Yassine Filali
yassine.filali@ensta-paris.fr

Lab session # LAB 4
ALTEGRAD 2020

12/15/20

1 Question 1

In the first connected component, we have a complete graph with 100 vertices, therefore the number of edges is $\frac{100*(100-1)}{2} = 50 * 99 = 4950$ edge. In the second connected component, we have a complete bipartite graph with 50 vertex in each partition set, therefore the number of edges is $50 * 50 = 2500$ edge. In total, the graph G has $4950 + 2500 = 7450$ edge.
The total number of triangles is exactly the number of triangles in the complete graph, therefore it is equal to the number combinations of 3 vertices out of 100. Therefore he total number of triangles in G is $\begin{pmatrix} n \\ 3 \end{pmatrix} = \frac{n!}{3!(n-3)!} = \frac{n*(n-1)*(n-2)}{6}$. In our case, It is equal to 161700

2 Question 2

The maximum possible value of the global clustering coefficient is 1. It is achieved in the case of a complete graph, where all the triplets are triangles and there are no open triplets.

3 Question 3

Let us consider $x = (1, 1, .., 1)^T$, we have that:

$$L_{rw}x = x - D^{-1}Ax = x - D^{-1}diag(D) = x - x = 0$$

Therefore x is the eigenvctor assocoiated to eigenvalue 0. Since the laplacian matrix is positive semi-definite, 0 is the minimum eigenvalue and x the minimum eigenvector. Since this vector offers no additional information, it can be safely ignored in the clustering process.

4 Question 4

The nature of the spectral clustering is deterministic since it always gives the same results after each execution.

5 Question 5

For the communities in (a), we have :

- $m = 13$
- $l_{green} = l_{blue} = 6$
- $d_{green} = d_{blue} = 13$
- $Q = \frac{l_{green}}{m} - (\frac{d_{green}}{2m})^2 + \frac{l_{blue}}{m} - (\frac{d_{blue}}{2m})^2 = 0.423$

For the communities in (b), we have :

- $m = 13$
- $l_{green} = 2$, $l_{blue} = 4$
- $d_{green} = 11$,$d_{blue} = 15$
- $Q = \frac{l_{green}}{m} - (\frac{d_{green}}{2m})^2 + \frac{l_{blue}}{m} - (\frac{d_{blue}}{2m})^2 = -0.0502$

6 Question 6

$$\Phi(C_4) = [4, 2, 0, 0, ..., 0]$$

$$\Phi(P_4) = [3, 2, 1, 0, ..., 0]$$

For the pair (C_4, C_4), $K(C_4, C_4) =< \Phi(C_4), \Phi(C_4) >= 16 + 4 = 20$.
For the pair (C_4, P_4), $K(C_4, P_4) =< \Phi(P_4), \Phi(C_4) >= 12 + 4 = 16$.
For the pair (P_4, P_4), $K(P_4, P_4) =< \Phi(P_4), \Phi(P_4) >= 9 + 4 + 1 = 14$.