Student ID : 20194293

Name : Go, Kyeong Ryeol

[AI 502] Stochastic Gradient Variational Bayes and Variational Autoencoder

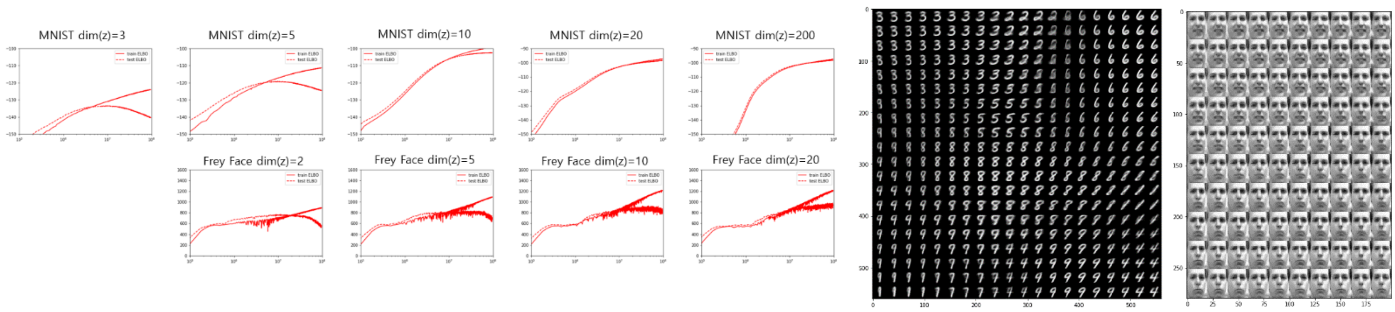
1. Experiment details

The main theme of the paper is that they introduced Auto-Encoding Variational Bayes (AEVB) algorithm that optimize a recognition model using Stochastic Gradient Variational Bayes (SGVB) estimator. The author highlighted the usefulness of the learning process on the fact that the model parameter can be efficiently learned through standard stochastic gradient methods as the approximate posterior inference can be done by the reparameterziation trick. Therefore, I decided to conduct two experiments that quantitatively and qualitatively verify the proposed approach based on Variational Autoencoder. Both of the experiments in based on MNIST and FreyFace dataset.

- For the first experiment, I decided to plot the estimates of variational lower bound along the number of training examples evaluated which is shown in Figure 2 in the paper. This is to track the learning trend of Evidence Lower Bound (ELBO) comparing to previous work and also to observe how the dimension of latent variables affects the value of ELBO when converged.

- For the second experiment, I visualized the learned manifold with two-dimensional latent space, which is shown in Figure 4 in the paper. This is to observe how the data of the same label behaves in the latent space and to capture factors of variation by each of the axis. Here, the inverse CDF of Gaussian was applied to unit square in the latent space as an input to the learned decoder and then the generated outputs were plotted in a single canvas in correspondence to the locations in latent space.

2. Reproduction results



The trend of ELBO curve is reproduced just as same as that referred in the paper. At the early stage, the test ELBO exceeds the train ELBO, but eventually, the situation is reversed. Moreover, the value of the converged ELBO is also reproduced well and it was able to further observe that more latent variables do not result in more overfitting due to the regularization term that encourages the learn posterior $q(z|x)$ not to deviate far from the prior $\mathcal{N}(0, I)$

Learned manifold of MNIST implies that the data with the same label is concentrated in certain region. In case of FreyFace, two factors of variations were captured which are the degree of joy in facial expression and the direction of their face from right to left.

(* Please check the folder named 'plots' to see bigger images in the report and generated images of both MNIST and FreyFace)

3. Discussion

To begin with, how can we quantitively measure the degree of disentanglement between the latent variables? The disentanglement would improve the interpretability so that the latent codes can play a role as significant features that may be utilized in transfer learning, zero-shot learning and so on. Next, how can we appropriately decide the dimension of latent space? This is just set to the hyperparameter to be tuned through heuristic manner or cross-validation. This can be resolved through the approaches based on meta-learning or Bayesian optimization.