YOU Borachhun
borachhunyou@gmail.com

1 Question 1

An undirected graph of n nodes without self-loops can have:

- The maximum number of edges of $\binom{n}{2} = \frac{n(n-1)}{2}$
- The maximum number of triangles of $\binom{n}{3} = \frac{n(n-1)(n-2)}{6}$.

2 Question 2

Two graphs with identical degree distributions are not necessarily isomorphic.

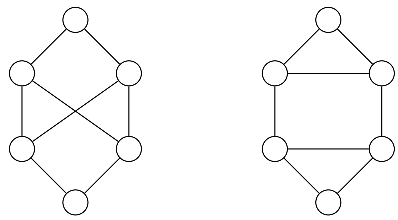

Figure 1: Example of graphs with the same degree distribution that are not isomorphic

As an example, in figure 1, both graphs have two nodes of degree 2 and four nodes of degree 3. However, the right graph has cycles of length 3 while the left one does not.

3 Question 3

$$\text{global clustering coefficient} = \frac{\text{\# closed triplets}}{\text{\# closed triplets} + \text{\# open triplets}}$$

- For C_3, i.e. cycle graph on 3 vertices, the number of closed triplets equals 1 and the number of open triplets equals 0. Thus, the global clustering coefficient equals 1.
- For $C_n, n \geqslant 4$, since there are no three nodes that are connected by 3 edges, the number of closed triplets equals 0. Thus, the global clustering coefficient equals 0.

4 Question 4

Based on [1], the smallest eigenvalue of $\mathbf{L_{rw}}$ is 0 and the associated eigenvector $\mathbf{u}_1 = (1, 1, ..., 1)^T$. Therefore, $\sum_{i=1}^{n} \sum_{j=1}^{n} \mathbf{A}_{ij}([\mathbf{u}_1]_i - [\mathbf{u}_1]_j)^2 = 0$.

5 Question 5

Left graph:

- Total number of edges: $m = 14$
- Number of clusters: $n_c = 2$
- Number of edges in each cluster: $l_1 = 6, l_2 = 6$
- Sum of degrees of nodes belong to each cluster: $d_1 = 14, d_2 = 14$
- Modularity: $Q = \left[\frac{6}{14} - \left(\frac{14}{2\times 14}\right)^2\right] + \left[\frac{6}{14} - \left(\frac{14}{2\times 14}\right)^2\right] = 0.357$

Right graph:

- Total number of edges: $m = 14$
- Number of clusters: $n_c = 2$
- Number of edges in each cluster: $l_1 = 5, l_2 = 2$
- Sum of degrees of nodes belong to each cluster: $d_1 = 17, d_2 = 11$
- Modularity: $Q = \left[\frac{5}{14} - \left(\frac{17}{2\times14}\right)^2\right] + \left[\frac{2}{14} - \left(\frac{11}{2\times14}\right)^2\right] = -0.0229$

6 Question 6

Path graph on 4 vertices P_4 contains 3, 2, 1 shortest paths of distance 1, 2, 3 respectively. Star graph on 4 vertices S_4 contains 3 shortest paths of both distance 1 and 2. Then we have $\phi(P_4) = [3, 2, 1]$ and $\phi(S_4) = [3, 3, 0]$. Therefore, the shortest path kernel:

$$k(P_4, P_4) = \langle\phi(P_4), \phi(P_4)\rangle = 14$$

$$k(P_4, S_4) = \langle\phi(P_4), \phi(S_4)\rangle = 15$$
$$k(S_4, S_4) = \langle\phi(S_4), \phi(S_4)\rangle = 18$$

7 Question 7

A kernel value equal to 0 means that there is no similarity between the two graphs that are compared.

Figure 2: Graph G (left) and G' (right) with graphlet kernel equals zero

Let's consider the graphlet kernel with graphlets of size 3 (i.e. induced subgraphs with 3 nodes). Figure 2 shows an example of two graphs, G and G', for which the graphlet kernel $k(G, G') = 0$ holds. G only contains a graphlet with 3 edges while G' contains graphlets with 1 or 2 edges. Both graphs do not contain any graphlet with 0 edge. By using the notation in the lab paper, we have:

$$k(G, G') = f_G^T f_{G'} = \begin{pmatrix} a \\ 0 \\ 0 \\ 0 \end{pmatrix}^T \times \begin{pmatrix} 0 \\ b \\ c \\ 0 \end{pmatrix} = 0$$

where $a, b, c > 0$ (depends on the number of random samples of graphlets).

References

[1] Ulrike von Luxburg. A tutorial on spectral clustering. *Statistics and computing*, pages 5–6, 2007.