Prof. Dr. Ingo Scholtes
Chair of Informatics XV
University of Würzburg

Exercise Sheet 08

Published: January 12, 2023
Due: January 19, 2023
Total points: 10

1. **Basics of Markov chain models**

   Consider a graph $G = (V, E)$ with the following adjacency matrix:

$$\mathbb{A} = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 2 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

   (a) Assuming that the row and column numbers correspond to nodes (enumerated from 1 to 5), calculate the probability that a random walker starting in node 1 will traverse the following sequence of nodes: $(1, 2, 3, 2, 4, 1, 5, 1, 2)$. $\boxed{\text{1P}}$

   (b) Does a unique stationary distribution exist for a random walk in graph G? If yes, calculate the stationary distribution, if not explain why. $\boxed{\text{1P}}$

   (c) What is the relative frequency at which we expect the sequence of nodes $(1, 2)$ to appear in an infinite random walk on graph G? $\boxed{\text{1P}}$

2. **Random Walks and Node Centralities**

   Consider a (small) directed network in the `netzschleuder` (e.g. the `highschool` data). Check whether a unique stationary distribution exists. If necessary ensure that the network is aperiodic.

   (a) Compute the stationary visitation probabilities of nodes and rank the nodes based on those probabilities. Compare the resulting ranking with a ranking based on the in -or out-degree of nodes or other path-based measures. Explain what you observe. $\boxed{\text{2P}}$

   (b) Make the network undirected and compare the ranking of nodes based on stationary visitation probabilities to a degree-based ranking. Can you explain your observation in an analytical way? $\boxed{\text{2P}}$

3. **Diffusion speed in networks**

   Consider microstates generated by the (one-dimensional) Watts-Strogatz model with parameters $n = 100$, $s = 3$, and three rewiring probabilities $p = 0$, $p = 0.1$, and $p = 1$.

   (a) Use the transition matrix to compute the stationary distributions $\vec{\pi}$ of a random walk on these three microstates. *Hint:* Use `numpy` or `scipy` functions to calculate eigenvectors. $\boxed{\text{1P}}$

   (b) Assume that $\vec{\pi}^{(0)}$ is a random initial distribution of a random walk, where we assign probability one to a randomly chosen node. We further assume that $\vec{\pi}^{(t)}$ denotes the visitation probabilities of nodes after t steps. Write a `python` function that computes the total variation distance between two distributions $\vec{\pi}$ and $\vec{\pi}'$. Use your function to calculate the total variation distances $|\vec{\pi} - \vec{\pi}^{(t)}|$ between the stationary distribution and the visitation probabilities after t steps for different values of t and 50 different random initial distributions. Repeat your experiment for the three microstates mentioned above and plot the evolution of the average total variation distance to the stationary distribution over time. $\boxed{\text{2P}}$

   In which network is the speed of convergence to the stationary distribution largest? Which one exhibits the slowest convergence speed?