281 Live Session

Week 11 — 2023/3/22

Agenda

- Final project updates
- Dimensionality overview
- Exercise: gradient descent

Project Updates

Today: 2 minute show-and-tell with example images

Next update: PCA plots of variance for 1-2 feature vectors

Discussion Questions

- Definition of covariance
    - why are the eigenvectors of the cov. matrix important?
- Relationship between # of data points vs # of features
- What is an eigenface?
- Why is the nearest neighbor different between the pixel value basis vs eigenface basis?
- What is the difference between PCA and tSNE?

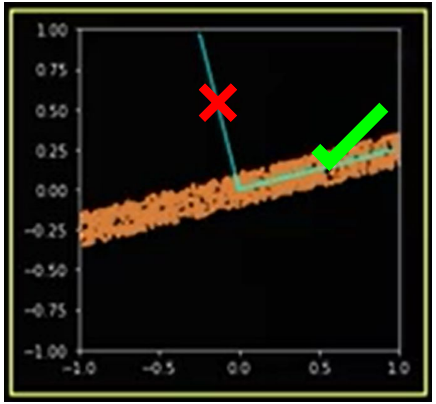
Dimensionality Reduction / Reorganization

Goal is to spread out data along axes of <u>high</u> variance

And compress data along axes of <u>low</u> variance

Ideal axes:

- All (or almost all) variance explained
- Total variance spread across many dimensions



Applications of PCA

Computer Vision

- Image recovery and denoising
- Image composition
- Image colorization
- Image alignment and rectification
- Multi-focus imaging
- Face recognition
- MRI parallelization and background suppression
- Structure from motion
- Motion recovery
- Video denoising and restoration
- Hyperspectral video
- Background/foreground separation

Other

- Neuroscience signal decomposition (EEG)
- Quantitative finance
- Medical data correlation analysis
- Chemical analysis
- General data compression, visualization, and simplification

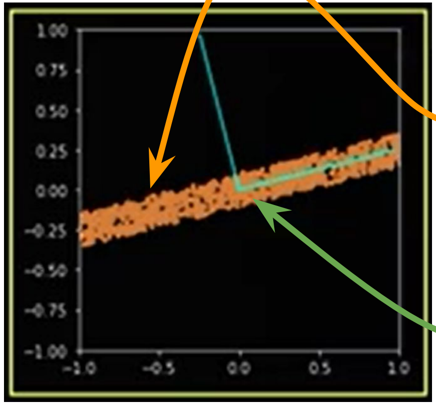
Bouwmans, Thierry, et al. "On the applications of robust PCA in image and video processing." *Proceedings of the IEEE* 106.8 (2018): 1427-1457.

Overview — principal component analysis (PCA)

Co-Variance Definition

$$C = \begin{pmatrix} x_1 & \cdots & x_n \\ y_1 & \cdots & y_n \end{pmatrix} \begin{pmatrix} x_1 & y_1 \\ \vdots & \vdots \\ x_n & y_n \end{pmatrix}$$

$$C = \begin{pmatrix} \sum_i x_i^2 & \sum_i x_i y_i \\ \sum_i x_i y_i & \sum_i y_i^2 \end{pmatrix}$$



$$D = \begin{pmatrix} x_1 & \cdots & x_n \\ y_1 & \cdots & y_n \end{pmatrix}$$

$$\hat{D} = BD$$

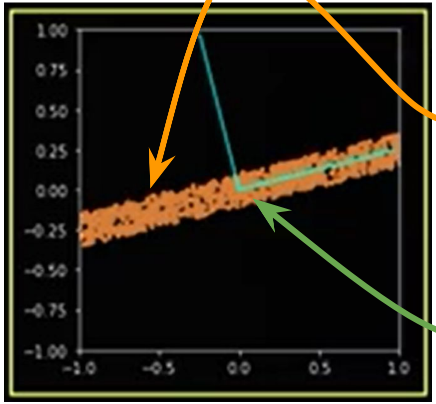
new basis

Overview – principal component analysis (PCA)

Co-Variance Definition

$$C = \begin{pmatrix} x_1 & \dots & x_n \\ y_1 & \dots & y_n \end{pmatrix} \begin{pmatrix} x_1 & y_1 \\ \vdots & \vdots \\ x_n & y_n \end{pmatrix}$$

$$C = \begin{pmatrix} \sum_i x_i^2 & \sum_i x_i y_i \\ \sum_i x_i y_i & \sum_i y_i^2 \end{pmatrix}$$



$$D = \begin{pmatrix} x_1 & \dots & x_n \\ y_1 & \dots & y_n \end{pmatrix}$$

$$\hat{D} = \underbrace{B}_{\text{new basis}} D$$

$C = DD^t$ co-variance matrix (m x m)

$\hat{C} = D^t D$ co-variance matrix (n x n)

$$(DD^t)D\vec{e} = \lambda D\vec{e}$$
$$C(D\vec{e}) = \lambda D\vec{e}$$

$\vec{e}$ is an eigenvector of $\hat{C}$ then $D\vec{e}$ is a eigenvector of C

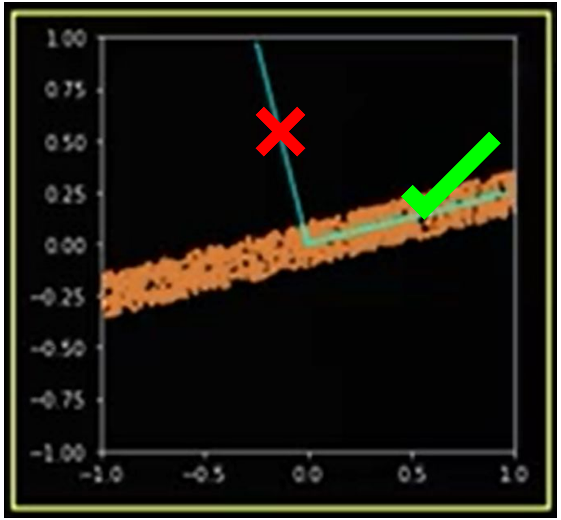
M (number of variables)

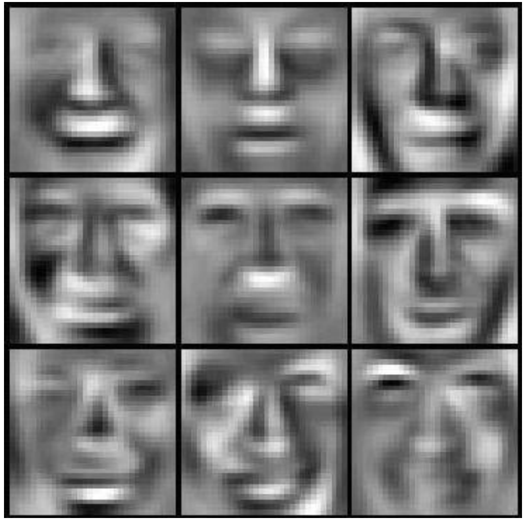
N (number of data points)

Eigenfaces

Basis faces: axes of variance will depend upon the dataset

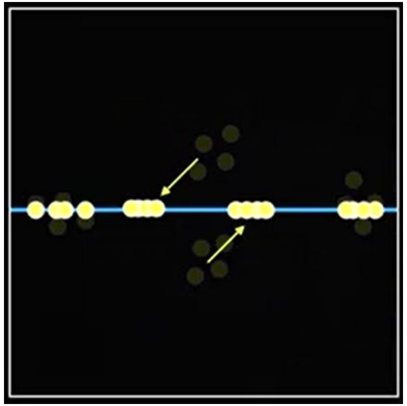
(uniformity of alignment, rotation, lighting, etc)

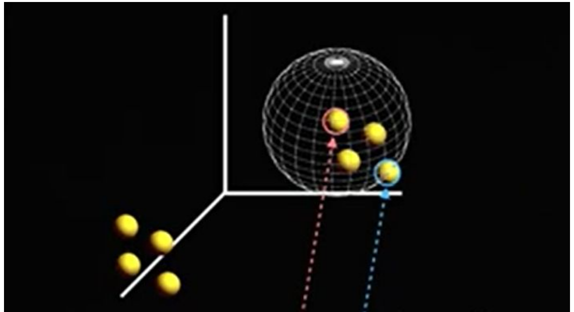




Source: mikedusenberry.com

Overview – t-distributed stochastic neighbor embedding (tSNE)





$$p_{i,j} = \frac{\exp(-\|\vec{x}_i - \vec{x}_j\|^2 / 2\sigma_i^2)}{\sum_{k \neq i} \exp(-\|\vec{x}_i - \vec{x}_k\|^2 / 2\sigma_i^2)}$$

$$C = \sum_i \sum_j p_{i,j} \log\left(\frac{p_{i,j}}{q_{i,j}}\right)$$

Kullback-Leibler

original distribution

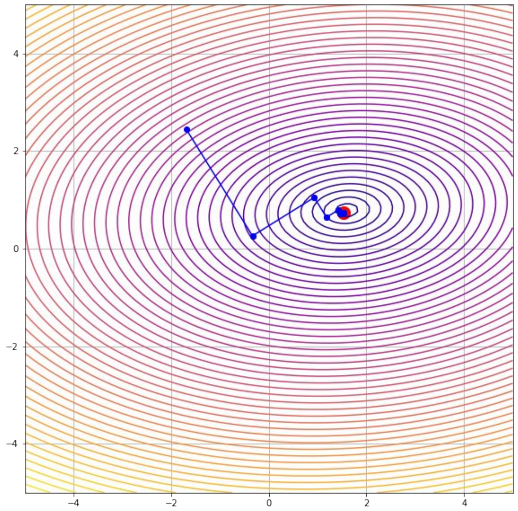
new distribution

same ratio

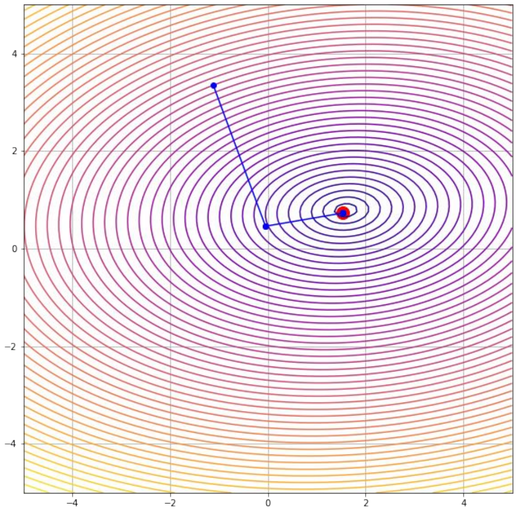
Discussion Questions

- Definition of covariance
    - why are the eigenvectors of the cov. matrix important?
- Relationship between # of data points vs # of features
- What is an eigenface?
- Why is the nearest neighbor different between the pixel value basis vs eigenface basis?
- What is the difference between PCA and tSNE?

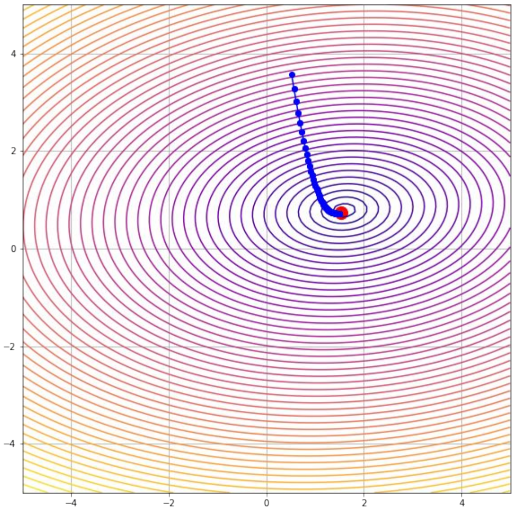
Group Exercise – Gradient Descent



Steepest Descent



Conjugate Gradient Descent



Gradient Descent

Upcoming ToDo's

Finish Assignment 6 (Due March 28^{th})

Start Assignment 7 (Due April 4^{th})

Watch Async lectures for Unit 11

Final Project Updates: PCA plots of variance for 1-2 feature vectors