MVA / ALTEGRAD 2022 / Lab session 4: Graph Mining

Imane Elbacha

29 November, 2022

1 Question 1:

What is the number of edges in graph G consisting of two connected components, where one of the connected components is a complete graph on 100 vertices and the other connected component is a complete bipartite graph with 50 vertices in each partition set?

- for a complete graph: n vertices equate $n(n-1)/2$ hence for 100 vertices there is 4950

- For complete bipartite graph there is 50*50/2= 1250 edges

- In total there are 6200 edges

2 Question 2:

What is the maximal possible global clustering coefficient? On what type of graph is this clustering coefficient achieved?

The clustering coefficient is a real number between zero and one that is zero when there is no clustering, and one for maximal clustering, which happens when the graph consists of disjoint cliques

In graph theory, a clique is a subset of undirected graph's vertices where each pair of different vertices is next to the other. In other words, a clique of a graph G is a full induced subgraph of G.

3 Question 3

Consider a connected graph, i.e., a graph consisting of a single connected component. What is the eigenvector corresponding to the smallest eigenvalue of Lrw?

4 Question 4

Given a graph G is the output of the spectral clustering deterministic or stochastic?

In this algorithm of spectral clustering we use K-means and this clustering algorithm is non-deterministic. The random selection of data points used to create the initial centroids of K-Means accounts for its non-deterministic nature.

5 Question 5:

Compute (showing your calculations) the modularity of the clustering results shown in Figure 1. Note that different colors correspond to different clusters.

For graph (a)

- for a: nc=2 m=9 lc=[5,4] for the degrees=[2,1,2,2,1,2,2,2,2] dc=[8,8]

- for cluster 1: $m_1 = 5/9 - (8/2*9)^2 = 0.35$

- for cluster 2: $m_2 = 4/9 - (8/2*9)^2 = 0.25$

- modality=0.6

For graph (b)

- for a: nc=2 m=9 lc=[4,5] for the degrees=[2,1,2,2,1,2,2,2,2] dc=[7,9]

- for cluster 1: $m_1 = 4/9 - (7/2 * 9)^2 = 0.29$

- for cluster 2: $m_2 = 5/9 - (9/2 * 9)^2 = 0.30$

- modality=0.59