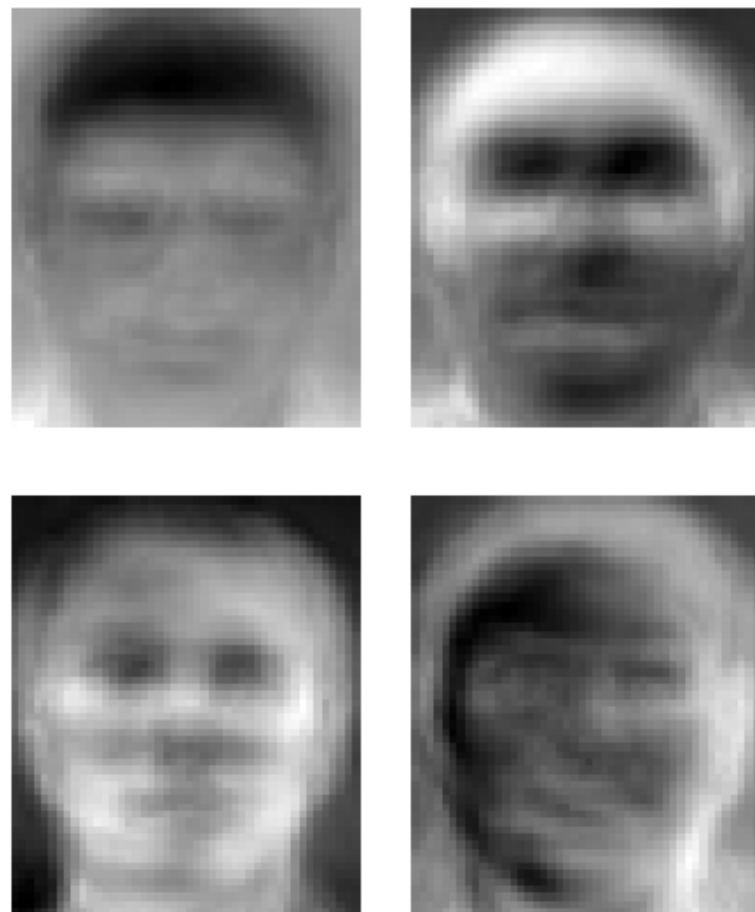

Feature Projection

Key points:

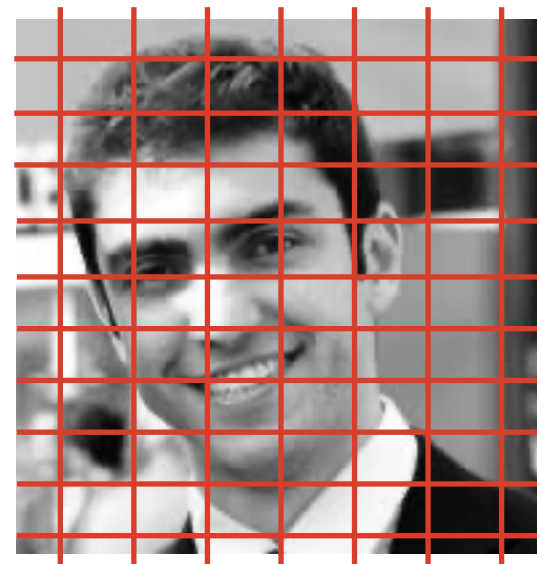
- Intuition for PCA and feature projection and difference to feature selection
- Issues with PCA and alternatives

Feature Projection: Projecting High-Dimensional Data

Eigenfaces: Principal Component Analysis on face images to get “basis vectors” of face image space



Linear combination of basis vectors



One image is one column



faces



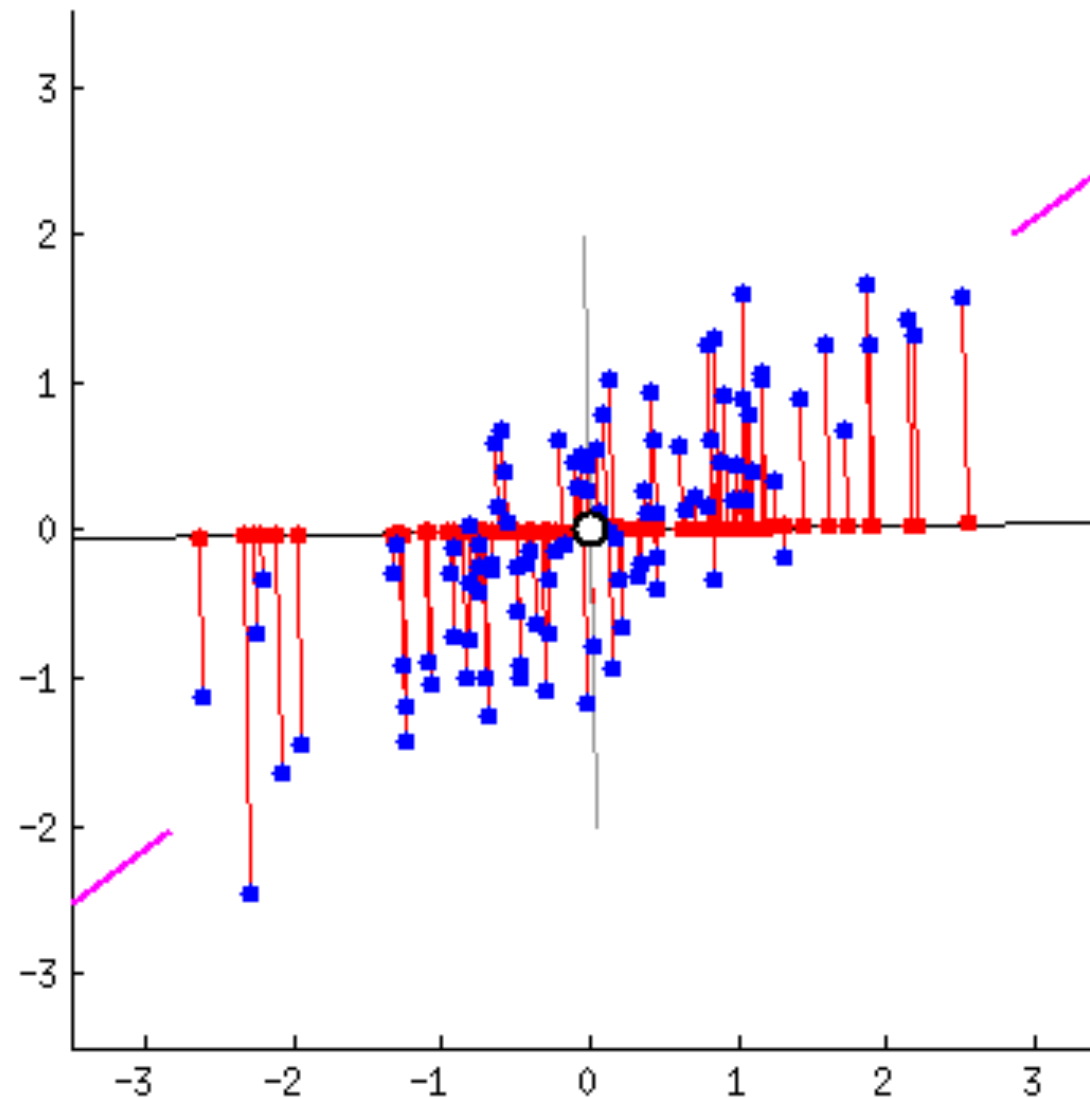
many faces

PCA

Find basis vectors

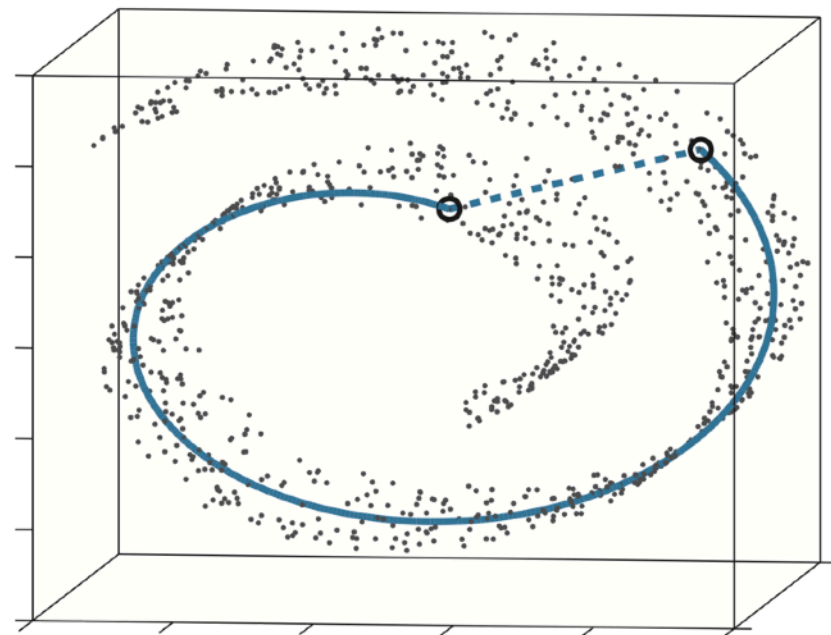
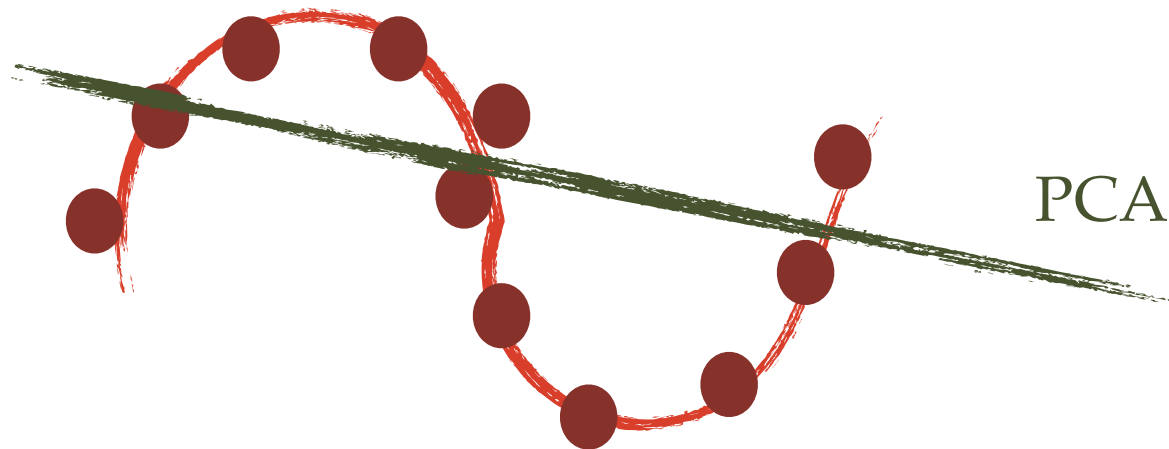
pixels

Feature Projection: Intuition for PCA



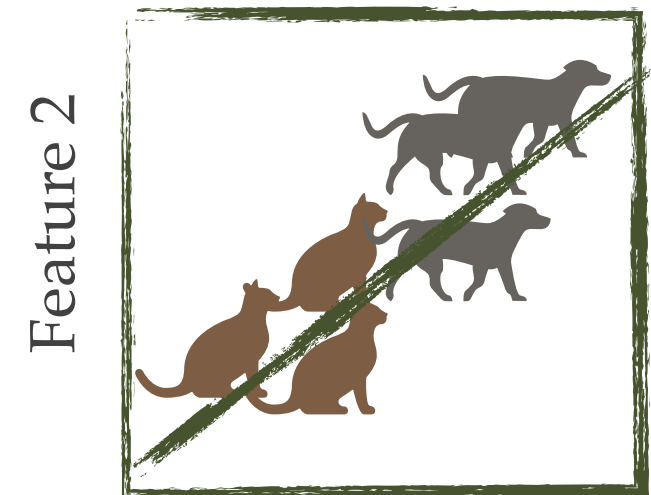
Caveats With PCA

Non-linearity

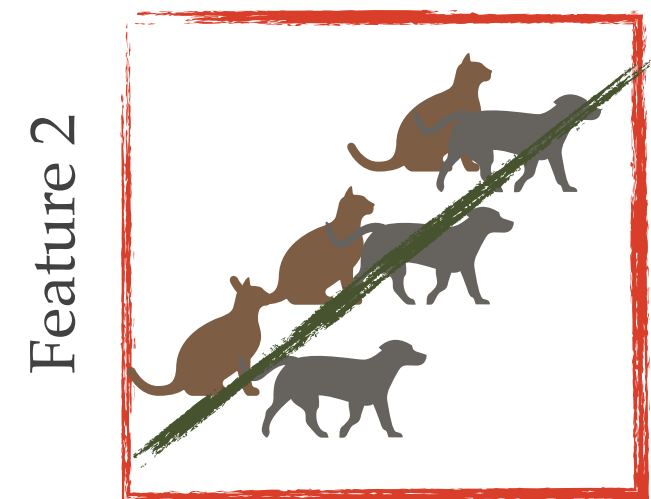


- Data is linearly uncorrelated
- But there is still a non-linear dependence

Higher variance feature is not more discriminative



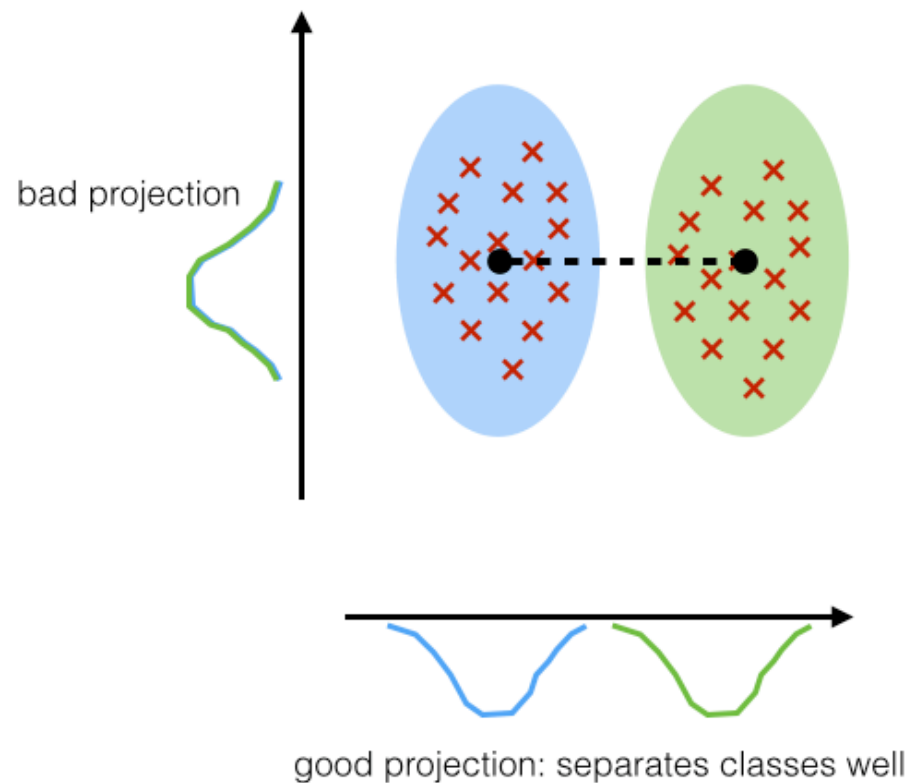
Feature 1



Feature 1

Other Members of the Dimensionality Reduction Zoo

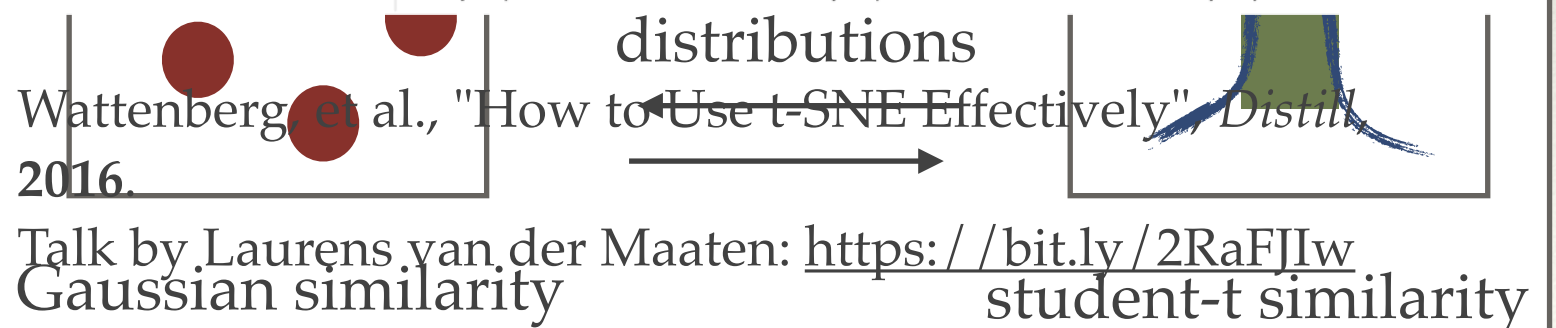
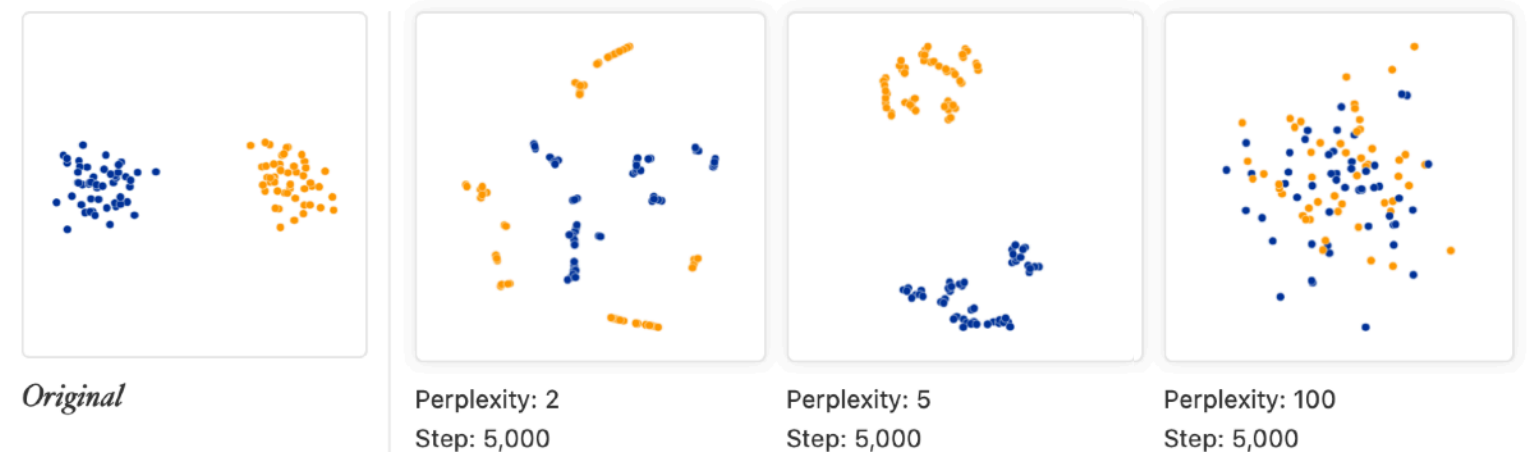
Linear Discriminant Analysis (LDA, Supervised Technique)



- math like for PCA
- maximizing component axes for class separation

t-distributed stochastic neighbor embedding (t-SNE)

- Non-linear
- Conditional probabilities that represent similarities
- Sensitive to perplexity (number of close neighbors)



Feature Projection

Key points:

- Intuition for PCA and feature projection
 - *Uncorrelate data using eigendecomposition of covariance matrix*
 - *Can combine PCs to reconstruct data*
- Issues with PCA and alternatives
 - *Nonlinearity and if variance is not discriminative*
 - *LDA, nonlinear techniques like tSNE*