Computational Data Analysis

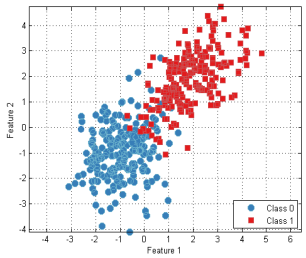
Machine Learning

Yao Xie, Ph.D.
Assistant Professor
Harold R. and Mary Anne Nash Early Career Professor
H. Milton Stewart School of Industrial and Systems Engineering

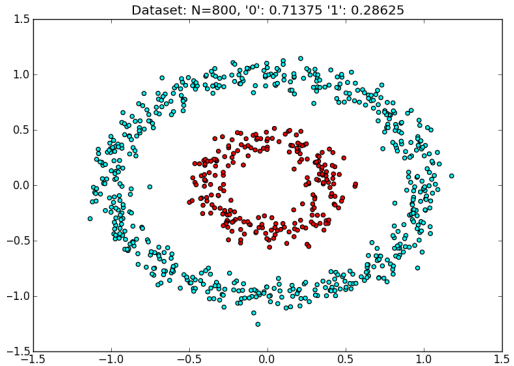
Nonlinear Dimensionality Reduction

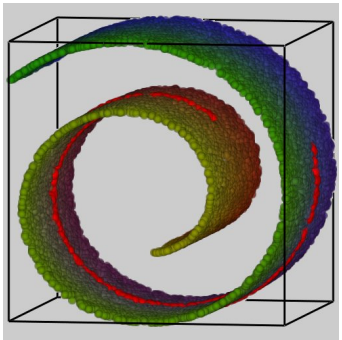
Limitation of PCA and SVD

- Suitable when variables are linearly correlated



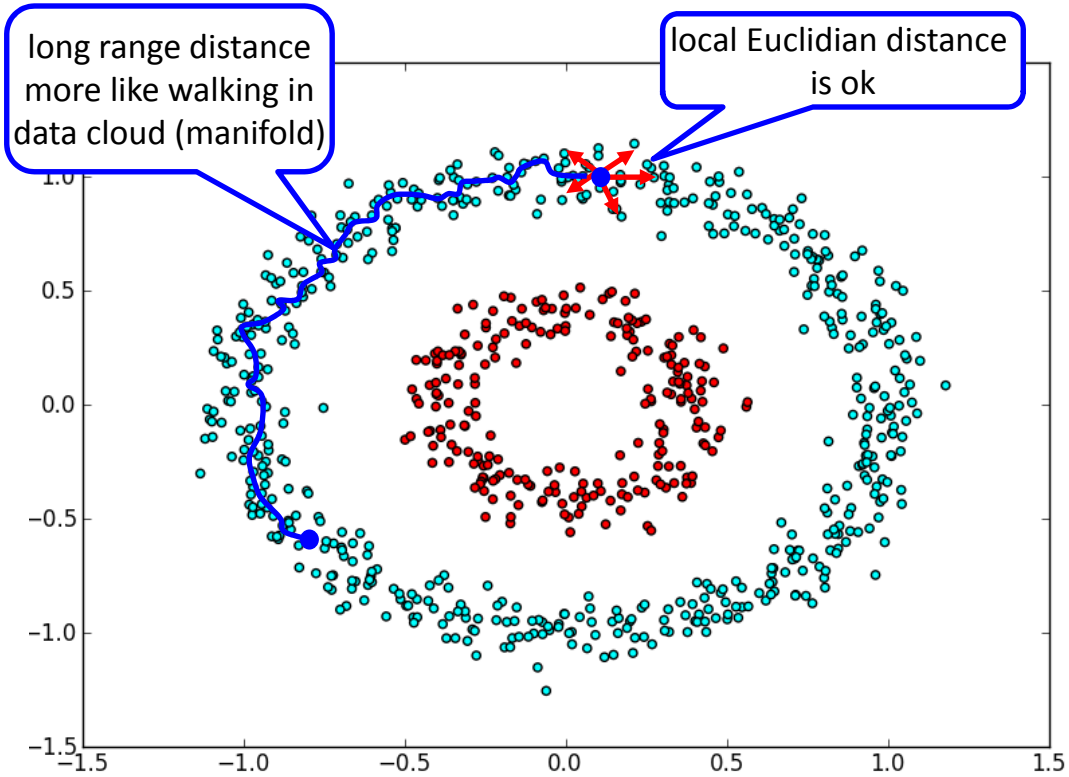
- Not suitable when nonlinear structures are present





http://www.datawrangling.org/python-montage-code-for-displaying-arrays/

What's a reasonable distance measure

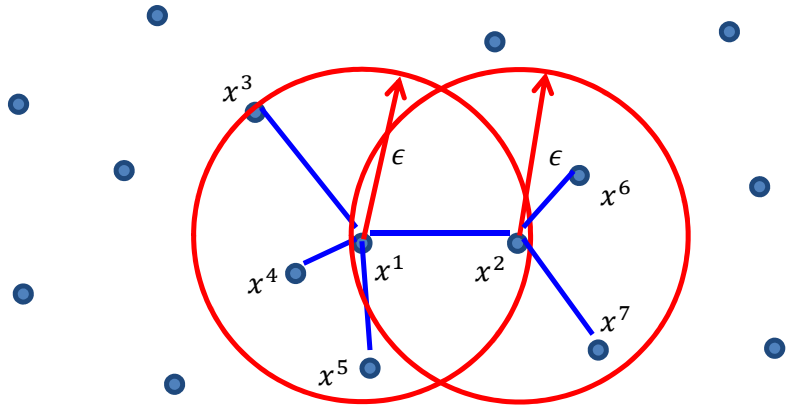


Recall: nearest neighbor graph

(p.23, spectral clustering)

- Given m data points, threshold ϵ, construct matrix $A \in R^{m \times m}$

$$A^{ij} = \begin{cases} 1, \text{if } \left\| x^i - x^j \right\| \leq \epsilon \\ 0, \qquad \text{otherwise} \end{cases}$$



Isomap

- Given m data points, $\{x^1, x^2, \dots x^m\} \in R^n$
- Step 1: build a weighted graph A using nearest neighbors, and compute pairwise shortest distance matrix D
- Step 3: use a centering matrix $H = I - \frac{1}{m} 11^\top$ to get

$$C = -\frac{1}{2m} H (D)^2 H$$

$$D_{ij}^2 := (D_{ij})^2$$

Is the entrywise square of the distance matrix

- Step 4: compute leading eigenvectors $w^1, w^2, \dots$ and eigenvalues $\lambda_1, \lambda_2 \dots$ of C

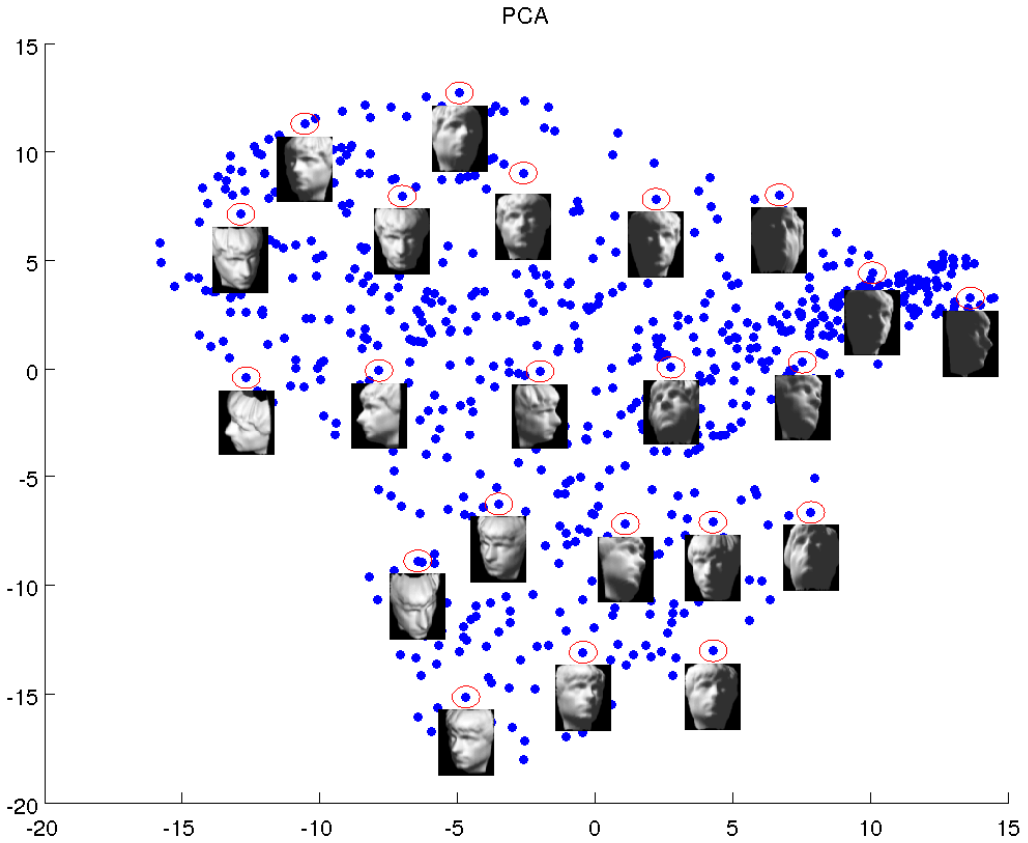
$$Z^T = (w^1, w^2 \dots) \begin{pmatrix} \lambda_1^{1/2} & & \\ & \lambda_2^{1/2} & \\ & & \ddots \end{pmatrix}$$

In Step 3 we are calculating the gram matrix

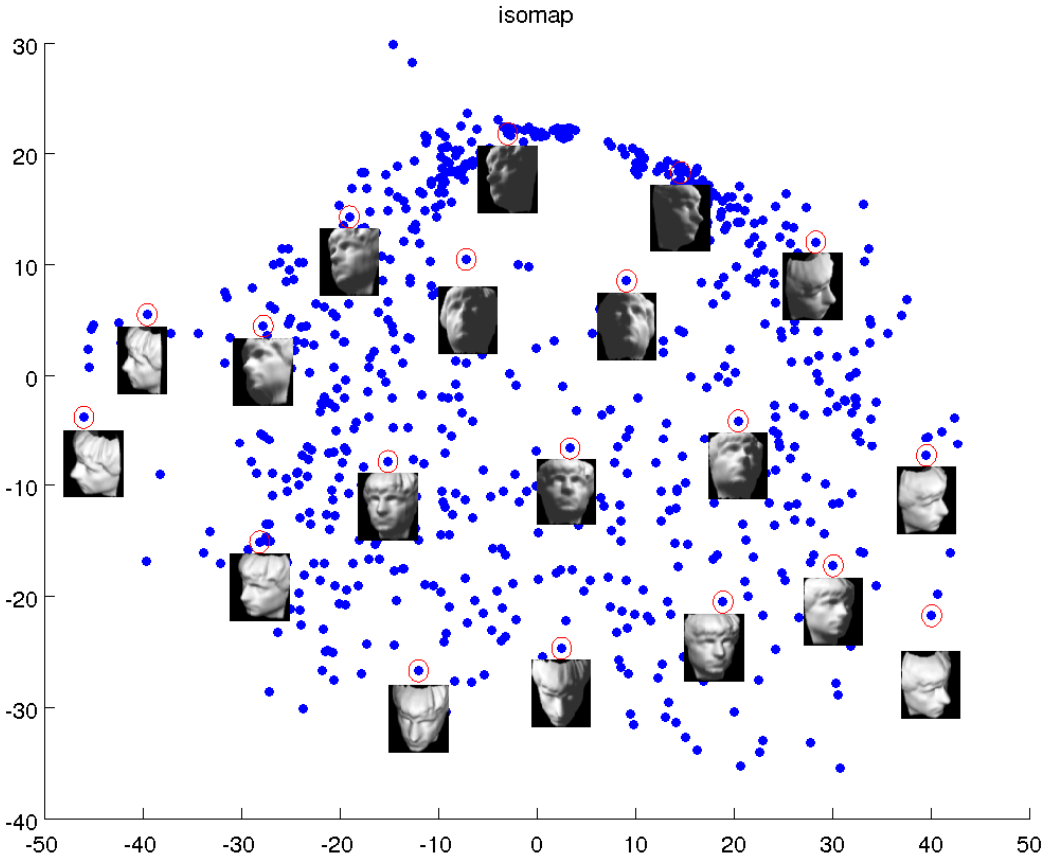
The Gram matrix of the columns of an m×n matrix M is MT M, an n×n matrix.

The rank of M and that of the Gram matrix of its columns are equal. In particular, if the Gram matrix is invertible, then M has full rank, so its columns are linearly independent. This is what the Gram matrix is usually used for.

Is the principal direction interpretable?



Result by isomap




SUMMARY