Benjamin Lapostolle
benjamin.jean.lapostolle@gmail.com

Lab session # 2
ALTEGRAD 2023

10/24/23

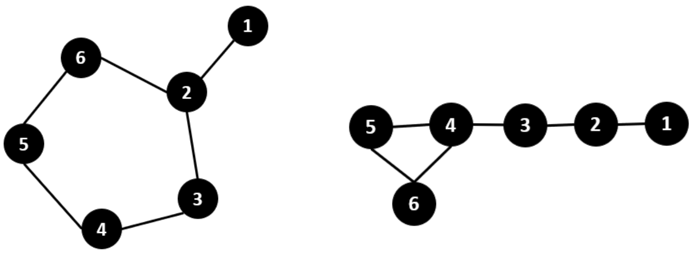
1 Question 1

A graph of n nodes has a maximum number of edges and triangle if all nodes are connected to each other. In this case:

- Maximum number of edges: $\frac{n(n-1)}{2}$, because it is the number of ways to pick 2 vertices among the n, which is equal to $\binom{n}{2}$.
- Maximum number of triangles: $\frac{n(n-1)(n-2)}{6}$, because it is the number of ways to pick 3 vertices among the n, which is equal to $\binom{n}{3}$.

2 Question 2

This statement if false and we give an example in the figure 1 to show it. Those two graphs of 6 points have the same distribution, however we can show that there is no bijective mapping between the two graphs such that the condition of isomorphism is respected. In fact, the node 1 in 1b must be associated to the node 1 in 1a. And we select a different node, then because the node 1 has only one connection and that the other nodes have at least two connections, the other selected one will have one missing connection. By the same argument, we can show that the node 6 in 1b must be associated to the node 2 in 1a (otherwise, one connection among the 3 will not be present). However, node 1 and node 6 which were not connected are now connected and because the number of edges among the two graphs is the same, this will result in a missing connection no matter how we associate the remaining nodes. Hence, those two graphs cannot be isomorphic.



(a) degree distribution: $d_1 = 1, d_2 = 4, d_3 = 1$

(b) degree distribution: $d_1 = 1, d_2 = 4, d_3 = 1$

Figure 1: Two graphs of 6 points with the same degree distribution

3 Question 3

In a cycle graph of n nodes, we can clearly see that there is no closed triangles when $n >= 4$, hence the clustering coefficient will be equal to 0. For a cycle graph of 3 nodes, we have one closed triangle hence the clustering coefficient will be equal to 1.

4 Question 4

To answer this question, we can use the results of the proposition 3 in this paper[1] stating properties of laplacians.

Let u be the eigenvector associated to the smallest eigenvalues of the normalized Laplacans. According to the proposition, L_{rw} is positive semi-definite and 0 is an eigenvalue of this matrix. Hence, u is associated with the eigenvalue 0. Moreover, still according to the proposition, if u is an eigenvector of L_{rw}, then $w = D^{1/2}u$ is an eigenvector of L_{sym} associated to the same eigenvalue.
Moreover, we have :

$$\forall f \in \mathbb{R}^n, f^T L_{sym} f = \frac{1}{2} \sum_{i,j=1}^{n} A_{ij} \left(\frac{f_i}{\sqrt{d_i}} - \frac{f_j}{\sqrt{d_j}} \right)^2$$

In particular, for $f = D^{1/2}u$, and because $L_{sym}f = 0$, we have:

$$\frac{1}{2} \sum_{i,j=1}^{n} A_{ij} \left(\frac{u_i\sqrt{d_i}}{\sqrt{d_i}} - \frac{u_j\sqrt{d_j}}{\sqrt{d_j}} \right)^2 = \frac{1}{2} \sum_{i,j=1}^{n} A_{ij} \left(u_i - u_j\right)^2 = 0$$

5 Question 5

We index the parameters x by x_b and x_r, corresponding to the blue and red communities. The modularity is given by:

$$Q = \sum_{i=0}^{n_c} (\frac{l_c}{m} - (\frac{d_c}{2m})^2)$$

For the first graph, we have:

- $m = 14$
- $l_r = 6, d_r = 14$
- $l_b = 6, d_b = 14$

Hence, we have $Q_1 = \frac{5}{14}$

For the second graph, we have:

- $m = 14$
- $l_r = 2, d_r = 11$
- $l_b = 5, d_b = 17$

Hence, we have $Q_2 = -\frac{9}{392}$, which is closer to 0 indicating that this clustering is worse in the sense of the defined modularity than the first clustering.

6 Question 6

To compute the shortest path kernel for the pairs (P_4, P_4), (P_4, S_4) and (S_4, S_4), we need to compute the representation of P_4 and S_4 in the shortest path space.
We have:

- $S_4 \rightarrow [3, 3, 0]$
- $P_4 \rightarrow [3, 2, 1]$

So, when computing the dot product we obtain:

- $K(S_4, S_4) = 9 + 9 = 18$
- $K(S_4, P_4) = 9 + 6 = 15$
- $K(P_4, P_4) = 9 + 4 + 1 = 14$

7 Question 7

If the graphlet kernel of two graphs is 0, because the coefficients of the representations in the graphlet space of the graphs are positive, it indicates that if a coefficient of a graph is strictly positive, then the other corresponding coefficient must be 0. From this observation, we conclude that if the graphlet kernel of two graphs is 0, the two graphs must not share subgraphs isomorphic to the same graphlet of size $k = 3$. Two examples are presented in the figure 2. By taking the basic garphlets of size 3, we see that the dot product is O.

References

[1] Ulrike Von Luxburg. A tutorial on spectral clustering. *Statistics and Computing*, 17(4):395–416, 2007.

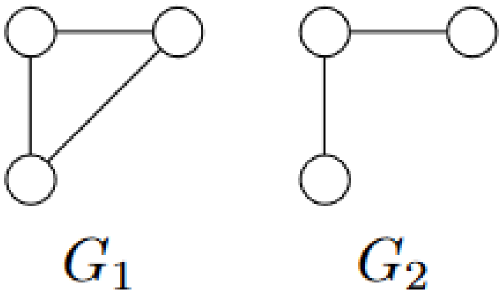


Figure 2: Two graphs whose graphlet kernel is equal to 0