Data Science and Deep Learning (2024)

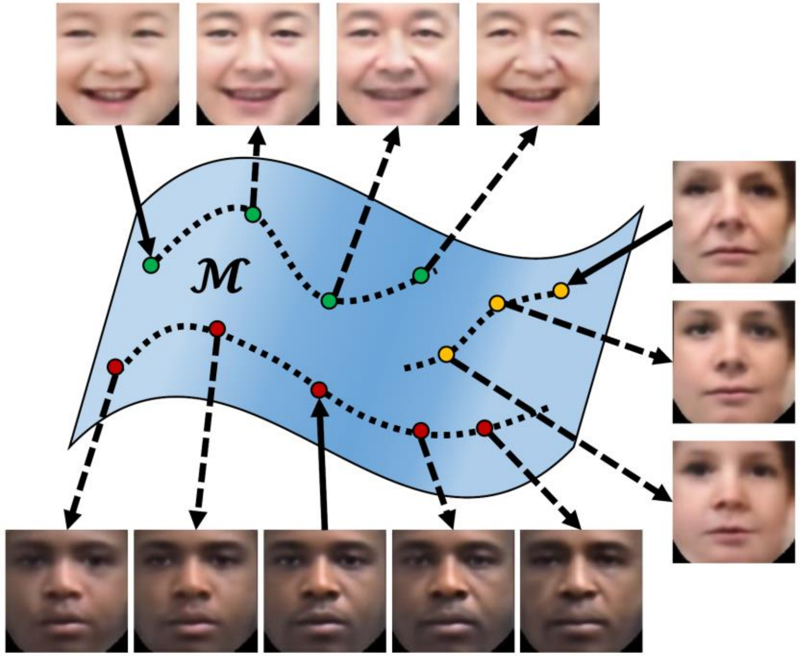
Lecture 3

Representation Learning

Stan Z. Li

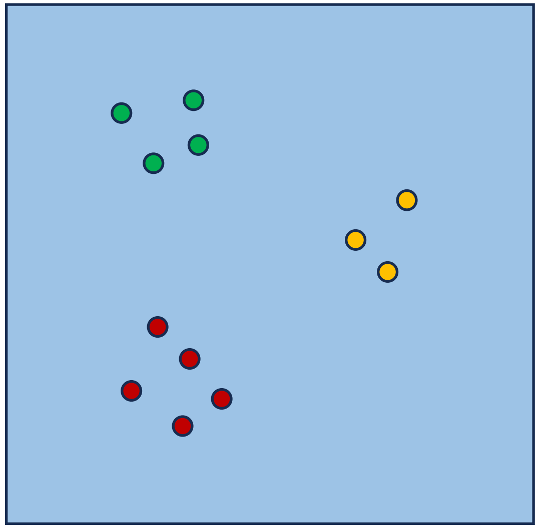
西湖大學
WESTLAKE UNIVERSITY

Encoding and Decoding



Feature Extraction
By Encoder
Neural Network

Data Generation
by Decoder
Neural Network



Samples on low-dim but complex manifold in high-dim data space

Features in lower-dim Euclidean embedding space

Key Problems in Representation Learning

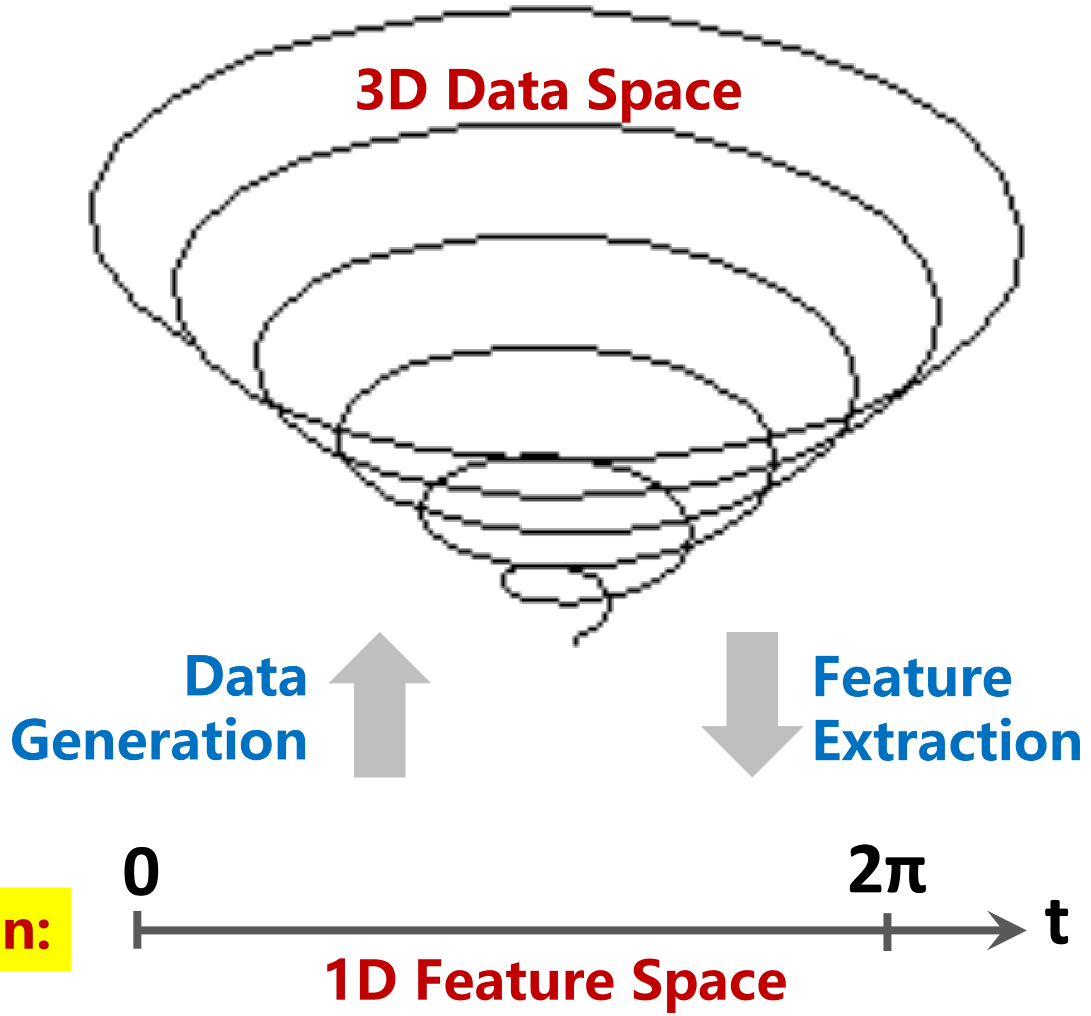
- Q1(a). Representation Learning: How to **encode** observable data into a latent space (embedding)
- Q1(b). Generative Learning: How to **decode** a latent code into observable data
- Q2. Neural Networks: How to design a (parametric) neural network **architecture** to perform the desired (nonlinear) transformation
- Q3. Optimization: How to learn (find) **optimal parameters** for the neural network

Purposes of Representation Learning

1. **Feature Learning**: To automatically discover the representations or features from raw data that are most useful for classification or other tasks.
2. **Dimensionality Reduction**: To reduce the dimensionality of the data to improve computational efficiency, reduce noise, and facilitate data visualization.

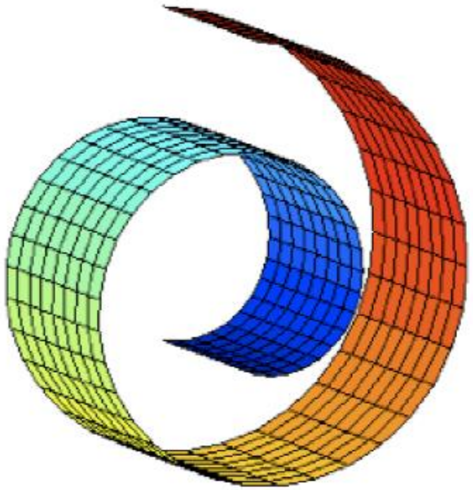
3.......

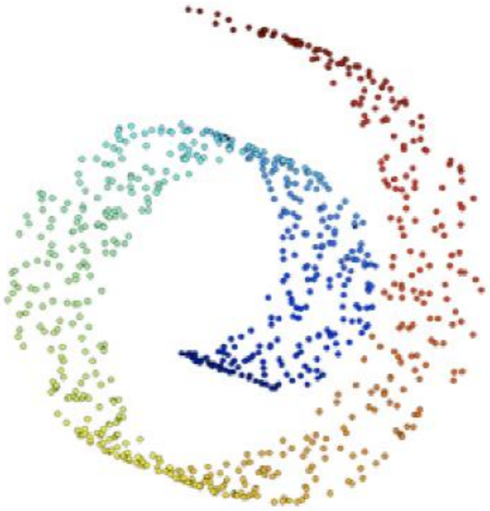
1D Manifold in 3D Space



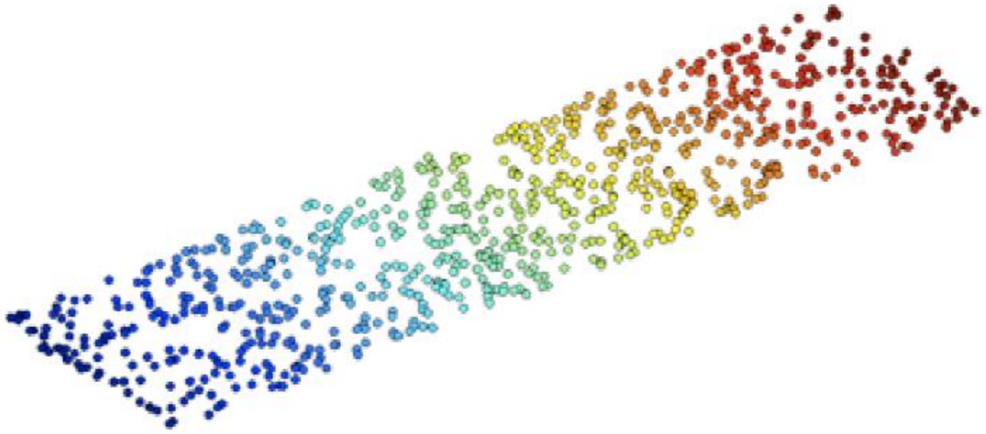
The Best Representation:

2D Manifold in 3D Space





The Best Representation:



Swiss Roll:

$x=\varphi\cos(\varphi)$, $y=\varphi\sin(\varphi)$, $z=\psi$

$1.5\pi \le \varphi \le 4.5\pi$, $0 \le \psi \le 10$

Manifold: 2D rectangle

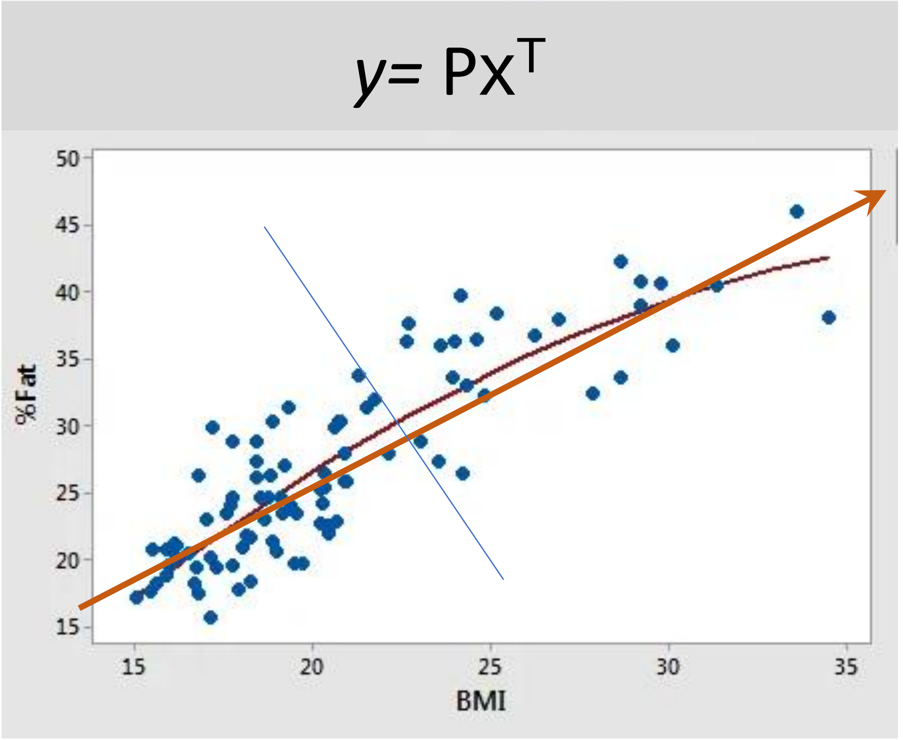
generated by two latent variables φ, ψ

Linear Projection

- **x --- high dimensional data (N*1 vector)**
- **y --- low dimensional feature (n*1 vector)**
- **P --- projection matrix (n*N matrix)**
- **$y = Px^T$**
- **P can be learned from training data,**

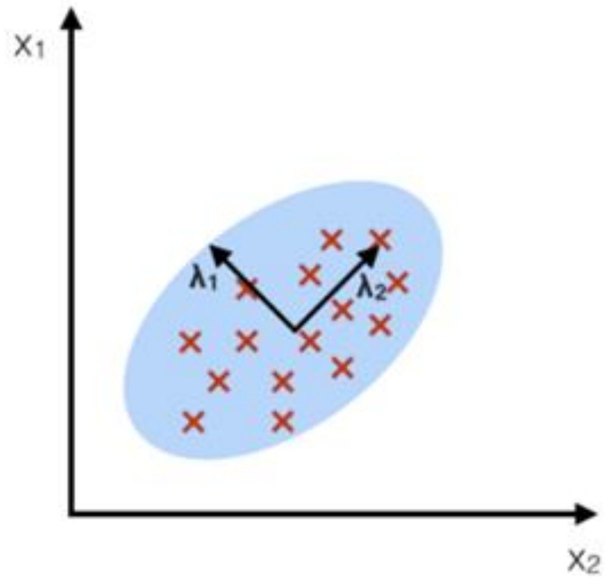
  e.g. Principal Component Analysis (PCA)

Linear Regression Fitting



Principal Component Analysis (PCA)

component axes that
maximize the variance

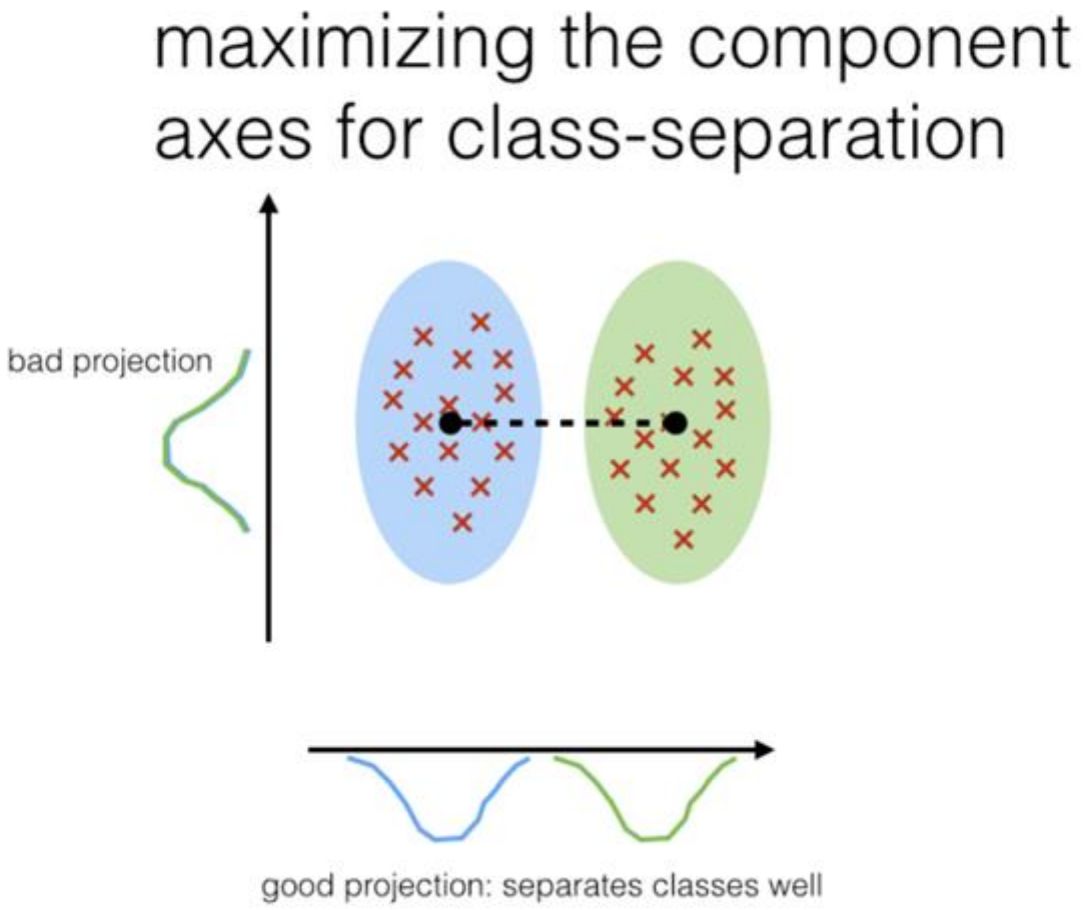


Step 1: $\mathbf{x}-\bar{\mathbf{x}} \rightarrow \mathbf{x}$

Step 2: $\mathbf{y} = \mathbf{P}\mathbf{x}^T = \boldsymbol{f}(\mathbf{x})$

Reconstruction by $\mathbf{x}^T = \mathbf{P}^{-1}\,\mathbf{y}$ **?**

Linear Discriminant Analysis (LDA)



Thanks