Machine Learning for Graphs and Sequential Data Exercise Sheet 10

Graphs & Networks, Generative Models

Problem 1: An unweighted, undirected graph without self-loops represented by an adjacency matrix $A \in \{0,1\}^{N \times N}$ is given. Prove that the number of triangles in the graph is equal to $\frac{1}{6}\operatorname{trace}(A^3)$ and that this term is in turn equal to $\frac{1}{6}\sum_i \lambda_i^3$ where λ_i are the eigenvalues of the adjacency matrix A. *Hint:* Show first that A^k_{ij} is the number of walks of length k from node i to node j.

Problem 2: Given is an Erdös-Renyi graph consisting of N nodes, with the edge probability $p \in [0,1]$. Derive the probability p_k that a node in the graph has degree equal to exactly k.

Problem 3: Given is an Erdös-Renyi graph consisting of N nodes with edge probability $p \in [0,1]$. What is the expected number of triangles in this graph?

Problem 4: Given are 6 graphs $\{G_1, \ldots, G_6\}$, which exhibit the properties listed in Table 1. Five of them have been synthetically generated, while one is a real graph. Assign the graphs $\{G_1, \ldots, G_6\}$ to the following models (one each) and briefly justify each answer!

a) Erdös-Renyi model
b) Stochastic block model with 5 clusters
c) Stochastic block model with 10 clusters
d) Stochastic block model with core-periphery structure
e) Initial attractiveness model
f) Real graph

Hint: for information about the "eigengap" see Sec. 8.3 in this tutorial

Problem 5: Compare the two following graph generation processes.

- Graph G_1 is generated by a stochastic block model. It consists of N nodes partitioned into $K = 2$ communities. Both communities consist of exactly $N/2$ nodes, and $\boldsymbol{\eta} = \begin{pmatrix} a & b \\ b & a \end{pmatrix}$.
- Graph G_2 is an Erdös-Renyi graph of N nodes and edge probability p.

Given the probabilities a and b, for which values of p will the expected number of triangles in G_2 be *larger* than the expected number of triangles in G_1?

Problem 1: An unweighted, undirected graph without self-loops represented by an adjacency matrix $A \in \{0,1\}^{N \times N}$ is given. Prove that the number of triangles in the graph is equal to $\frac{1}{6}\,\text{trace}(A^3)$ and that this term is in turn equal to $\frac{1}{6}\sum_i \lambda_i^3$ where λ_i are the eigenvalues of the adjacency matrix A. *Hint:* Show first that A_{ij}^k is the number of walks of length k from node i to node j.

$A_{ij}^{k} =$

$\text{trace}(A) = \sum_{i=1}^{N} \lambda_i$

$A^3 v = A^2 A v = \lambda A^2 v = \lambda^3 v$

Problem 2: Given is an Erdös-Renyi graph consisting of N nodes, with the edge probability $p \in [0, 1]$. Derive the probability p_k that a node in the graph has degree equal to exactly k.

$$d \sim \text{Binomial}(N-1, p)$$

Probability of vertex has degree k $p_k = \binom{N-1}{k} \cdot p^k \cdot (1-p)^{N-1-k} \approx \frac{z^k e^{-z}}{k!}$, $z = p(N-1)$

Problem 3: Given is an Erdös-Renyi graph consisting of N nodes with edge probability $p \in [0, 1]$. What is the expected number of triangles in this graph?

$\binom{N}{3}$?

Problem 4: Given are 6 graphs $\{G_1, \ldots, G_6\}$, which exhibit the properties listed in Table 1. Five of them have been synthetically generated, while one is a real graph. Assign the graphs $\{G_1, \ldots, G_6\}$ to the following models (one each) and briefly justify each answer!

a) Erdös-Renyi model ~ G
b) Stochastic block model with 5 clusters
c) Stochastic block model with 10 clusters
d) Stochastic block model with core-periphery structure
e) Initial attractiveness model ~ pow law
f) Real graph

Hint: for information about the "eigengap" see Sec. 8.3 in this tutorial

Table 1: Graphs $\{G_1, \ldots, G_6\}$

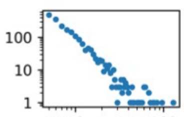
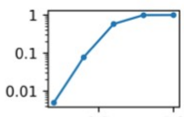
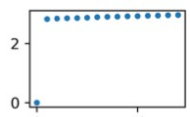
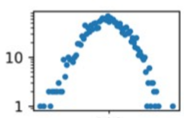
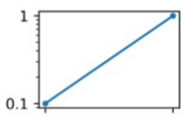
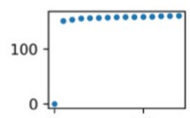
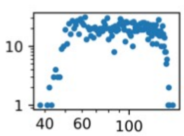
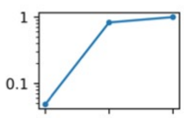
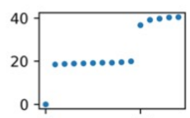
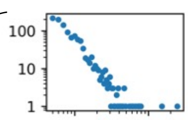
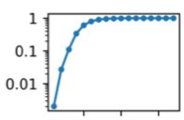
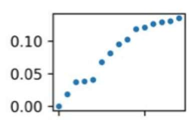
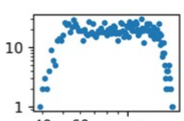
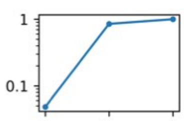
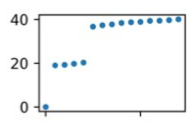
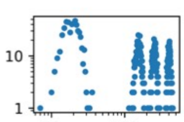
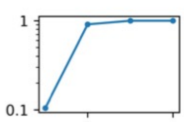
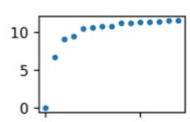
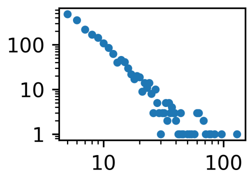
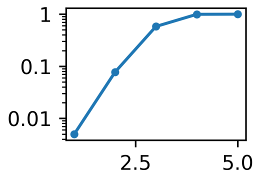
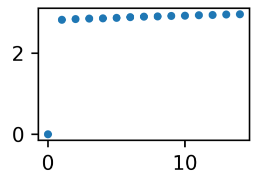
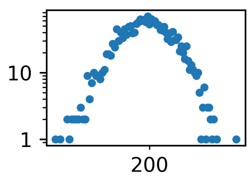
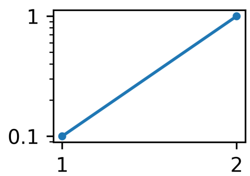
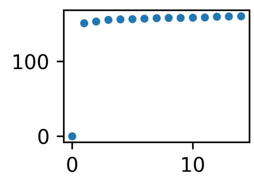
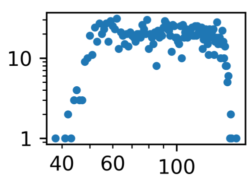
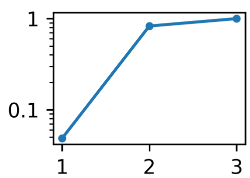
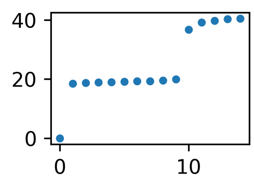
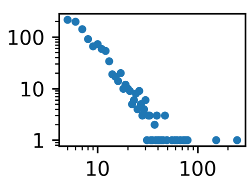
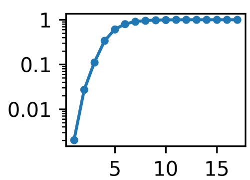
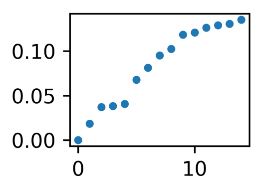
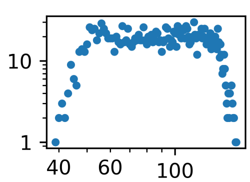
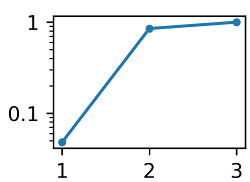
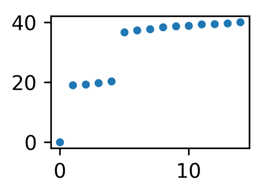
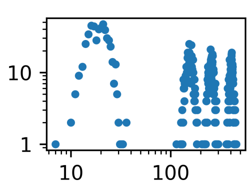
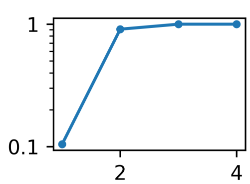
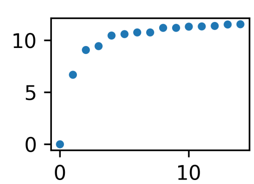
ID	Degree distribution	Hop plot	Smallest eigenvalues of Laplacian	Clustering coeff.
po-L G_1				0.013 ~~d~~ e
G_2				0.100 (a)
G_3				0.145 (c)
po-L G_4				0.278 f
G_5				0.275 (b)
G_6				0.191 ~~e~~ d

Table 1: Graphs $\{G_1, \ldots, G_6\}$

ID	Degree distribution	Hop plot	Smallest eigenvalues of Laplacian	Clustering coeff.
G_1				0.013
G_2				0.100
G_3				0.145
G_4				0.278
G_5				0.275
G_6				0.191

Problem 5: Compare the two following graph generation processes.

- Graph G_1 is generated by a stochastic block model. It consists of N nodes partitioned into $K = 2$ communities. Both communities consist of exactly $N/2$ nodes, and $\boldsymbol{\eta} = \begin{pmatrix} a & b \\ b & a \end{pmatrix}$.
- Graph G_2 is an Erdös-Renyi graph of N nodes and edge probability p.

Given the probabilities a and b, for which values of p will the expected number of triangles in G_2 be *larger* than the expected number of triangles in G_1?

$\eta \begin{pmatrix} a & b \\ b & a \end{pmatrix}$ $\quad \# \triangle = \binom{\frac{N}{2}}{3} a^3$ 内部 $\quad + \quad \times 2$

$\# \triangle = \binom{\frac{N}{2}}{2} \binom{\frac{N}{2}}{1} b^2 \cdot a$ 外部

2 in C_1 $\quad$ 1 in C_2