EECS 16B CSM

Bryan Ngo

Principal Component Analysis

EECS 16B CSM

Bryan Ngo

UC Berkeley

2020-11-15



Logistics

EECS 16B CSM

Bryan Ng

Principal Component Analysis

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

EECS 16B CSM

Bryan Ngo

Principal Component Analysis

Principal Component Analysis

Motivation

EECS 16B CSM

Bryan Ngc

Principal Component Analysis

- used for statistical analysis
- clustering
- correlation

How to PCA

EECS 16B CSM

Bryan Ng

Principal Component Analysis

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

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

Visualization

EECS 16B CSM

Bryan Ngo

Principal Component Analysis

