An introduction to Autoencoders

Baptiste Gregorutti

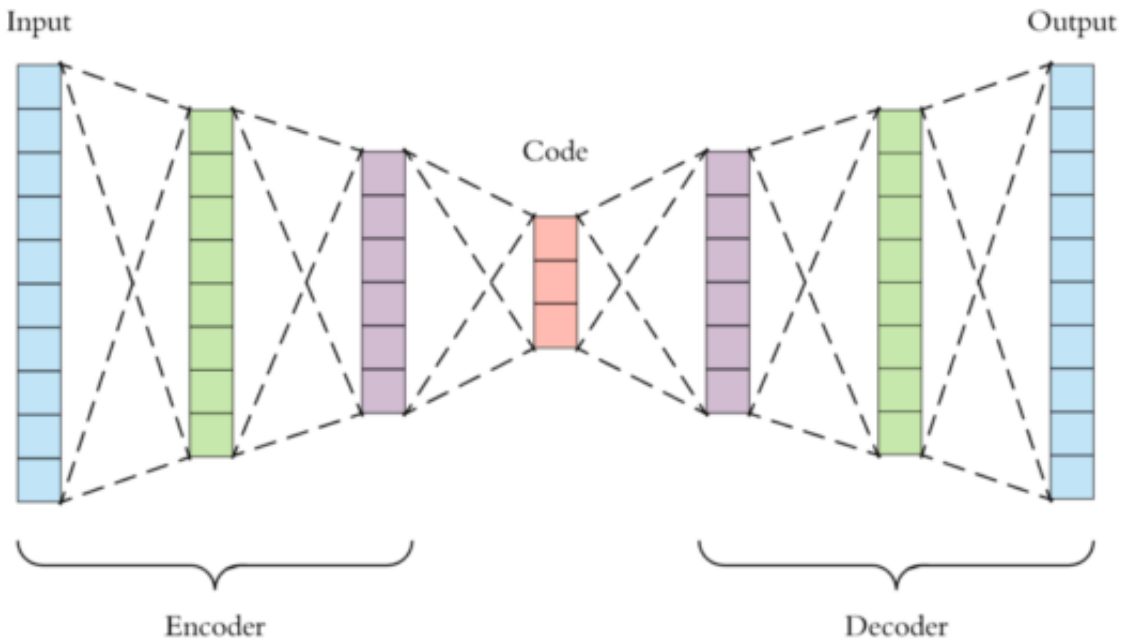
June 2, 2020

Autoencoders Content

- Mathematical definitions
- Relations with the PCA (linear transformation) : AE performs a nonlinear transformation through a nonlinear activation function (e.g. ReLu, sigmoid)
- types of AE: MLP, Deep MLP, Convolutional, Recurrent
- applications to autoencoders

Definition

Autoencoders are a type of self-supervised learning model that can learn a **compressed representation** of input data.



Definition

Let $\mathcal{X}$ be the original space and $\mathcal{F}$ be a latent space.

- Encoding function $\phi : \mathcal{X} \longrightarrow \mathcal{F}$
- Decoding function $\psi : \mathcal{F} \longrightarrow \mathcal{X}$

We have to find the optimal encoding and decoding function by minimizing

$$\phi^\star, \psi^\star \in \underset{\phi,\psi}{\arg\min} \|X - (\psi \circ \phi)X\|^2$$

Two networks:

- Encoding network $\mathbf{z} = \sigma(\mathbf{W}\mathbf{x} + \mathbf{b})$
- Decoding network $\mathbf{x}' = \sigma(\mathbf{W}'\mathbf{z} + \mathbf{b}')$
- Loss function: $\mathcal{L}(\mathbf{x}, \mathbf{x}') = \|\mathbf{x} - \mathbf{x}'\|^2$

Hyperparameters of autoencoders

There are 5 hyperparameters:

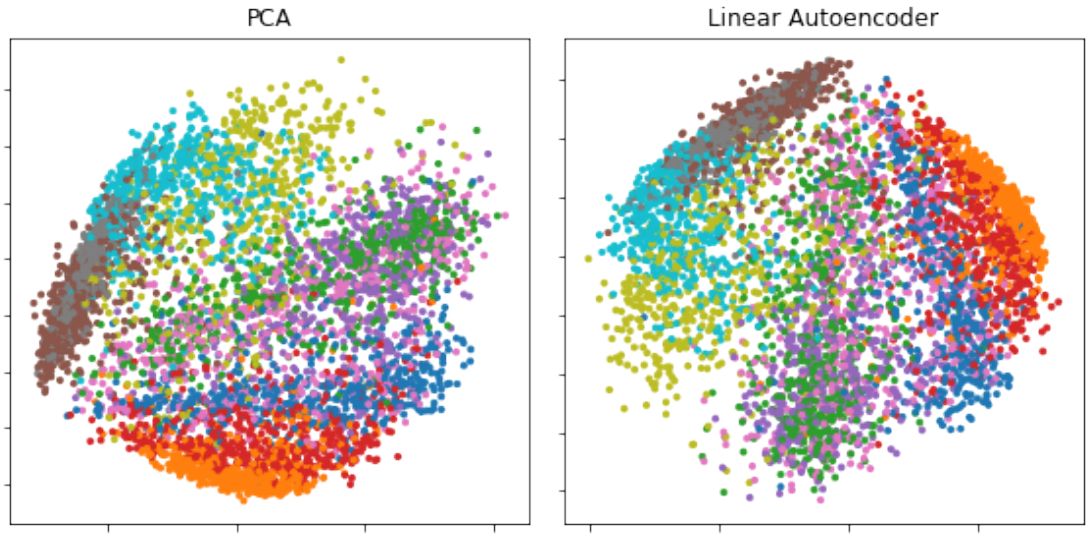
- Dimension of the latent space
- Activation function
- Number of layers
- Number of nodes per layers
- Loss function

In practice ?

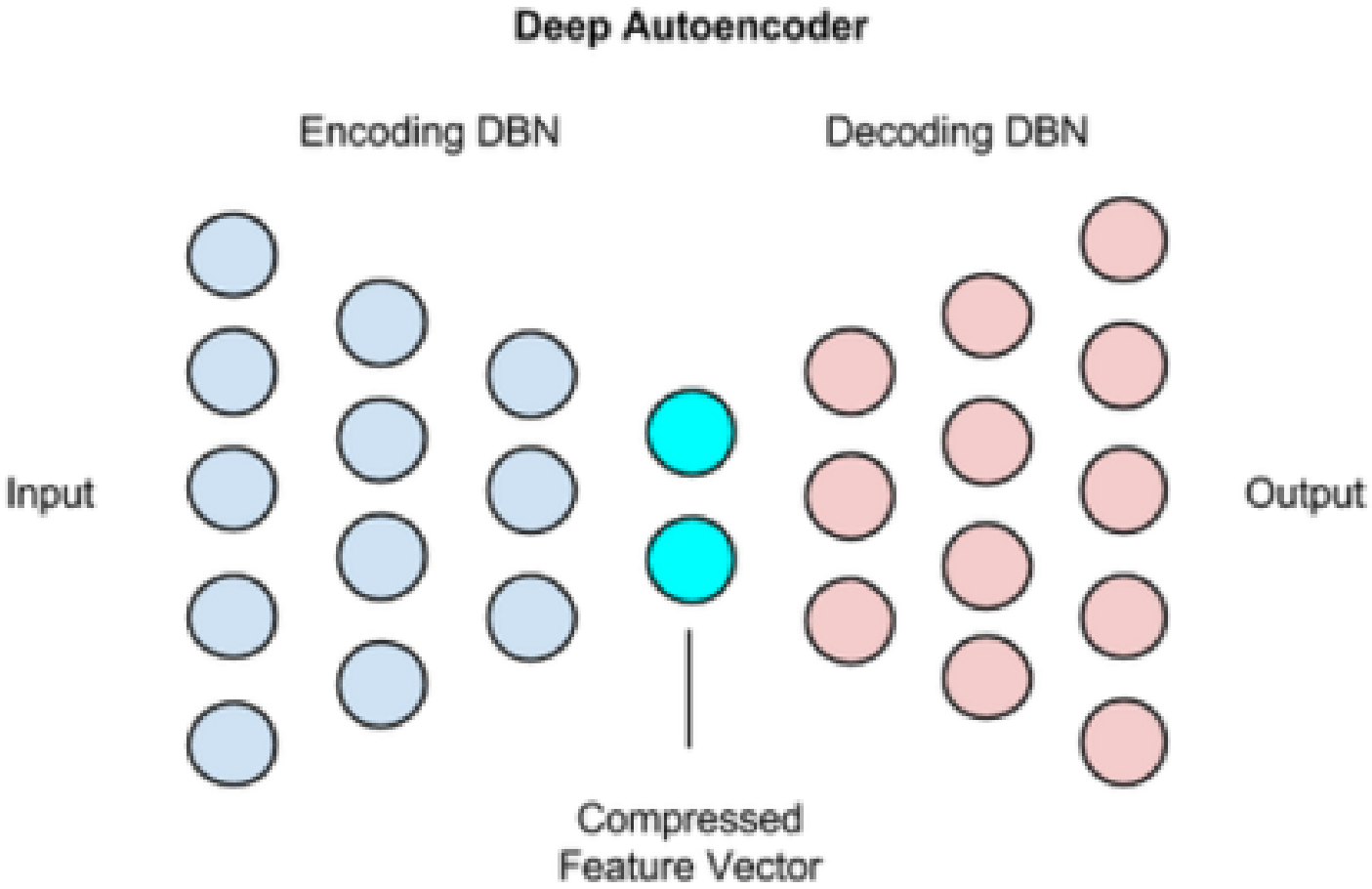
```
1     input_img = Input(input_dim)
2     encoded = Dense(lat_dim, activation='relu')(input_img)
3     decoded = Dense(inp_dim, activation='sigmoid')(encoded)
4     autoencoder = Model(input_img, decoded)
5
```

Relationships with the Principal Components Analysis

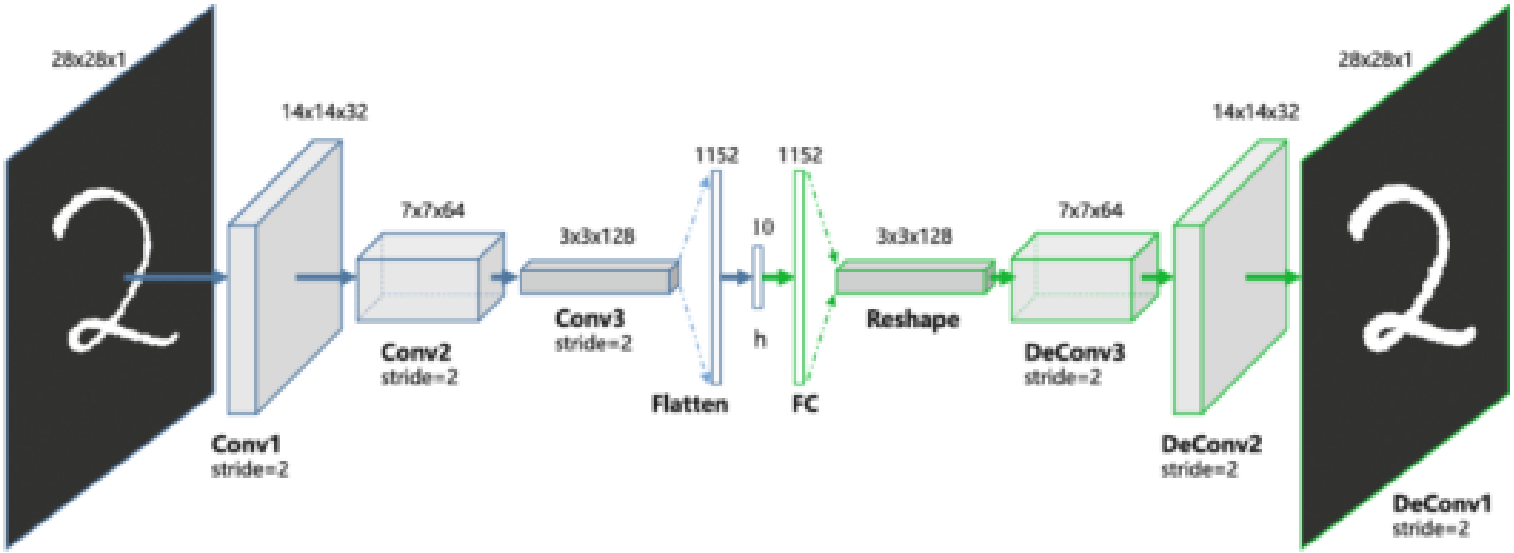
An autoencoder with a **single fully-connected hidden layer**, a **linear activation function** and a **squared error cost function** trains weights that span the same subspace as the one spanned by the principal component loading vectors



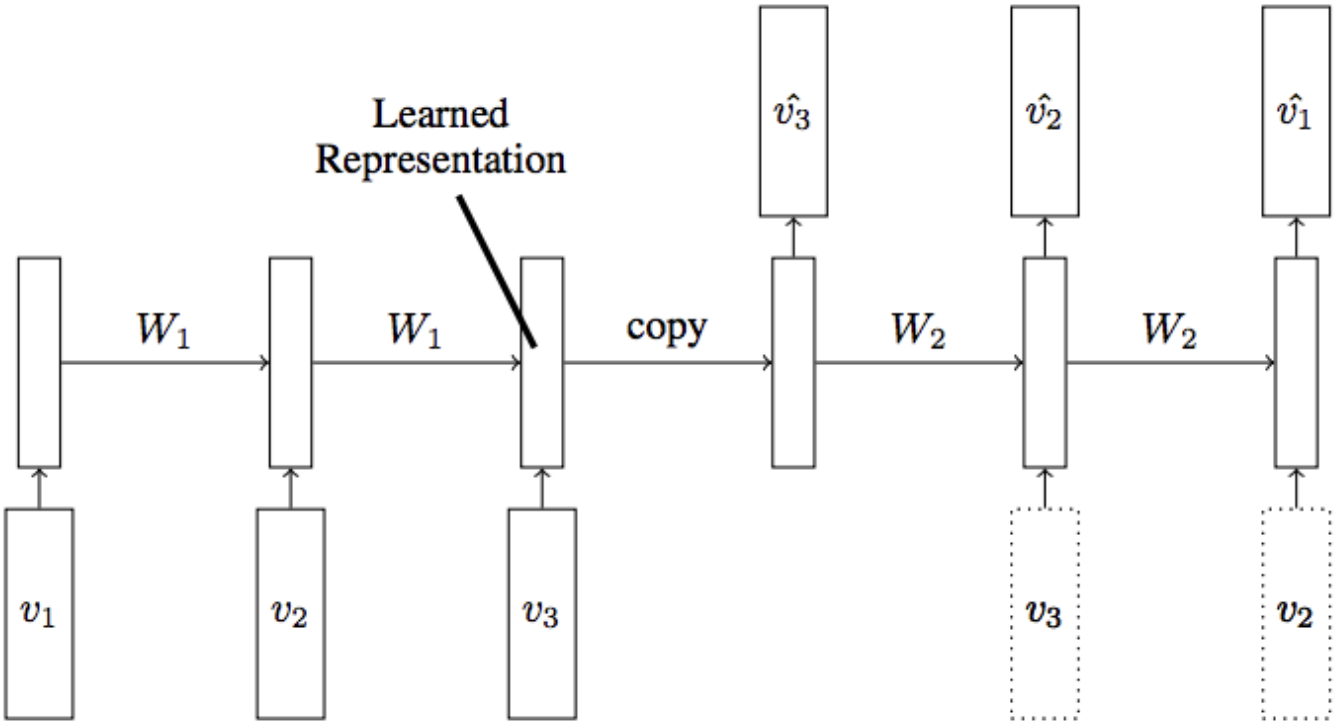
Deep autoencoders



Convolutional autoencoders



Recurrent autoencoders



Applications to autoencoders

Applications:

- Data compression
- Image reconstruction
- Image colorization
- Denoising

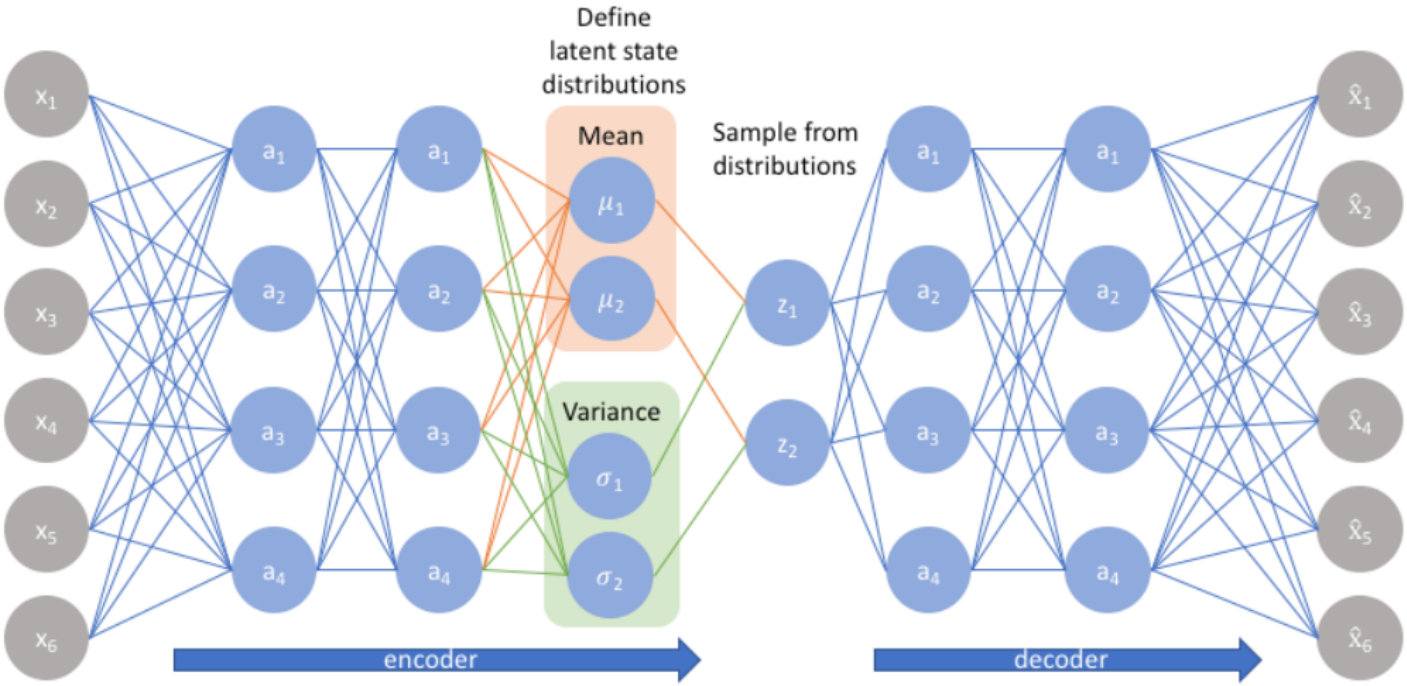
Take away message

- An autoencoder can learn non-linear transformations with a non-linear activation function and multiple layers
- It doesn't have to learn dense layers. It can use convolutional layers to learn which is better for video, image and series data.
- It is more efficient to learn several layers with an autoencoder rather than learn one huge transformation with PCA.

Variational autoencoders

Idea

A variational autoencoders is a **generative model**, mainly used for content generation (images, sounds, etc.)



References

- Keras blog
- Comprehensive Introduction to Autoencoders, Matthew Stewart
- How To Perform Data Compression Using Autoencoders?, Sayantini Deb
- A Gentle Introduction to LSTM Autoencoders, Jason Brownlee
- Understanding Variational Autoencoders, Joseph Rocca