12/9/19

EE5601: Representation Learning

- Review
- Principal Component Analysis (PCA) [Shlens 2009.pdf]
  - Motivation
  - Problem formulation
  - Solution
- Motivation for PCA
  - Find principal components of data
  - Identify hidden structure
  - Find directions along which variance is high
  - Efficiently represent data
- How to find these principal components? (Assume zero mean data)

$$\{\underline{x}_i\}_{i=1}^{N} \; ; \; \underline{x}_i \in \mathbb{R}^d$$

$$X = \underset{d}{\begin{bmatrix} | & & | \\ \underline{x}_1 & \cdots & \underline{x}_N \\ | & & | \end{bmatrix}}$$

$$\underset{N}{}$$

- Interested in finding how correlated our data points are. Use covariance matrix to find this.

$$C_{XX} = \begin{bmatrix} \frac{1}{N}\sum_{i=1}^{N} x_{1i}^2 & \cdots & \frac{1}{N}\sum_{i=1}^{N} x_{1i}x_{di} \\ \frac{1}{N}\sum_{i=1}^{N} x_{2i}x_{1i} & \cdots & \frac{1}{N}\sum_{i=1}^{N} x_{2i}x_{di} \\ \vdots & \ddots & \vdots \\ \frac{1}{N}\sum_{i=1}^{N} x_{di}x_{1i} & \cdots & \frac{1}{N}\sum_{i=1}^{N} x_{di}^2 \end{bmatrix}$$