1 Question 1

Let two nodes n_1, n_2 in the same component. Since there are no other nodes in the component, a random walk is always an alternation of n_1 and n_2. Therefore random walks spanning from n_1 and n_2 are very similar and their embeddings should be equal. This means that the cosine similarity of their embeddings should be around 1. On the other hand if two nodes are in different components, the random walks spanning from them have nothing in common as the set of nodes that constitute them are disjoints. Therefore their embedding vectors should be orthogonal and the cosine similarity shoud be around 0.

2 Question 2

The two embedding matrixes are related to each other through a symetry along the axis $y = 0$, $X_2 = Sym(X_1)$. Since a symetry is an orthogonal matrix in $\mathbb{R}^2$, the distances between two rows i and j are equal for both X_1 and X_2, meaning that both embedding matrixes are equally informative.

3 Question 3

A message passing layer adds to a node the feature information of its neighbours. One message passing layer adds first order proximity information to a node, two message passing layers adds second order proximity information to a node and more than two layers adds higer order proximity information to a node. A GNN typically contains more than one message passing layer because second order information helps classify a node by adding relevant information on its local environment. Adding too many message passing layers provides a node of the global structure of the graph and therefore deteriorates the quality of the classification. This is why in task 12 when we initialize the features of the nodes to the same value, the classification becomes very bad.

4 Question 4

The adjacency matrix A of G is $\begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$ therefore $\hat{A} = I_d + A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$,

$D^{-\frac{1}{2}} = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{3}} & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{3}} & 0 \\ 0 & 0 & 0 & \frac{1}{\sqrt{2}} \end{pmatrix}$ and $\tilde{A} = \begin{pmatrix} \frac{1}{2} & \frac{1}{\sqrt{6}} & 0 & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{3} & \frac{1}{3} & 0 \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{1}{\sqrt{6}} & \frac{1}{2} \end{pmatrix}$

Then $Z^0 = Relu(\tilde{A}XW^0)$=$Relu\left(\begin{pmatrix} \frac{1}{2} & \frac{1}{\sqrt{6}} & 0 & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{3} & \frac{1}{3} & 0 \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{1}{\sqrt{6}} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 0.5 & -0.2 \\ 0.5 & -0.2 \\ 0.5 & -0.2 \\ 0.5 & -0.2 \end{pmatrix}\right) = \begin{pmatrix} 0.454 & 0 \\ 0.537 & 0 \\ 0.537 & 0 \\ 0.454 & 0 \end{pmatrix}$

Finally $Z^1 = Relu(\tilde{A}Z^0W^1) = Relu\left(\begin{pmatrix} \frac{1}{2} & \frac{1}{\sqrt{6}} & 0 & 0 \\ \frac{1}{\sqrt{6}} & \frac{1}{3} & \frac{1}{3} & 0 \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{\sqrt{6}} \\ 0 & 0 & \frac{1}{\sqrt{6}} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 0.454 & 0 \\ 0.537 & 0 \\ 0.537 & 0 \\ 0.454 & 0 \end{pmatrix} \begin{pmatrix} 0.3 & -0.4 & 0.8 & 0.5 \\ * & * & * & * \end{pmatrix}\right)$ which implies $Z^1 = \begin{pmatrix} 0.068 \\ 0.199 \\ 0.236 \\ 0.289 \end{pmatrix}$ We notice that Z^1 contains no information on the nodes as values for the two nodes of degree 1 are very different.