Abstract

Generating synthetic data 'similar to' existing data is a difficult problem - how does a computer know what a cat looks like in order to draw another picture of a cat? One solution is called a Variational Autoencoder, which encodes latent 'catness' seperately from 'dogness', and can generate new cats by sampling something with a similar level of 'catness'. One issue with this is that it often is overly specific - instead coding for 'tail', 'paws', 'furry' etc. Sparsity is meant to make each encoded variable maximally meaningful, so that each code tells us more about our final output. My goal was to modify an existing VAE into a sparse VAE.

Sparse Variational Autoencoder

Ryan Pyle

August 24, 2017

Outline

- Variational Autoencoders
- A Spare VAE
- Results

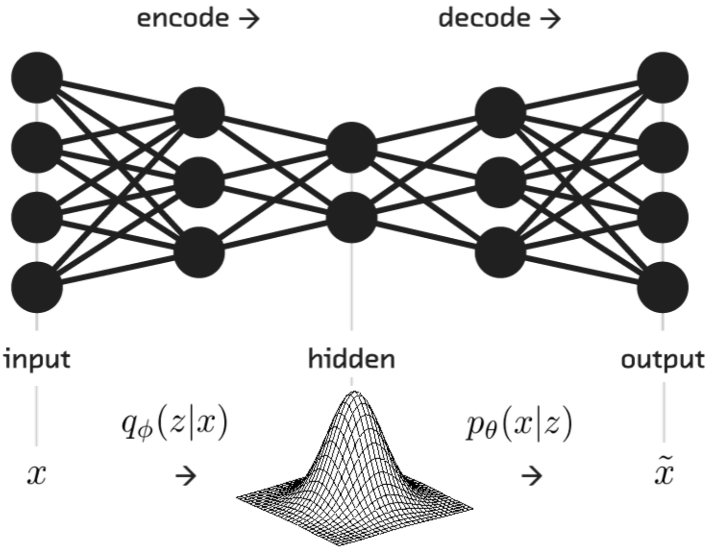
VAE

Variational Autoencoders (VAEs) are a powerful class of machine learning tool for grouping and producing data. The main idea of a VAE is to use bayesian inference to model the underlying distribution of the data, so that new samples can be generated by drawing from that underlying distribution.

VAE

- High dimensional data x
- Low dimensional latent variables z
- Encoder network $q(z|x)$
- Decoder network $p(x|z)$
- Coupled together, by sampling z for p from distribution generated by q
- Can be used to generate new data

VAE


encode →
decode →
input
hidden
output
$q_\phi(z|x)$
$p_\theta(x|z)$
x
→
→
$\tilde{x}$

A Sparse VAE

- Sparsity could give a more sensible hidden layer
- For VAE, sparsity analogue is amount of latent variables z active at once
- One possibility is to give $p(z)$ a 'sparse' prior
- Issue : VAE is tractible because of simplifications made assuming $p(z) \sim N(0, 1)$

A Sparse VAE

- In VAE, in order to sample we wish to have P(Z) be a known distribution eg N(0,1)
- Q(Z|X) is constrained to be 'like' P(Z) in terms of KL divergence
  - This is meant to enforce Q(Z|X) $=$ P(Z)
- This is only true if the distributions are additive
- This means we should use an additive, or Levy Alpha Stable distribution

A Sparse VAE

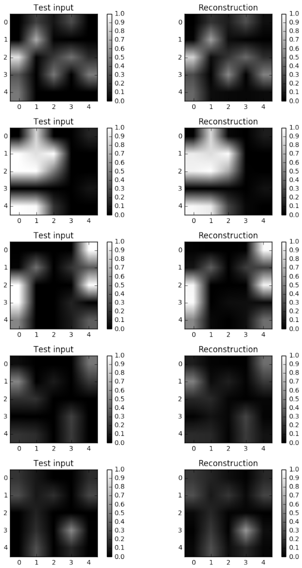
The only real choice is the Cauchy Distribution, giving objective function

$$C = E_{z\sim Q}[log(P(X|z)] - D_{kl}(Q(Z|X)||P(Z)) \tag{1}$$

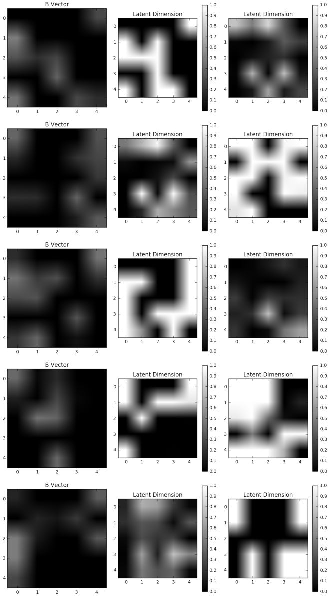
The first term can be solved in terms of log likelihood or approximated by using SSE, while

$$D_{kl}(Q(Z|X)||P(Z)) = E_{z\sim Q}[log(Q(Z|X)) - log(P(Z))] \tag{2}$$

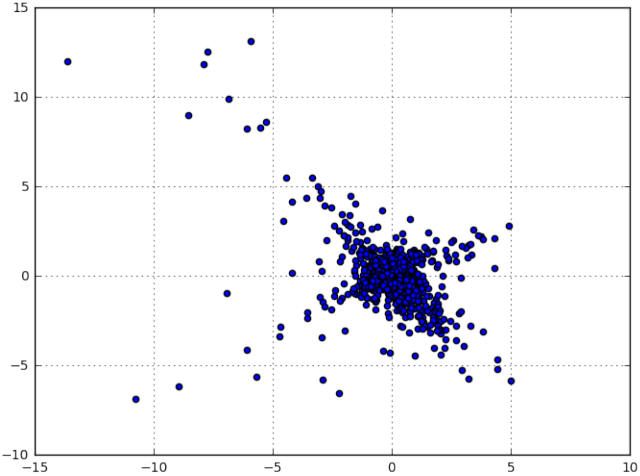
Synthetic Data - Normal VAE



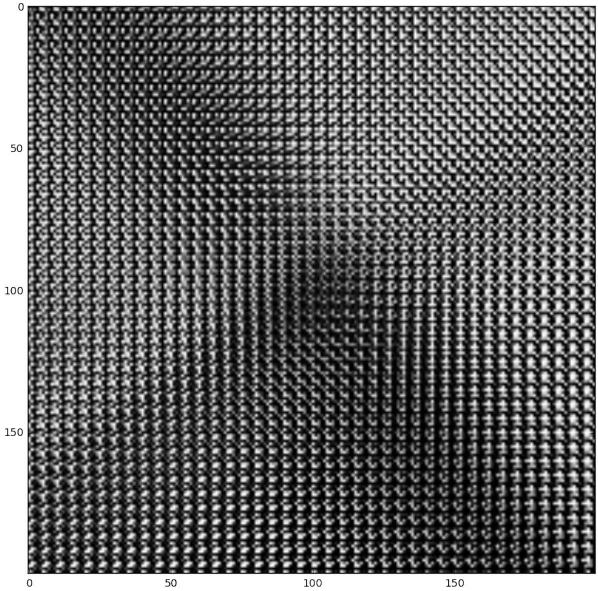
Synthetic Data - Normal VAE



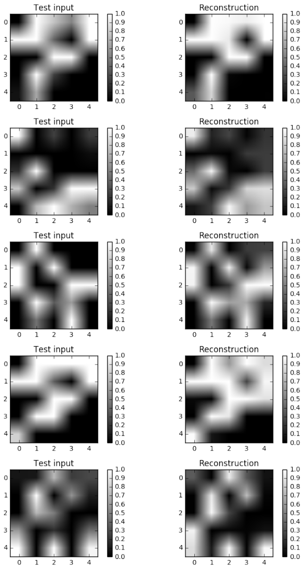
Synthetic Data - Normal VAE



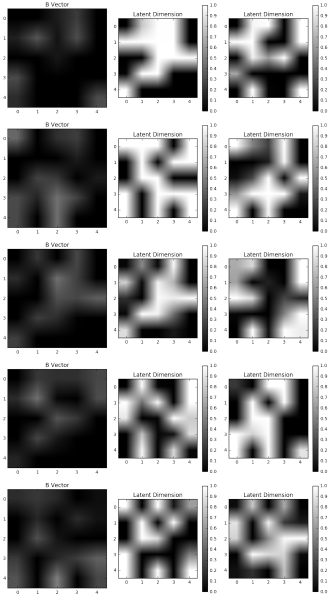
Synthetic Data - Normal VAE



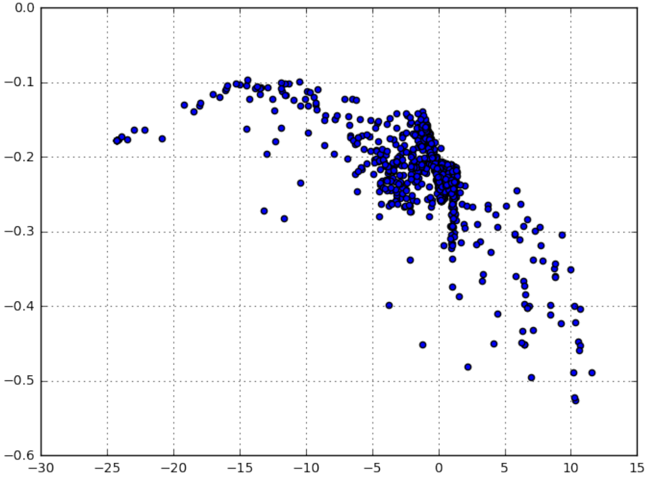
Synthetic Data - Sparse VAE



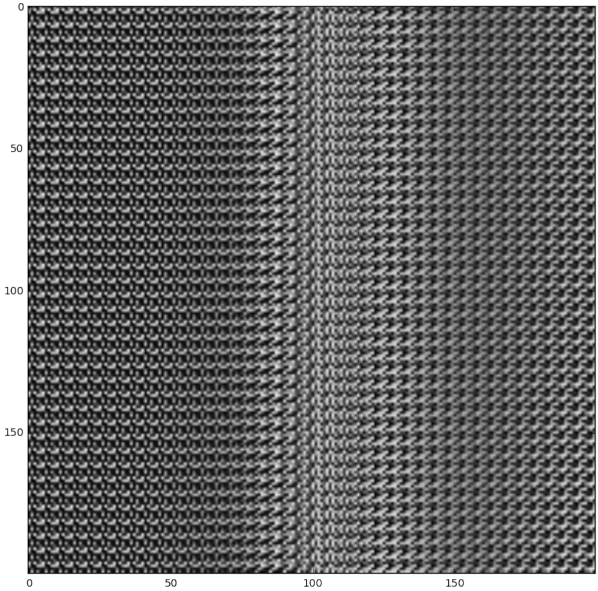
Synthetic Data - Sparse VAE



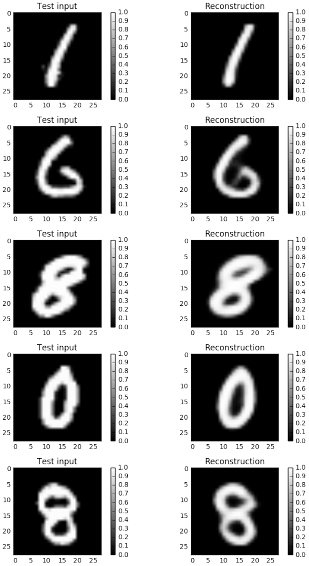
Synthetic Data - Sparse VAE



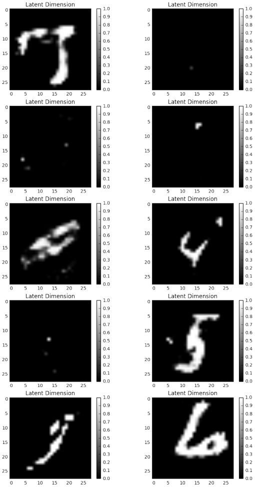
Synthetic Data - Sparse VAE



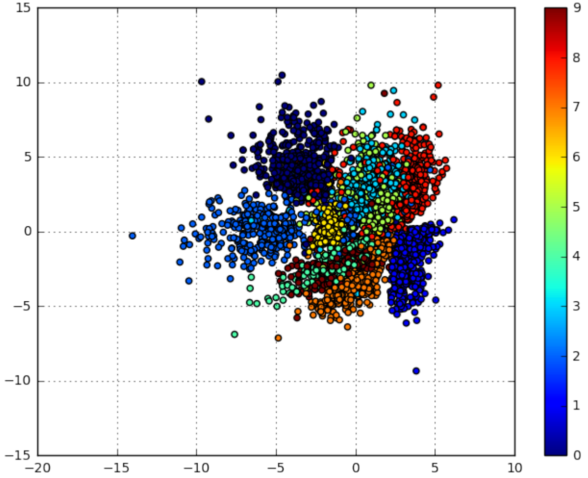
MNIST - Normal VAE



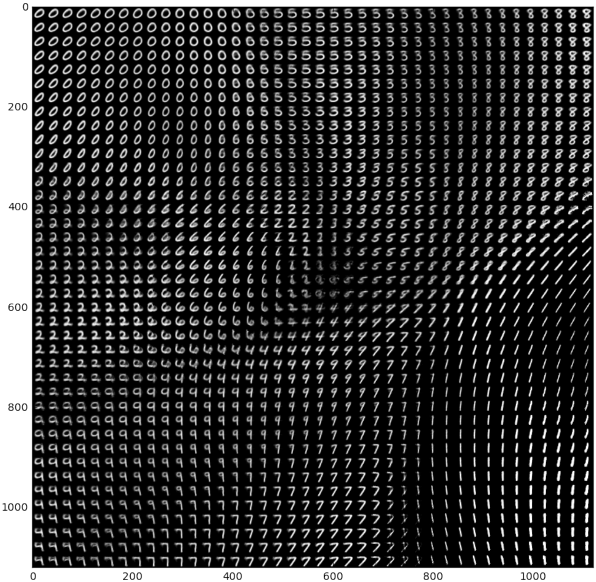
MNIST - Normal VAE



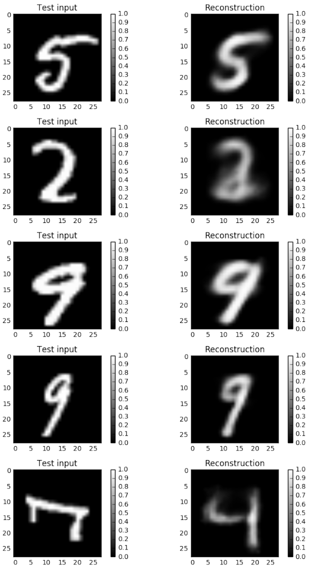
MNIST - Normal VAE



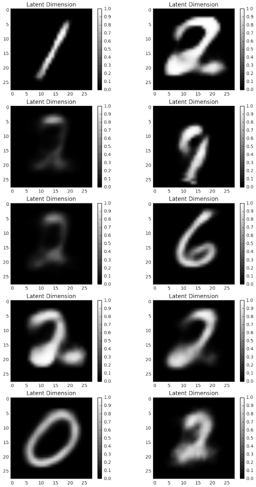
MNIST - Normal VAE



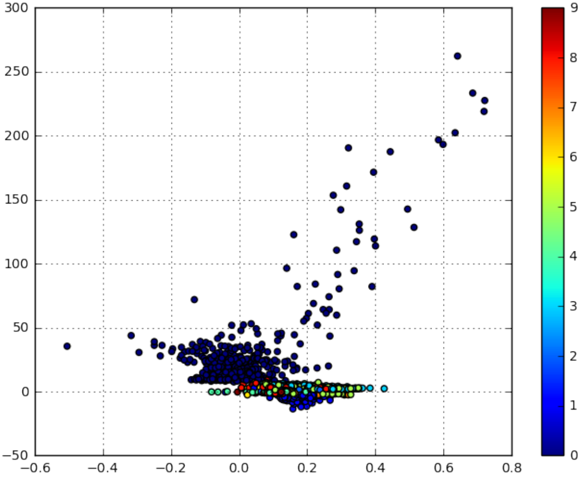
MNIST - Sparse VAE



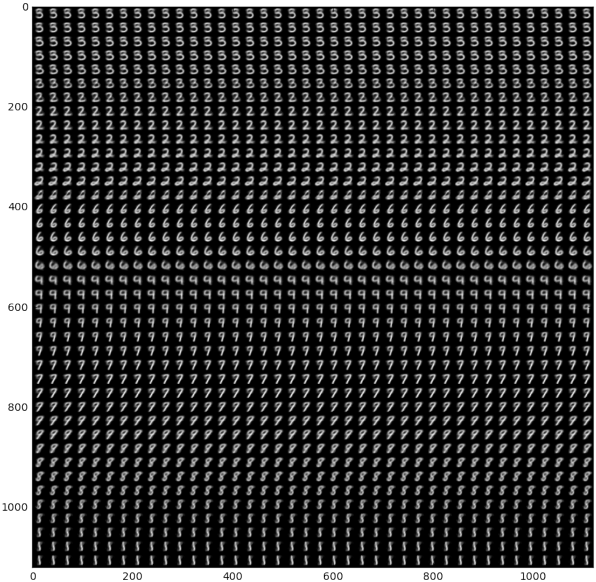
MNIST - Sparse VAE



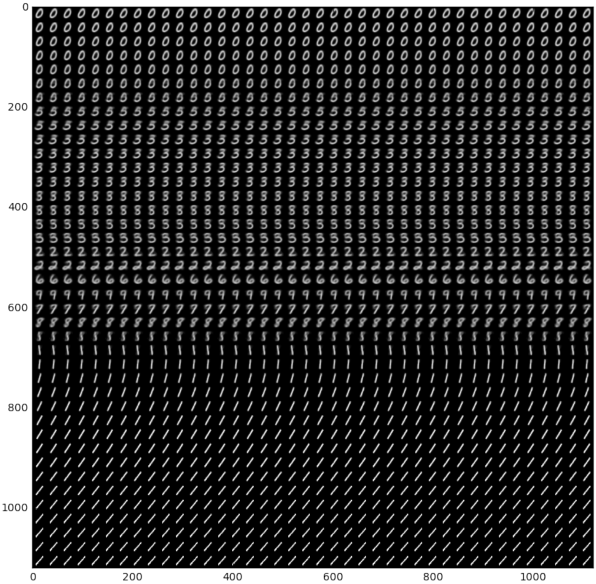
MNIST - Sparse VAE



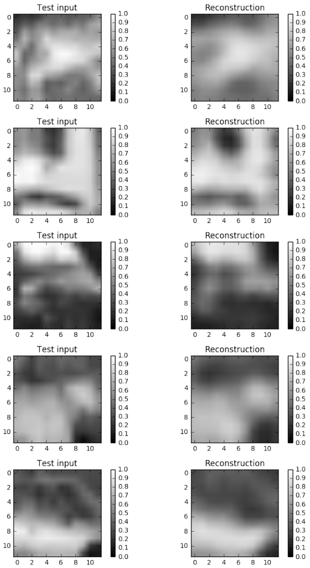
MNIST - Sparse VAE



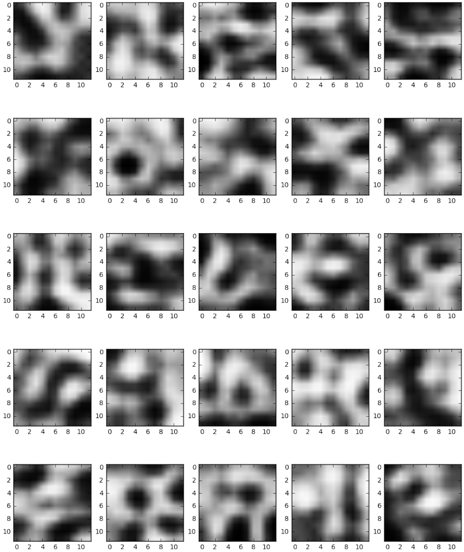
MNIST - Sparse VAE



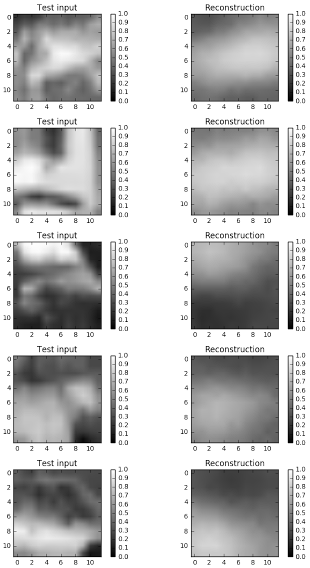
CIFAR - Normal VAE



CIFAR - Normal VAE



CIFAR - Sparse VAE



CIFAR - Sparse VAE