

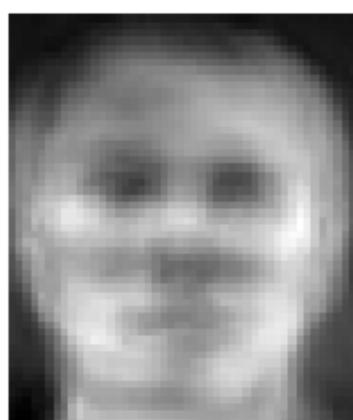
Feature Projection

Key points:

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

Feature Projection: Projecting High-Dimensional Data

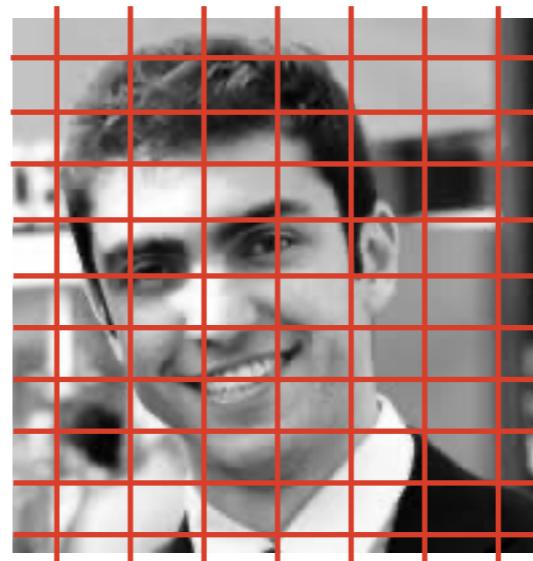
Eigenfaces: Principal Component Analysis on face images to get “basis vectors” of face image space



Linear
combination
of basis
vectors

PCA

Find basis
vectors



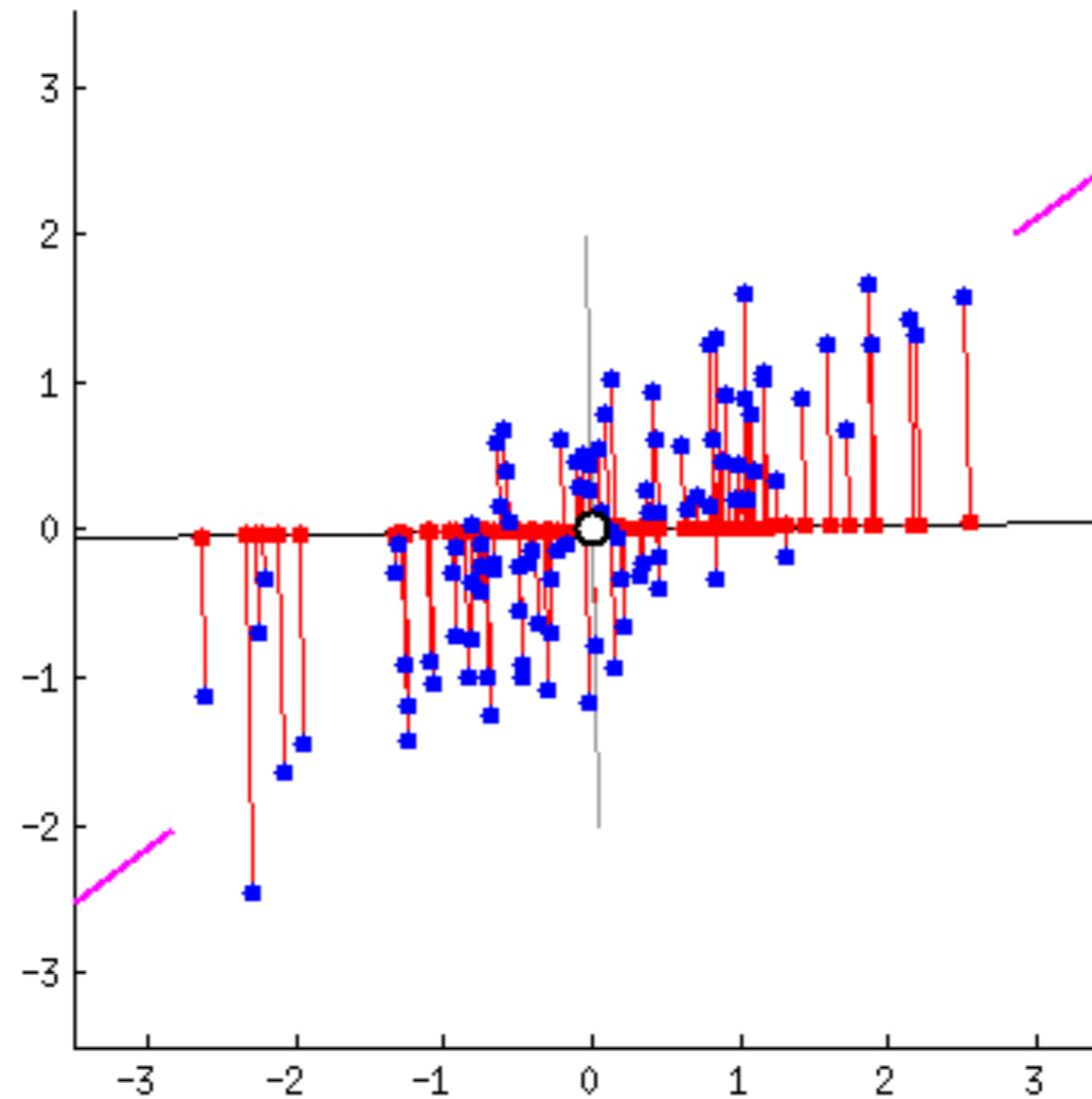
pixels

One image
is one column

many
faces

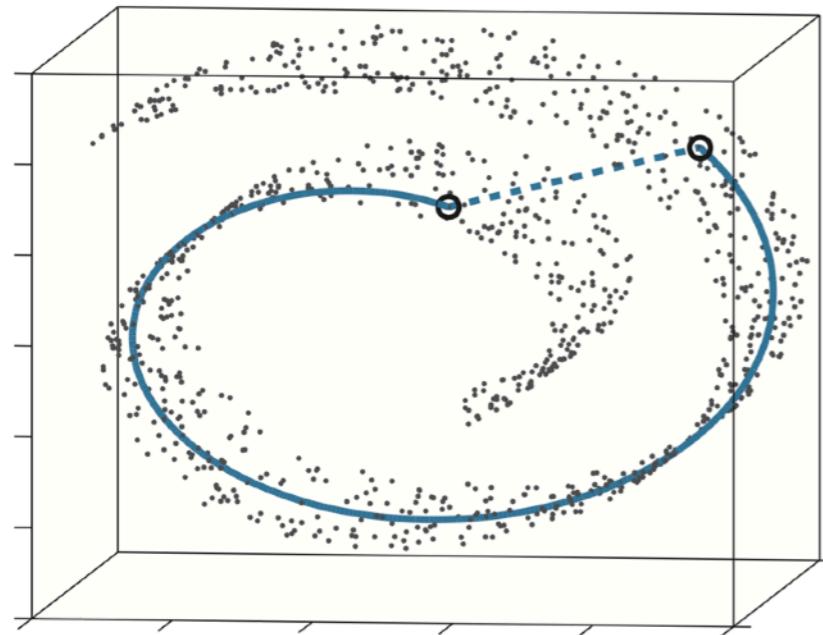
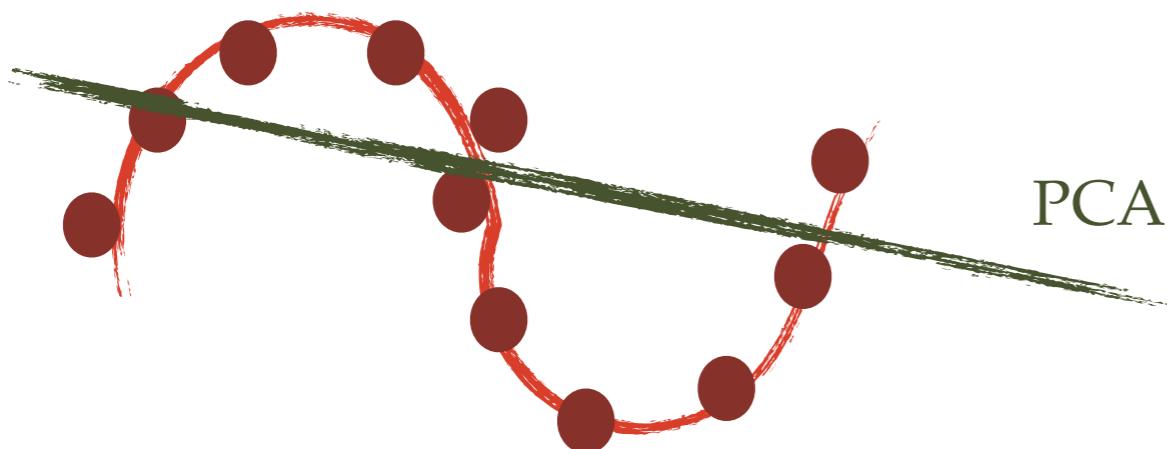


Feature Projection: Intuition for PCA

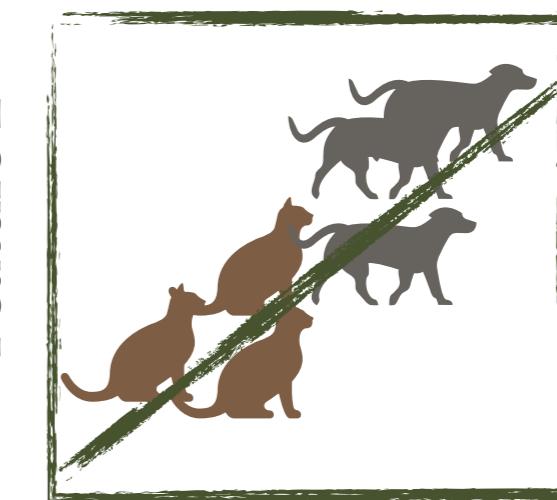


Caveats With PCA

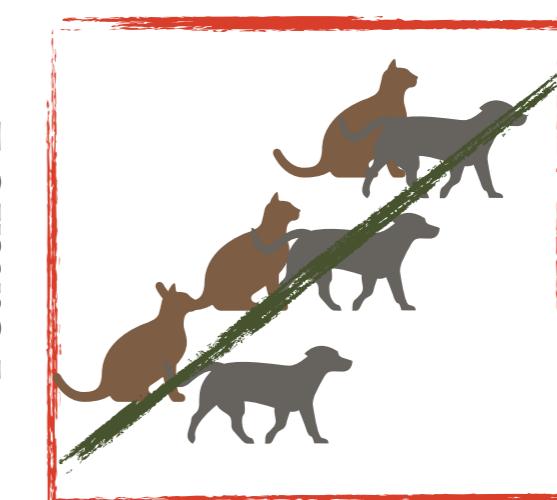
Non-linearity



Higher variance feature is not more discriminative



Feature 1

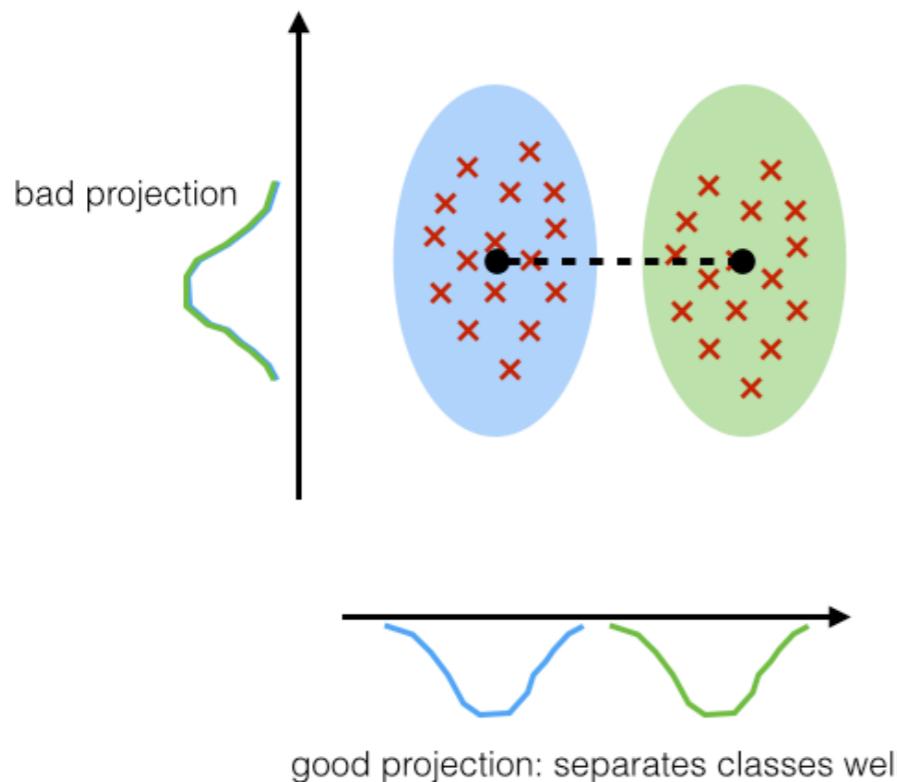


Feature 2

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

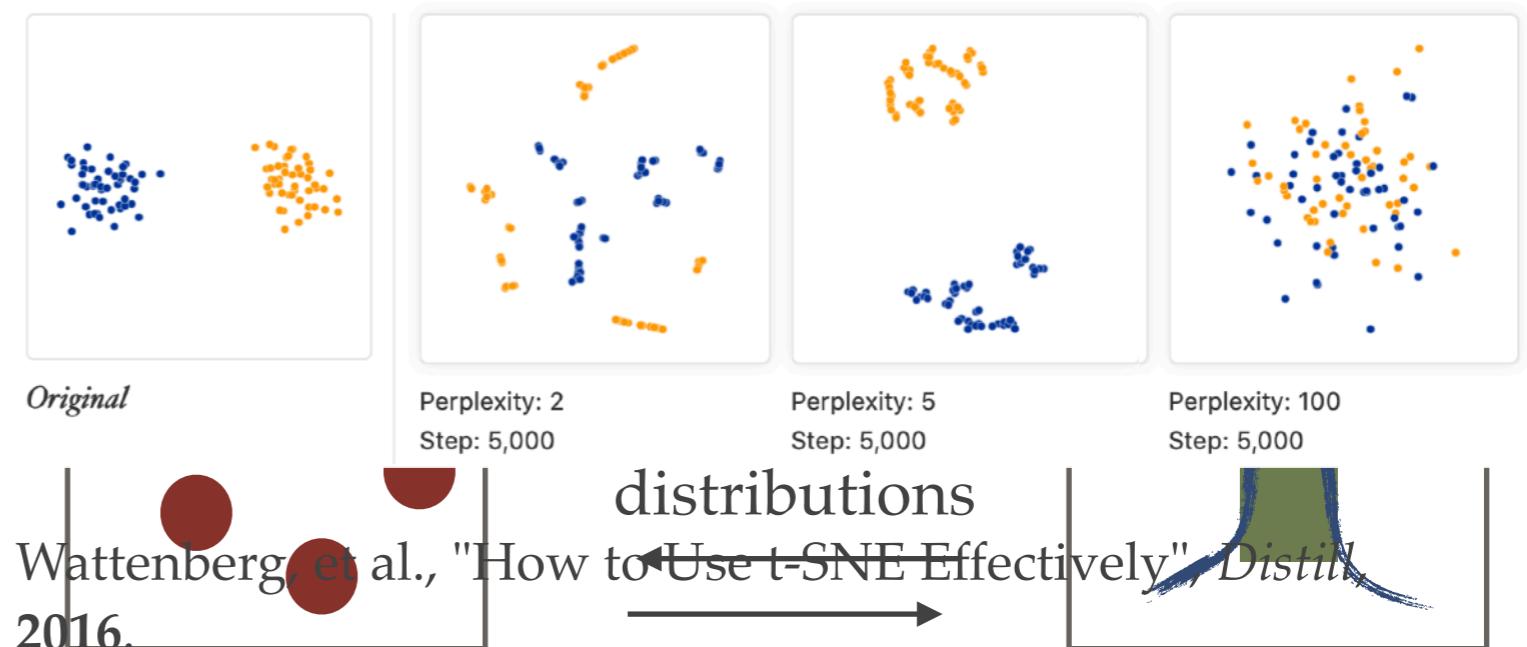
Other Members of the Dimensionality Reduction Zoo

Linear Discriminant Analysis (LDA, Supervised Technique)



t-distributed stochastic neighbor embedding (t-SNE)

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



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

Talk by Laurens van der Maaten: <https://bit.ly/2RaFJIw>
Gaussian similarity student-t similarity

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*