Kevin WARDAKHAN
kevin.wardakhan@icloud.com

Lab session # 4
ALTEGRAD 2025

11/18/25

1 Question 1

I have created both graph, and are presented in the following.

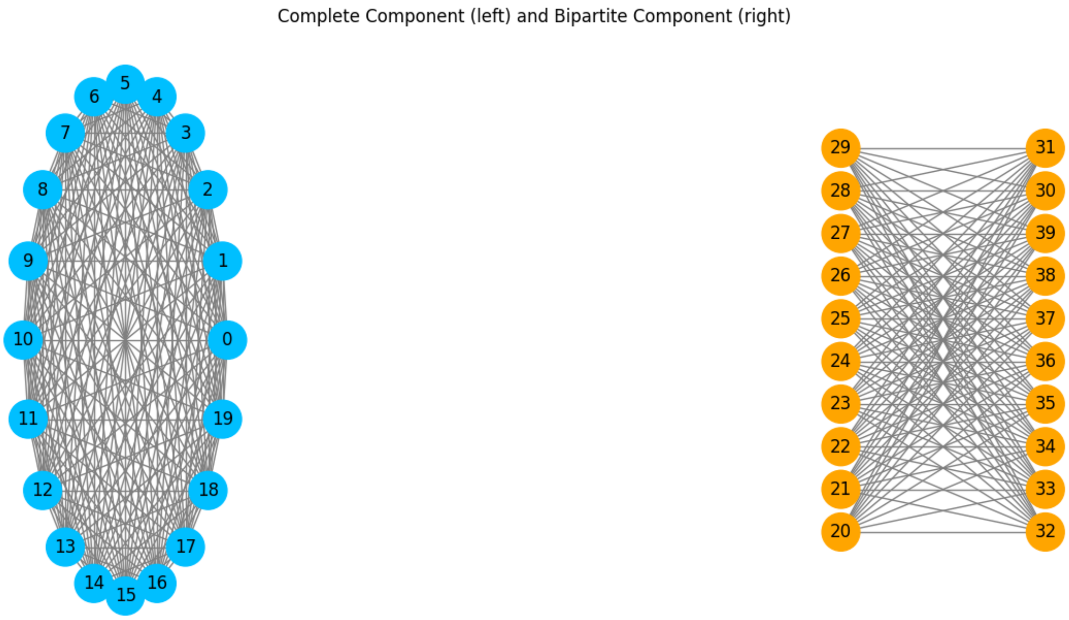

Figure 1: Representation of the graph G.

To count the number of triangles in $\bar{G}$, we first describe its structure. Let A be the vertex set of the original K_{20}, and let B and C be the two parts of size 10 of the original complete bipartite graph $K_{10,10}$. In $\bar{G}$ we have:

- A is an independent set of 20 vertices,
- B and C each induce a complete graph K_{10},
- there are no edges between B and C,
- every vertex of A is adjacent to every vertex of B and to every vertex of C.

We now count the triangles in $\bar{G}$.
There are four types of triangles to consider.
First, we look at triangles entirely contained in B. Since $B = K_{10}$, any choice of three vertices in B forms a triangle, so the number of such triangles is

$$\binom{10}{3}.$$

By the same argument, triangles entirely contained in C are also counted by

$$\binom{10}{3}.$$

Next, we consider triangles with one vertex in A and two vertices in B. We first choose the vertex in A and then choose an unordered pair of vertices in B, which gives

$$\binom{20}{1}\binom{10}{2}$$

triangles.

Finally, triangles with one vertex in A and two vertices in C are counted in exactly the same way, giving

$$\binom{20}{1}\binom{10}{2}$$

additional triangles.
There are no triangles involving vertices from both B and C, because there are no edges between B and C, and there are no triangles with at least two vertices in A, because A is independent.
Adding all contributions, the total number of triangles in $\bar{G}$ is

$$2\binom{10}{3} + 2\binom{20}{1}\binom{10}{2} = 2 \cdot 120 + 2 \cdot 20 \cdot 45 = 240 + 1800 = 2040.$$

Therefore, the number of triangles in $\bar{G}$ is $\boxed{2040}$.

2 Question 2

We are interested in the stationary points of $R(A, \cdot)$, that is, the vectors x such that

$$\nabla_x R(A, x) = 0$$

Let's first compute the gradient of the Rayleigh quotient.
Set

$$f(x) = x^\top Ax, \qquad g(x) = x^\top x.$$

Then $R(A, x) = f(x)/g(x)$. Since A is symmetric (by definition of an adjacency matrix),

$$\nabla_x f(x) = (A + A^\top)x = 2Ax, \qquad \nabla_x g(x) = 2x.$$

By the quotient rule,

$$\nabla_x R(A, x) = \frac{g(x)\, \nabla_x f(x) - f(x)\, \nabla_x g(x)}{\big(g(x)\big)^2}.$$

Which finally gives

$$\nabla_x R(A, x) = \frac{(x^\top x)\, 2Ax - (x^\top Ax)\, 2x}{(x^\top x)^2} = \frac{2}{(x^\top x)^2}\left[(x^\top x)Ax - (x^\top Ax)x\right].$$

Suppose $x \neq 0$ is a stationnary point of $R(A, \cdot)$.
We have $x^\top x > 0$, so the denominator in the previous expression is nonzero. Hence

$$(x^\top x)Ax - (x^\top Ax)x = 0,$$

or equivalently

$$Ax = \frac{x^\top Ax}{x^\top x}\, x.$$

Define

$$\lambda = \frac{x^\top Ax}{x^\top x} = R(A, x).$$

Then the stationary condition becomes

$$Ax = \lambda x,$$

Thus, x is the eigen vector of $R(A, \cdot)$ associated with the eigenvalue λ.
Conversely, if we suppose that x is the eigen vector of $R(A, \cdot)$ associated with the eigenvalue λ, we get :

$$R(A, x) = \frac{x^\top Ax}{x^\top x} = \lambda$$

Differentiating a constant thus gives

$$\nabla_x R(A, x) = 0$$

We have shown that x is a stationary point of the Rayleigh quotient $R(A, \cdot)$ if and only if x is an eigenvector of A, and in that case the associated eigenvalue is given by

$$\lambda = R(A, x) = \frac{x^\top Ax}{x^\top x}.$$

3 Question 3

Both figures correspond to the same graph, but with different clusterings. In each case there are two clusters; we denote $c = 1$ for the blue cluster and $c = 2$ for the orange one. We will compute, for each figure, the modularity of the given partition and then compare the values. Visually, the clusters in Figure a seem more natural than those in Figure b, so we expect the modularity to be larger for Figure a.
Recall the modularity of a partition is

$$Q = \sum_{c=1}^{n_c} \left(\frac{l_c}{m} - \left(\frac{d_c}{2m} \right)^2 \right),$$

Computation of the modularity for Figure a.

Reading the graph and the blue/orange partition of Figure a, we obtain

$$m = 13, \qquad n_c = 2, \qquad l_1 = 7,\ l_2 = 5, \qquad d_1 = 15,\ d_2 = 11.$$

Hence

$$\begin{aligned} Q_a &= \left(\frac{l_1}{m} - \left(\frac{d_1}{2m} \right)^2 \right) + \left(\frac{l_2}{m} - \left(\frac{d_2}{2m} \right)^2 \right) \\ &= \left(\frac{7}{13} - \left(\frac{15}{26} \right)^2 \right) + \left(\frac{5}{13} - \left(\frac{11}{26} \right)^2 \right) \\ &= \frac{139}{338} \approx \boxed{0.411}. \end{aligned}$$

Computation of the modularity for Figure b.

For the clustering shown in Figure b we read

$$m = 13, \qquad n_c = 2, \qquad l_1 = 2,\ l_2 = 7, \qquad d_1 = 8,\ d_2 = 18.$$

Therefore

$$\begin{aligned} Q_b &= \left(\frac{l_1}{m} - \left(\frac{d_1}{2m} \right)^2 \right) + \left(\frac{l_2}{m} - \left(\frac{d_2}{2m} \right)^2 \right) \\ &= \left(\frac{2}{13} - \left(\frac{8}{26} \right)^2 \right) + \left(\frac{7}{13} - \left(\frac{18}{26} \right)^2 \right) \\ &= \frac{20}{169} \approx \boxed{0.118}. \end{aligned}$$

We have $Q_a \approx 0.411$ and $Q_b \approx 0.118$, so the partition of Figure a has a much higher modularity. In Figure a most edges lie inside the two colored groups and only one edge links the two communities, which means that the partition captures well the community structure of the graph. In Figure b more edges are cut by the partition, so the corresponding modularity is lower and the clustering is less coherent, which was expected.

4 Question 4

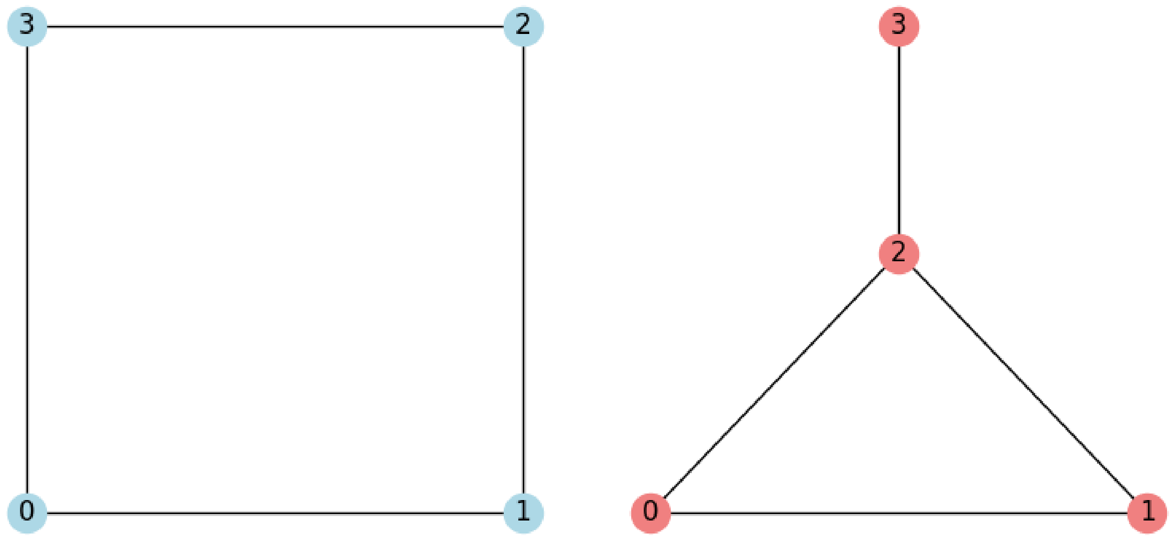

Figure 2: Representation of graph G (square) and G' (triangle with extra edge).

These two graphs are non-isomorphic because, for example, node 2 has degree 2 in graph G but degree 3 in graph G'. Nonetheless, they have the same shortest path kernel representation:

$$\phi(G) = \phi(G') = [4, 2, 0, ..., 0]$$

5 Question 5

Let us perform one full iteration of the Weisfeiler–Lehman subtree kernel algorithm for both graphs G and G'.

Iteration number 0

By simply counting the number of occurences of each labels in each graph, we can define ϕ_0 : graph $\rightarrow \mathbf{N}^n$ where n is the number of distinct label in the graph. This function returns a vector containing at index i the number of nodes in the Graph having for label the i-th distinct label. We obtain after computation :

$$\phi_0(G) = (2, 1, 1, 1, 1) \quad \text{and } \phi_0(G') = (1, 1, 1, 2, 1)$$

So, the kernel for iteration 0 is :

$$k_0(G, G') := \langle \phi_0(G), \phi_0(G') \rangle = 7$$

Iteration number 1

We have to relabel the nodes of each graph, depending on their current label and the labels of their neighbors. We get:

- For G :

Vertex	$\ell^{(0)}$	$M^{(0)}$ (sorted neighbor labels)	$s^{(0)} = (\ell^{(0)}, M^{(0)})$	$\ell^{(1)}$
v_1	2	$[1, 3, 4, 5]$	$(2, [1, 3, 4, 5])$	a
v_2	1	$[2, 5]$	$(1, [2, 5])$	b
v_3	5	$[1, 2, 3]$	$(5, [1, 2, 3])$	c
v_4	3	$[1, 2, 4, 5]$	$(3, [1, 2, 4, 5])$	d
v_5	4	$[1, 2, 3]$	$(4, [1, 2, 3])$	e
v_6	1	$[3, 4]$	$(1, [3, 4])$	f

Table 1: One WL iteration for graph G.

Vertex	$\ell^{(0)}$	$M^{(0)}$ (sorted neighbor labels)	$s^{(0)} = (\ell^{(0)}, M^{(0)})$	$\ell^{(1)}$
u_1	2	$[1,4,5]$	$(2,[1,4,5])$	g
u_2	1	$[2,5]$	$(1,[2,5])$	b
u_3	5	$[1,2,4]$	$(5,[1,2,4])$	h
u_4	3	$[4,4]$	$(3,[4,4])$	i
u_5	4	$[3,5]$	$(4,[3,5])$	j
u_6	4	$[2,3]$	$(4,[2,3])$	k

Table 2: One WL iteration for graph G'. The tuple $(1,[2,5])$ receives the same new label b as in G.

- For G' :

Thus,

$$\phi_1(G) = \phi_1(G') = (1,1,1,1,1)$$

Which finally gives :

$$\boxed{k(G,G') := k_0(G,G') + k_1(G,G') = 7 + 1 = 8}$$

We see that at iteration 0 the kernel has a relatively large value of 7. This reflects that the two graphs are quite similar at the level of their original node labels: they share the same set of labels and several of them appear with comparable frequencies. After one WL refinement (iteration 1), only one refined label is common to both graphs, so the contribution to the kernel is just 1. This shows that once local neighborhood structure is taken into account, the two graphs are actually quite different.