EECS 16B CSM

Bryan Ngo

UC Berkeley

2020-11-15

Logistics

- Linear algebra review
- SVD review
- Applications of SVD
- PCA

Principal Component Analysis

Motivation

- used for statistical analysis
- clustering
- correlation

How to PCA

Given $\boldsymbol{A} \in \mathbb{R}^{n \times m}$, n measurements with m samples,

1. find $\overline{n_i}$ to center $\boldsymbol{A}$ around mean
2. find SVD $\widetilde{\boldsymbol{A}} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^\top$
3. plot eigenvectors/principal components $\boldsymbol{v}_1, \boldsymbol{v}_2$ against centered points
4. data is scaled by σ_1, σ_2
5. more stretched along vector $\implies$ larger correlation

Visualization