GMM-VAE Experiment Report

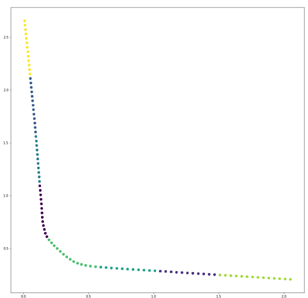
Alptekin Orbay

May 2020

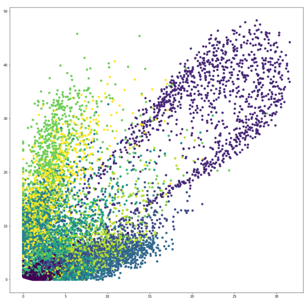
To illustrate problem clearly, I have begun with AutoEncoder and Variational Encoder on synthetic dataset along with MNIST. MNIST is used to check validity of experiments as synthetic dataset is one-dimensional. Note that Synthetic dataset is normalized by standard normalization. For simplicity, MNIST dataset is colored by its labels in figures meanwhile a Gaussian Mixture Model is fitted into Synthetic Dataset to obtain label of each sample for visualization purposes.

1 AutoEncoder

The below figures indicates that MNIST samples cannot be placed into latent space in a meaningful order while Synthetic is able to be mapped. The generated samples show that digits cannot be produced fully. This may be caused by different reasons like lack of regularization, low dimensional latent space etc. The experiments are conducted to make a warm start in the project and compare results with GMM-VAE results.

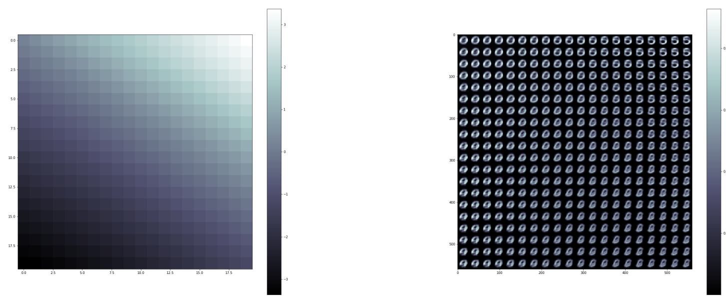


(a) Synthetic



(b) MNIST

Figure 1: Encoded Latent Space of Test Sets

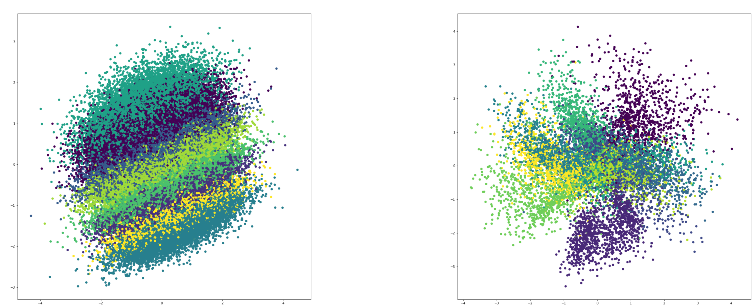


(a) Synthetic (b) MNIST

Figure 2: Generated Samples From Decoder

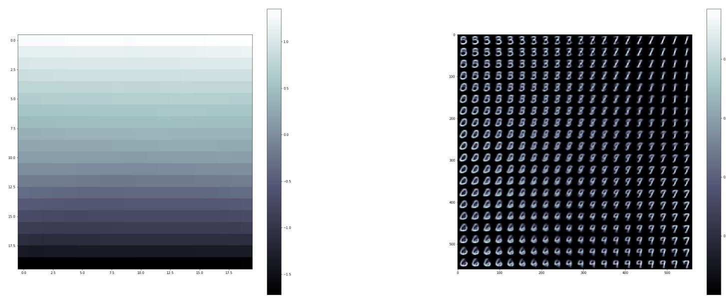
2 Variational AutoEncoder

The figures below show that Variational AutoEncoder is better than Vanilla AutoEncoder in terms of generating samples and mapping into latent space. Now, we have vanilla AutoEncoder and a variational model. So, We can investigate effects of modifying Variational AutoEncoder into GMM-VAE.



(a) Synthetic (b) MNIST

Figure 3: Encoded Latent Space of Test Sets

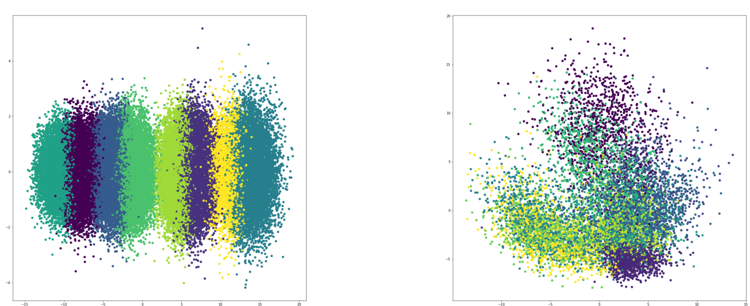


(a) Synthetic (b) MNIST

Figure 4: Generated Samples From Decoder

3 GMM Variational AutoEncoder

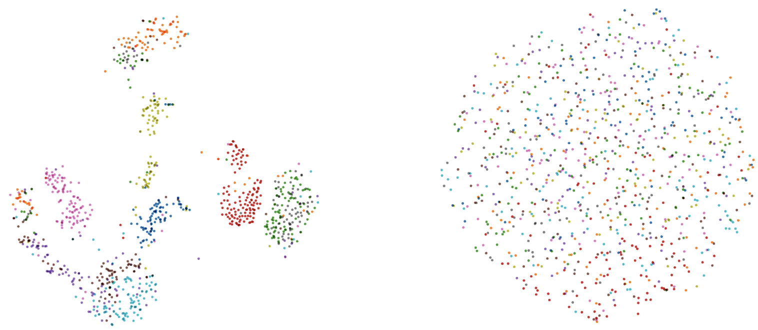
Rui Shu's blog post is very helpful to understand the effects of imposing GMM structure into VAE. In this section, I tried to construct the model as proposed in the paper. As seen in the below figures, Synthetic samples is encoded into latent space according to the their labels. In the second figure, the results proves the correctness of the method. MNIST dataset is trained on GMM-VAE and Vanilla-VAE with 128 latent dimension setting. VAE cannot enforce clusters but GMM-VAE explicitly impose clusters. The illustration is done with https://projector.tensorflow.org.



(a) Synthetic (b) MNIST

Figure 5: Encoded Latent Space of Test Sets into 2-D with PCA

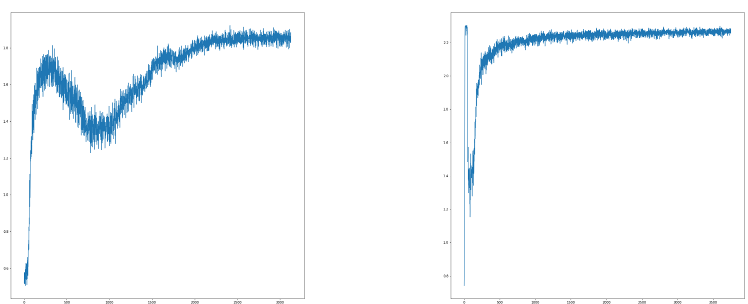
In the training, the cluster number is set to 10 for both datasets. The cluster prior term in loss function is tracked and illustrated in the figure below. There is a regime difference about 1.5 point. It is because the trick proposed in the paper



(a) GMM-VAE (b) VAE

Figure 6: Encoded Latent Space of MNIST Test Set

is used. The term is ignored until it exceeds a threshold. Note that MNIST reaches over 2 point and Synthetic reaches about 1.8 point at convergences.

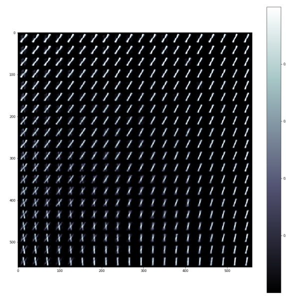


(a) Synthetic (b) MNIST

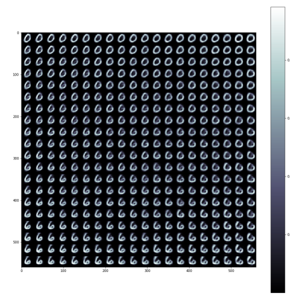
Figure 7: Encoded Latent Space of MNIST Test Set

3.1 Generative Part

For the generative part, two datasets can be generated with two parameters. In the experiments, two dimensional Gaussian Noise and a cluster indicator is arranged to obtain samples. As expected, the cluster indicator really defines class and styles are dependent on noises.

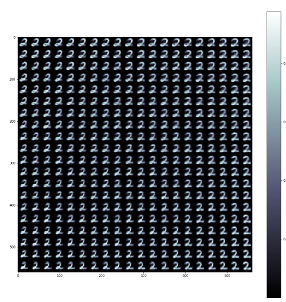


(a) Cluster 9

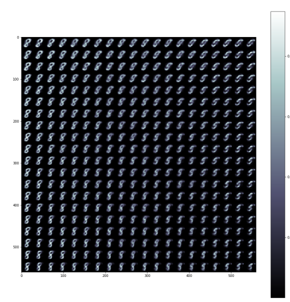


(b) Cluster 2

Figure 8: Generated MNIST Samples with different cluster indicators and same noises.

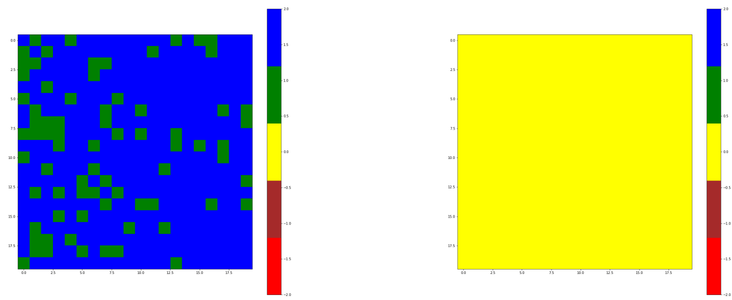


(a) Cluster 8



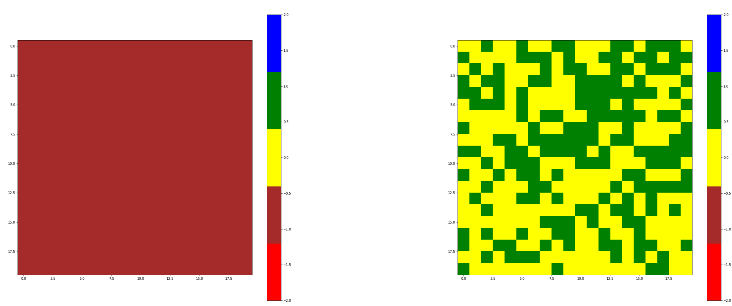
(b) Cluster 3

Figure 9: Generated MNIST Samples with different cluster indicators and same noises.



(a) Cluster 4 (b) Cluster 7

Figure 10: Generated Synthetic Samples with different cluster indicators and same noises.



(a) Cluster 8 (b) Cluster 0

Figure 11: Generated Synthetic Samples with different cluster indicators and same w.