Variational Autoencoder

Dominik Seitz

November 13, 2019

Table of Content

Introduction

- Deep learning based generative models have had huge success producing highly realistic images, texts and sounds
- Examples for deep generative models are Generative Adversarial Networks and Variational Autoencoders

Introduction

Introduction

- How can we perform efficient inference and learning in directed probabilistic models, in the presence of continuous latent variables with intractable posterior distributions, and large datasets?
- Variational Bayes approach involves the optimization of an approximation to the intractable posterior
- Allows to efficiently learn the model parameters, without the need of expensive iterative inference schemes (such as MCMC) per datapoint

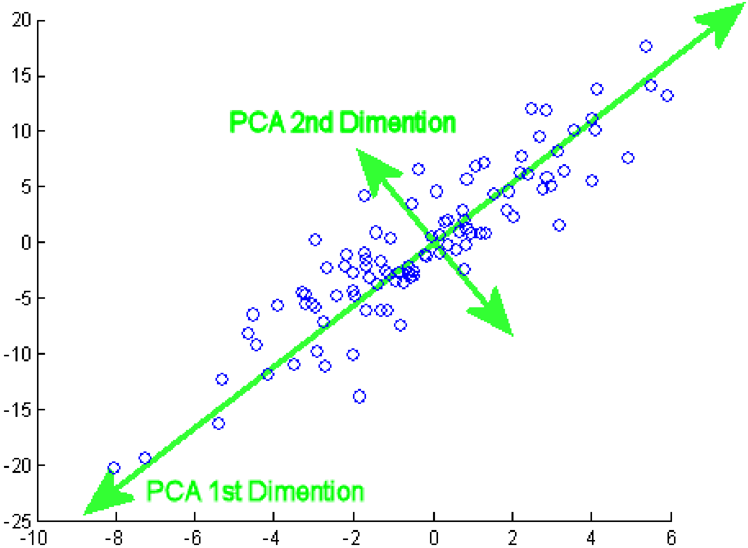
Objective of VAEs

- Reduce dimensionality / find a latent representation of data
- Be able to sample from this latent distribution to generate new unseen data
- Avoid very costly approaches usually used to perform this task

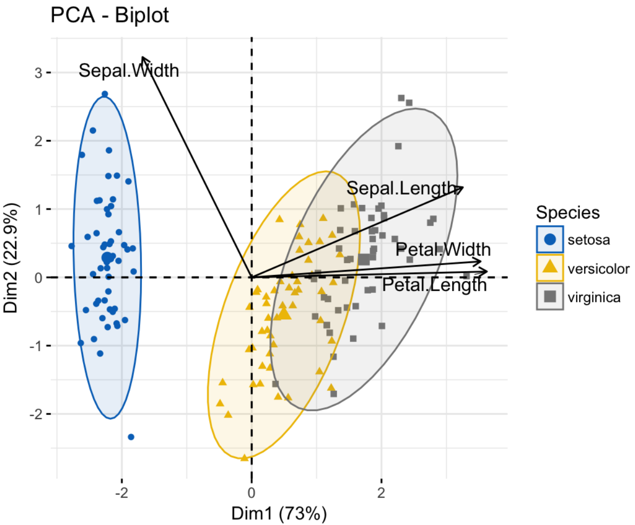
Basics for understanding VAEs

- Dimensionality Reduction
- Variational Bayes
- Kullback Leibler-Divergence
- Latent Variable Spaces

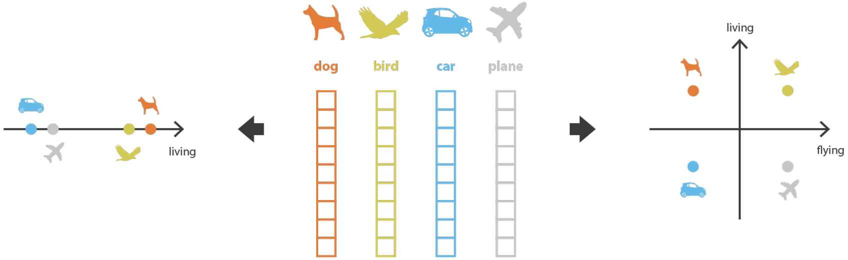
Dimensionality Reduction: PCA



Dimensionality Reduction: PCA



Dimensionality Reduction



near optimal encoding
in one dimension
(too much information lost)

initial data with many features

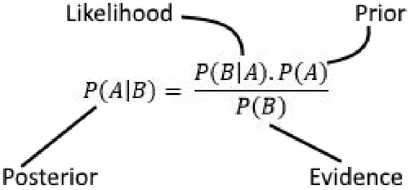
near optimal encoding
in two dimensions
(less information lost)

Variational Bayes

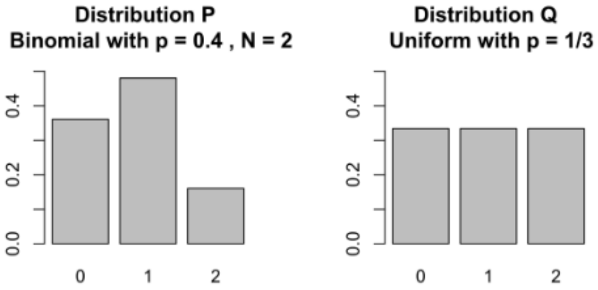
First some basics

- Bayes Theorem
- Prior/Likelihood/Posterior
- Kullback-Leibler Divergence

Bayes Theorem



Kullback-Leibler Divergence: An Example

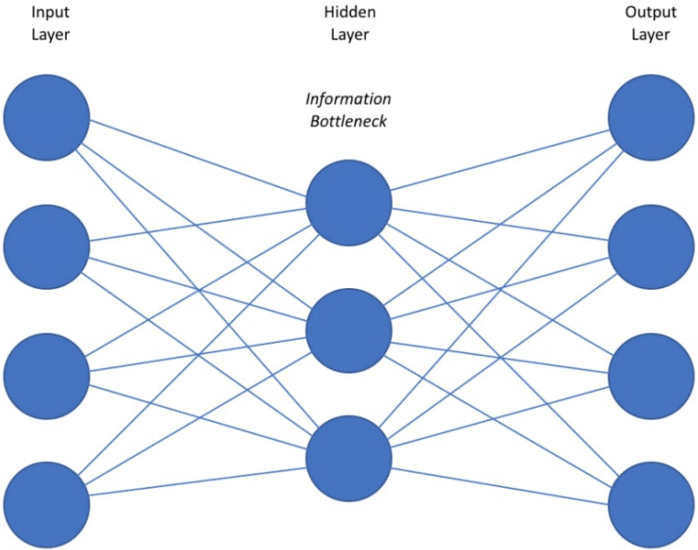


x	0	1	2
Distribution *P*(x)	0.36	0.48	0.16
Distribution *Q*(x)	0.333	0.333	0.333

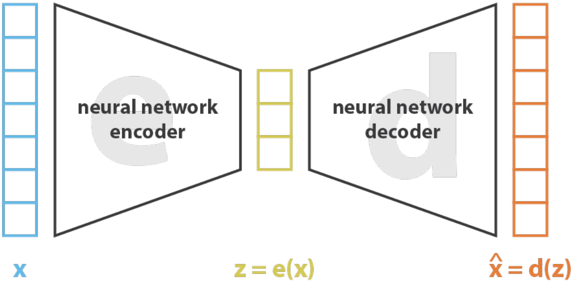
Kullback-Leibler Divergence: An Example

$$
\begin{aligned}
D_{\mathrm{KL}}(P \parallel Q) &= \sum_{x \in \mathcal{X}} P(x) \ln\left(\frac{P(x)}{Q(x)}\right) \\
&= 0.36 \ln\left(\frac{0.36}{0.333}\right) + 0.48 \ln\left(\frac{0.48}{0.333}\right) + 0.16 \ln\left(\frac{0.16}{0.333}\right) \\
&= 0.0852996 \\
D_{\mathrm{KL}}(Q \parallel P) &= \sum_{x \in \mathcal{X}} Q(x) \ln\left(\frac{Q(x)}{P(x)}\right) \\
&= 0.333 \ln\left(\frac{0.333}{0.36}\right) + 0.333 \ln\left(\frac{0.333}{0.48}\right) + 0.333 \ln\left(\frac{0.333}{0.16}\right) \\
&= 0.097455
\end{aligned}
$$

Basic Autoencoder

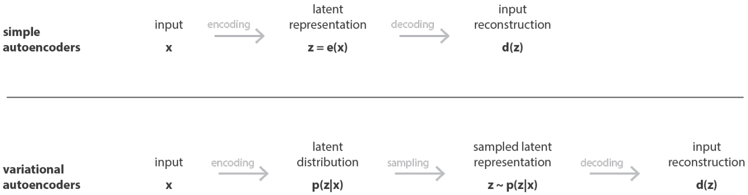


Autoencoders

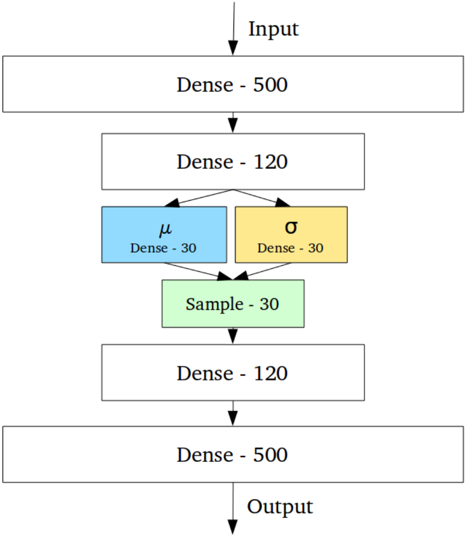


$$\text{loss} = ||\, x - \hat{x} \,||^2 = ||\, x - d(z) \,||^2 = ||\, x - d(e(x)) \,||^2$$

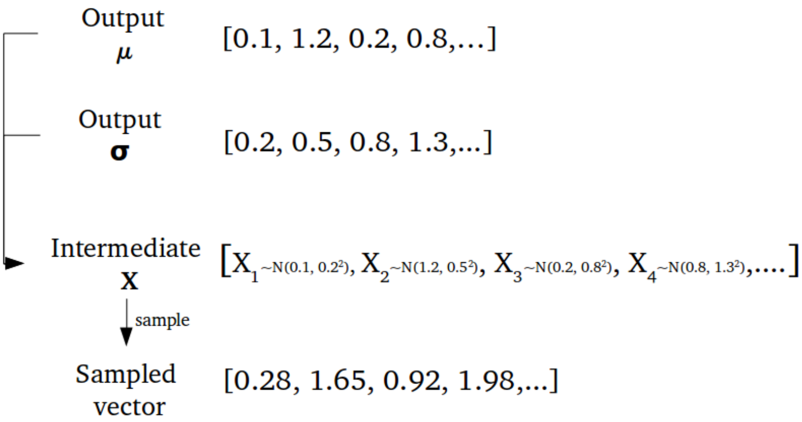
Variational Autoencoders



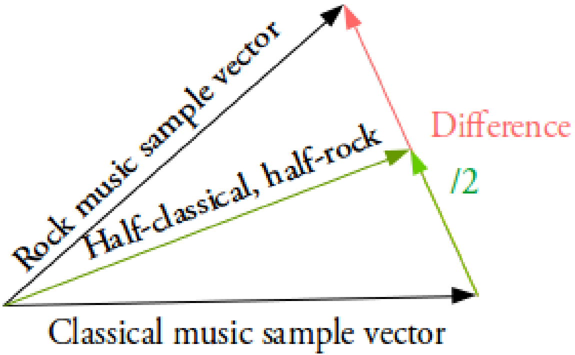
VAE Architecture Example



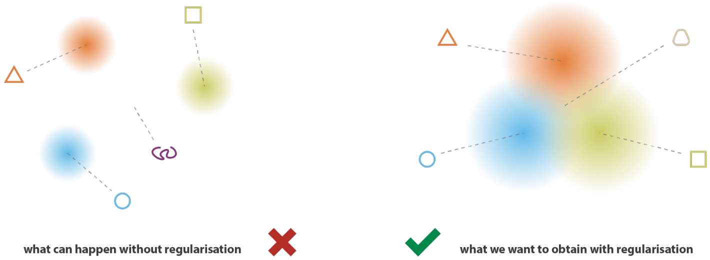
VAE Architecture Example



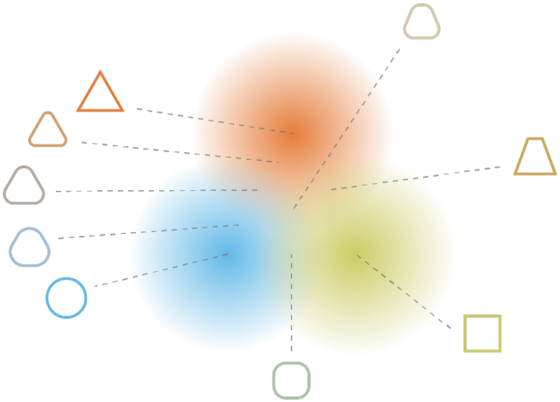
VAE Interpolation Potential



VAE: Advantage of regularization



VAE: Advantage of regularization



Summary

- Dimensionality reduction is the process of reducing/combining features that describe data
- Autoencoders are neural networks condense down the feature space and reconstruct it again with minimal information loss but are known to overfit
- VAE, in contrast, return a distribution given a sample (not a point like normal AE) and hence are more robust