit the density of posterior distribution. The authors of this work aim to solve these issues with different objective functions and some optimization tricks. We also investigate closely related posterior collapse problem, where the generative model learns to ignore a subset of the latent variable. The paper [2] gives a general introduction to this phenomenon. One of the solutions to this problem is to use annealing strategies for inference, for example, alpha or beta annealing.

The paper [1] states that due to the zero-forcing property of the KL the true posterior tends to be unimodal in usual variational inference, the drawbacks of biasing. The author introduces the hybrid method of alpha annealing and Annealed importance sampling, called Annealed Variational Objectives (AVO). The method uses a highly flexible parametric form of the posterior distribution (assuming we have a rich family of approximate posterior at the hands).

Goal

During this project we are going to implement method from [1] and repeat experiments of this paper.

Experiments:

1. Biased noise model.
2. Toy energy fitting.
3. Quantitative analysis on robustness to beta annealing.
4. Amortized inference on MNIST and CelebA datasets.

We also want to demonstrate posterior collapse on toy example and show how method from AVO [1] can mitigate this problem. Also we want to check if Variational Auto Encoder (VAE) can be improved by the AVO method on the MNIST dataset.

General plan

1. Re-implement Variational-Auto Encoder (VAE) [3] with parameters from [4].
2. Implement Hierarchical Variational Model (HVM) based on [5].
3. Implement Importance Weighted Auto-Encoder (IWAE) based on [6].

4. Implement Hierarchical Variational Inference (HVI) method based on [5].
5. Implement AVO method [1].
6. Analyze a behavior of Annealed Variational Objectives (AVO) objective on toy examples.
7. Conduct experiments with VAE and AVO on the MNIST and CelebA datasets.

Acknowledgements

The project represents the paper [1]. We use [2] as an introduction to the problem and [7] as an introduction to generalized variational inference problem.

We planned to use and rewrite some code from https://github.com/joelouismarino/iterative_inference/, https://github.com/jmtomczak/vae_householder_flow, https://github.com/AntixK/PyTorch-VAE, , https://github.com/haofuml/cyclical_annealing and https://github.com/ajayjain/lmconv. (We assume, that first two repositories were used in the original paper [1] closed source code)

We want to try to apply annealing strategies for some of SoTA AE for MNIST https://paperswithcode.com/sota/image-generation-on-mnist if we will have time.

References

[1] C. Huang, Shawn Tan, Alexandre Lacoste, and Aaron C. Courville. Improving explorability in variational inference with annealed variational objectives. *ArXiv*, abs/1809.01818, 2018.

[2] James Lucas, George Tucker, Roger Grosse, and Mohammad Norouzi. Understanding posterior collapse in generative latent variable models. 2019.

[3] Diederik P Kingma and Max Welling. Auto-encoding variational bayes, 2013.

[4] Jakub M Tomczak and Max Welling. Improving variational auto-encoders using householder flow. *arXiv preprint arXiv:1611.09630*, 2016.

[5] Rajesh Ranganath, Dustin Tran, and David M. Blei. Hierarchical variational models, 2015.

[6] Yuri Burda, Roger Grosse, and Ruslan Salakhutdinov. Importance weighted autoencoders, 2015.

[7] Jeremias Knoblauch, J. Jewson, and T. Damoulas. Generalized variational inference. *ArXiv*, abs/1904.02063, 2019.