Project report
Improving Explorability in Variational Inference with Annealed Variational Objectives

Mikhail Kurenkov, Timur Chikichev, Aleksei Pronkin

23 October 2020

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

Introduction

Variational Inference reducing representational bias amortized VI, Variational Autoencoders (VAE) Expressive families of variational distributions » losing the computational tractability Reducing the amortization error introduced by the use of a conditional network Non-parametric methods Importance Weighted Autoencoder (IWAE) » multiple samples » computationally difficult

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).

1 Experiments

VAE experiment on MNIST dataset

Same decoder and encoder 2 hidden layers with dimension 300 40 - latent space size LeakyReLU activation function Batch normalization Optimizer - Adam (lr - 1e-3), batch$_size = 64, epochs = 25HVI - 5stochastictransitionoperators(hiddensize - 40)BetaannealingBeta0 = 0.2Gamma = 2e - 4$

2 Results

3 Conclusion

4 Resources

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

Acknowledgements

The project represents the paper [1]. We use [2] as an introduction to the problem and [3] 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] J. Lucas, G. Tucker, R. Grosse, and Mohammad Norouzi. Understanding posterior collapse in generative latent variable models. In *DGS@ICLR*, 2019.

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