Statistical Data Analysis

Dr. Jana de Wiljes

2. Februar 2022

Universität Potsdam

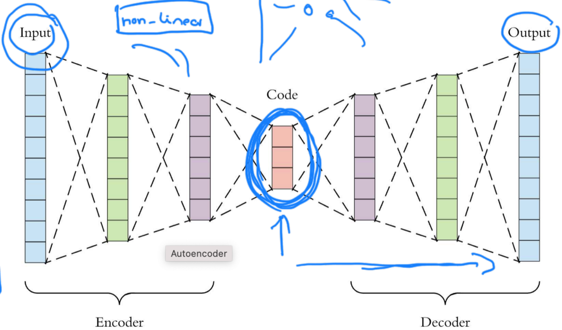
Autoencoders

Autoencoders

Snapshot information

- unsupervised artificial neural network (feed forward)
- Two steps:
  - **Encoder:** learns how to efficiently compress and encode data
  - **Decoder:** learns how to reconstruct the data back from the reduced encoded representation to a representation that is as close to the original input as possible.

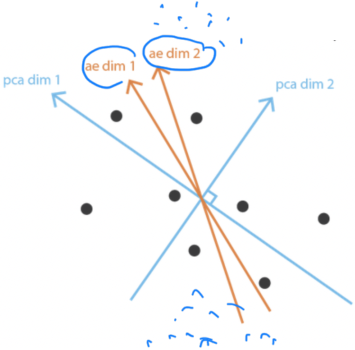
non-linear



L(Output - Input)

PCA vs Autoencoders

- PCA is essentially a linear transformation but Auto-encoders are capable of modelling complex non linear functions.
- PCA features are totally linearly uncorrelated while autoencoded features might have correlations
- PCA is faster and computationally cheaper than autoencoders.
- Autoencoder is prone to overfitting due to high number of parameters. (though regularization and careful design can avoid this)



Linear Autoencoders

Set up: have Data $Y \in \mathbb{R}^{n \times N}$

$$\boxed{Y_0 = Y - \bar{Y} 1_N^T}$$

$$x_i = W^T y_i, \quad x_i \in \mathbb{R}^m, \quad y_i \in \mathbb{R}^n$$

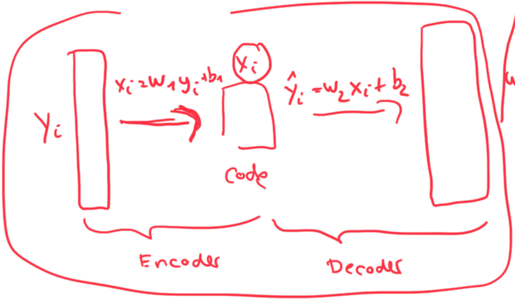
$$W \in \mathbb{R}^{n \times m}$$

PCA:

$$\boxed{p_1 = \max_{w_1} w_1^T Y_0 Y_0^T w_1}$$

$$\text{s.t. } \boxed{w_1^T w_1 = 1}$$

$$Y_0 Y_0^T = \boxed{P \Lambda P^{-1}} = \boxed{P \Lambda P^T}$$



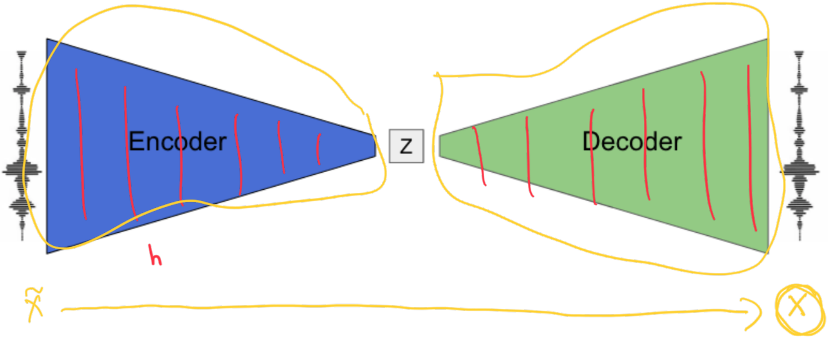
$$\min_{W_1, b_1, W_2, b_2} \| Y - (W_2 (\overbrace{W_1 Y + b_1 \cdot 1_N^T}^{X}) + b_2 \cdot 1_N^T) \|$$

$$\boxed{\min_{W_1, W_2} \| Y_0 - W_2 W_1 Y_0 \|}$$

Linear Autoencoders

Linear Autoencoders

General Autoencoders



$$\| A\tilde{x} - x \|$$

$$\| f(\tilde{x}) - x \|$$

$$\| g(f(\tilde{x})) - x \|$$

Regularised Autoencoders

$$L(x, g(f(x))) + \Omega(h, x) \tag{1}$$

where $h = \overset{W_1}{f}(x)$ is the encoder and $\overset{W_2}{g}(h)$ the decoder, L is a choice of loss function. The regularisation term can have the form

$$\Omega(h, x) = \lambda \sum_i \|\nabla_x h_i\| \tag{2}$$

An autoencoder with this regularisation is known as contractive autoencoder.

Denoising Autoencoders

Traditionally, autoencoders minimize some function

$$L(x, g(f(x))) \tag{3}$$

while a so called denoising autoencoder (DAE) instead minimizes

$$L(x, g(f(\tilde{x}))) \tag{4}$$