Project Proposal: Learning Latent Subspaces in Variational Autoencoders

Artem Shafarostov, Marina Pominova,
Alexander Lyzhov, Elizaveta Lazareva

October 2019

1 Introduction

We consider a generative modeling problem in which data belonging to each class has some latent structure. We want the structure to be easily recoverable from learned latent representations of data, and we also want the ability to generate data with any class-specific property.

”Learning latent subspaces in variational autoencoders” aims to solve this problem with solution grounded in the variational autoencoder - a Bayesian model which can generate the data from latent representations. The paper introduces new latent space structure and optimization scheme to solve the problem stated above.

2 Goal

Our general goal is to carefully reimplement the experimental part of the paper and make the code publicly available. We also intend to reimplement competing approaches to compare the results.

3 Methods

In this project we will compare the results of training three VAE-based models: Conditional VAE [1], Conditional VAE with Information Factorization and Conditional Subspace VAE which is introduced in [2]. We will give a brief description of each of them.

Let x to be the given data points with labels y and let z be a data points representations in some latent space.

Conditional VAE provides a method of structuring the latent space. By encoding the data and modifying the variable y before decoding it is possible to manipulate the data in a controlled way. The objective for Conditional VAE is a lower bound of marginal log-likelihood:

$$\log p_\theta(x|z) \geq \mathbb{E}_{q_\phi(z|x,y)}[\log p_\theta(x|z,y)] - D_{KL}(q_\phi(z|x,y)||p(z)) \tag{1}$$

In addition to Conditional VAE, the network $r_\psi(z)$ can be considered which is trained to predict y from z while $q_\phi(z|x)$ is trained to minimize the accuracy of r_ψ. Then, Conditional VAE's optimization problem (where $q_\phi(z|x,y)$ replaced with $q_\phi(z|x)$) is complemented by optomozation of:

$$\max_\phi \min_\psi L(r_\psi(z)(q_\phi(z|x)), y), \tag{2}$$

where L denotes the cross entropy loss. This model is named as Conditional VAE with Information Factorization.

To describe Conditional Subspace VAE in few words, suppose we are given a dataset of elements (x, y) with $x \in \mathbb{R}^n$ and $y \in Y = \{0, 1\}^k$ representing k features of x. Let $H = Z \times W = Z \prod_{i=1}^k W_i$ denote a probability space which will be the latent space of our model. The Conditional Subspace VAE is trained in a such way that it encodes all the information related to feature i labelled by y_i exactly in the subspace W_i. It can be done by maximizing a form of variational lower bound on the marginal log likelihood of the model along with minimizing the mutual information between Z and Y (all formulas are in [2])

4 Experiments

We will first test the model on toy datasets (e.g. from sklearn.datasets) for sanity check and ease of prototyping. Then we will move on to datasets tested in the paper (CelebA, TFD) including style-transfer settings and possibly do an evaluation on new datasets if it is needed. We will examine approximations and architectural decisions made by authors and try to consider possible alternatives.

References

[1] Diederik P Kingma and Max Welling. Auto-encoding variational bayes. arXiv preprint arXiv:1312.6114, 2013.

[2] Learning Latent Subspaces in Variational Autoencoders. Jack Klys, Jake Snell, Richard Zemel. https://arxiv.org/abs/1812.06190