PA 2018 - 06

Spectral Clustering:

Step 0:

Given a set of points $S = s_1, s_2, \ldots, s_n$ in R^l

Step 1:

Compute affinities $A \in R^{n \times n}$, $A_{ij} = e^{\frac{-||s_i - s_j||_2^2}{2\sigma^2}}$, $A'_{ij} = 0$

Step 2:

Let $D = diag(\sum_j A_{1j}, \sum_j A_{2,j} \ldots \sum_j A_{n,j}$ of the row affinities, and set $L = D^{-\frac{1}{2}} A D^{-\frac{1}{2}}$

Step 3:

Compute eigen decomposition of L.
Let $X = (x_1, x_2 \ldots x_k), X \in R^{n \times k}$ denote a matrix consisting of the k "largest eigenvectors" $x_1, \ldots x_k$ of L

k eigen vectors associated with the largest eigenvalues

Step 4:

Normalize X : $Y_{ij} = \frac{X_{ij}}{(\sum_j X_{ij}^2)^2}$

Step 5:

Perform k-means clustering with k centers on the points that are the rows of Y

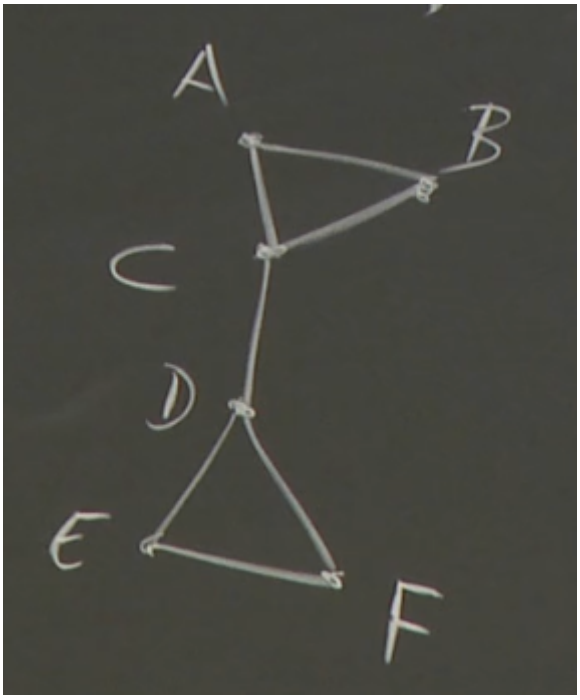
Step 6:

The points in the original space are assigned to the k clusters of their transformed version in %Y%.

Quick, informal note on the Graph Laplacian:

A tool that likes graph theory linear algorithm and probability theory.

Let`s look for example at a graph.



Popular graph algorithm:
Maximum Flow between two nodes?
Shortest Path between two nodes?

The set of mathematical tools that can be used on a graph is somewhat limited,. what exactly should we calculate on a list of nodes an edges ?

$\Rightarrow$ Spectral Graph Theory opens a door to linear algebra by treating a graph`s adjacency matrix as an actual matrix in an algebraic sense $\Rightarrow$ we can for example compute the eigenvalues to find out invariants about he underlying graph.

adjacency matrix: A usefully representation of the adjacency matrix is the Graph Laplacian L.

Remarks on the "Ideal Case" (Sec. 3. 1)

- The set of eigenvalues of a block matrix $\hat{L} = \begin{bmatrix} L^{(11)} & 0 & 0 \\ 0 & L^{(22)} & 0 \\ 0 & 0 & L^{(33)} \end{bmatrix}$ is identical to the set of $eigenvalue(L^{(11)}), eigenvalue(L^{(22)}), eigenvalue(L^{(33)})$(can be treated independently)
- The largest eigenvalue of a graph Laplacian is 1, the seconded largest is lower than 1 $\Rightarrow$ $\hat{L}$ has exactly 3 eigenvalues that are 1. These are the 3 largest ones that we are looking for.
- The associated 3 eigenvector are orthogonal and thus span a 3-dimensional subspaces. Not, however, that the exact choice of eigenvectors is not unique.

The points are clustering around the rotation axes span the subspace These axes are spanned by 90° . angles between them. Therefore. k-means can easily cluster the point in that subspace.

Sec. 3.2 clusters that lie close to another:

The distribution of eigenvalues indicates how well the data structure in to k cluster: consider the eigenvalue gap $\sigma = |\lambda_k - \lambda_{k+1}|$

Cheeger constant:

$$h(S_i) = min_I \frac{\sum_{j \in I, k \notin I} A_{jk}}{min\{\sum_{j \in I} d_j^{(i)}, \sum_{k \notin I} d_k^{(i)}\}}$$

$$\begin{bmatrix} L^{(11)} & 0 & 0 \\ 0 & L^{(22)} & 0 \\ 0 & 0 & L^{(33)} \end{bmatrix}$$