Report on Paper "A tutorial on spectral clustering" by Ulrike von Luxburg

Sourabh Antani

Introduction

This tutorial is setup up as self-contained introduction to spectral clustering. It derives spectral clustering from scratch and present different points of view to why spectral clustering works.

Similarity graphs

Given a set of data points $x_1, \ldots, x_n$ and some notion of similarity $s_{ij} \geq 0$ getween all pairs of data points x_i and x_j, the intuitive goal of clustering is to divide the data points into several groups such that the points in the same group are similar and the points in different groups are dissimilar to each other. A nice way of representing the data is in form of *similarity graph* $G = (V, E)$. Each vertex $v_i$in this graph represents a data point x_i. Two vertices are connected if the similarity s_{ij} between the corresponding data points x_i and x_j is positive or larger than a certain threshold and the edge is weighted by s_{ij}. The problem of clustering can now be reformulated using the similarity graph: we want to find the partition of the graph such that the edges between different groups have very low weights (which points to dissimilarity) and the edges within the group have high weights (hence similar).
Some Terminology:

$G = (V, E)$ undirected graph

$V = \{v_1, \ldots, v_n\}$ vertex set

E set of weighted edges such that each edge between two vertices v_i and v_j carries a non-negative weight $w_{ij} \geq 0$.

Weighted *adjacency matrix* of the graph is matrix $W - (w_{ij})_{i,j=1,\ldots,n}$. If $w_{ij} = 0$, vertices v_i and v_j are not connected.

$w_{ij} = w_{ji}$ since G is undirected.

The degree of a vertex $v_i \in V$ is defined as $d_i = \sum_{j=1}^{n} w_{ij}$.

degree matrix D is defined as the diagonal matrix with the degrees $d_1, \ldots, d_n$ on the diagonal.

$\bar{A}$ Complement $V \backslash A$ of a given subset $A \subset C$.

Indicator vector $\not\Vdash_A$ vector with entries $f_i = 1$ if $v_i = A$, $f_i = 0$ otherwise.

$W(A, B) = \sum_{i \in A, j \in B} w_{ij}$, for non-necessarily disjoint sets.

$|A|$ = number fo vertices in A = size of set A

$vol(A) = \sum_{i \in A} d_i$

Connected subset $A \subset V$ of a graph is connected of any two vertices in A can be joined by a path such that all intermediate points also line in A.

Connected Component Connected subset such that A and $\bar{A}$ are disjoint

Partition non-empty sets $A_1, \ldots, A_k$ form partition of graph inf $A_i \cap A_j = \emptyset$ and $A_1 \cup \cdots \cup A_k = V$

Different similarity graphs

ϵ-neighbourhood graph Connect points whose pairwise distances are less than ϵ. Weight data is not incorportated into the graph.

k-nearest neighbor graphs Connect the nearest k points. Usually directed since v_i being in the k *nearest neighbors* of v_j does not mean that v_j is in k *nearest neighbors* of $v_(i)$

Bibliography

[1] Cheung, Gene, Magli, Enrico, Tanaka, Yuichi, Ng, Michael K. "Graph Spectral Image Processing". Proceedings of the IEEE, vol 106, No. 5, pp. 907-930 May 2018