Mercier Marine
marine.mer98@gmail.com

Lab session # 5
ALTEGRAD 2021

01/29/22

1 Question 1:

In a Graph Autoencoder, Z is supposed to be a representation in lower dimension of the graph. A tradeoff is needed between longer embeddings which preserve more information while they induce higher time and space complexity and shorter embeddings. Nevertheless, if the number of column of Z is larger than the number of the nodes, the autoencoder is useless and performs poorly. Z would be a high dimensional representation of the graph, either whith aditional noise or repetitive information.

2 Question 2:

It's possible to initialize the node attribute matrix X to be the same as adjacency matrix A.

3 Question 3 :

There are $\binom{\binom{n}{2}}{M}$ graphs with n nodes and M edges have equal probability of $p^m(1-p)^{\binom{n}{2}-m}$. The number of edges in the graph follows a binomial distribution With M the number of edges of a graph :

$$\begin{aligned}\mathbb{E}[M] &= \sum_{m=0}^{\binom{n}{2}} m\binom{\binom{n}{2}}{m} p^m(1-p)^{\binom{n}{2}-m} \\ &= \binom{n}{2}p\end{aligned}$$

Likewise, the variance

$$V[M] = \binom{n}{2}p(1-p)$$

Therefore with p=0.2, the expected number of edges is $\mathbb{E}[M]_{p=0.2} = 21$ edges and the variance is 16.8. With p = 0.4, the expected number of edges is $\mathbb{E}[M]_{p=0.4} = 42$ and the variance is 12.6.

4 Question 4:

With sum readout function :

$$Z_G = \begin{pmatrix} z_{G_1} \\ z_{G_2} \\ z_{G_3} \end{pmatrix} = \begin{pmatrix} 0.7 & -0.49 & 2.39 \\ 3 & 0.4 & 2 \\ 1.5 & 0.2 & 1 \end{pmatrix}$$

With mean readout function :

$$Z_G = \begin{pmatrix} z_{G_1} \\ z_{G_2} \\ z_{G_3} \end{pmatrix} = \begin{pmatrix} 0.233 & -0.163 & 0.796 \\ 0.75 & 0.1 & 0.5 \\ 0.75 & 0.1 & 0.5 \end{pmatrix}$$

With max readout function :

$$Z_G = \begin{pmatrix} z_{G_1} \\ z_{G_2} \\ z_{G_3} \end{pmatrix} = \begin{pmatrix} 0.89 & 0.34 & 1.31 \\ 0.89 & 0.34 & 1.31 \\ 0.89 & 0.34 & 1.31 \end{pmatrix}$$

Therefore the sum readout function distinguishes thes graph best.