MATHEMATICAL FOUNDATIONS OF DATA SCIENCE

(CS6660)

M. TECH. IN DATA SCIENCES

IIT HYDERABAD

भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

MAUNENDRA SANKAR DESARKAR

Assignment-1 (Linear Algebra)

- Generate a Random Graph.

- How to generate?
  - For each possible edge, generate a random number p. If $p \leq \tau$, add the edge to the graph. Else, do not add the edge to the graph.
  - Note: τ is a parameter of the graph generation algorithm. Its value remains unchanged throughout the graph generation process.
  - In the adjacency matrix, set diagonal entries to zeros.

Task-1

- Fix a value of $\tau \leq 0.2$. Let G_n be a random directed graph of n nodes, generated using the procedure discussed in the previous slide, with this value of τ. Let A_n be the adjacency matrix. Use $n = 50, 100, 200$.
- Find $A_n^2, A_n^3, A_n^4, A_n^5$.

- What do the values in A_n^k indicate?
- For each A_n^k, plot the histogram of values. There should be one plot for each A_n^k. What do the histograms tell you?

Task-2

- Set $\tau' = 2\tau$. Generate 3 more graphs with n =50, 100 and 200 nodes. Let these adjacency matrices be B_n.
- Find B_n^2, B_n^3.

- Plot the histograms for B_n^2, B_n^3 for each n.
- How do A_n^2, A_n^3 compare with B_n^2, B_n^3? What does it tell you?
- Do you have any other observations from the plots?

Task-3

- Consider again the Adjacency Matrix A_n.
- For each A_n^k computed earlier, change the non-zero values in it to 1.
- Set $X_n = A_n$
- Compute the matrix $\mathrm{X}_n^2 = X_n \bigoplus A_n^2$. Here, $\bigoplus$ is an elementwise LOGICAL-OR operator. I.e., If $C = A \bigoplus B$ for Boolean matrices A and B, then $C_{ij} = A_{ij}$ $\mathrm{LOGICAL\text{-}OR}$ B_{ij}.
- Compute the matrix $\mathrm{X}_n^3 = X_n^2 \bigoplus A_n^3$
- Compute the matrix $\mathrm{X}_n^4 = X_n^3 \bigoplus A_n^4$
- Compute the matrix $\mathrm{X}_n^5 = X_n^4 \bigoplus A_n^5$
- Plot a graph showing (number of 1s in X_n^k)-vs-k.
- What observations can you make from the graph?
- Suppose for a particular integer y, all entries in X_n^y are 1. What does it indicate?

Submission

- Submit your code. One function/module for each of the following functionalities have to be submitted:
  - Graph generator
  - Function that accepts the adjacency matrix A, an integer k (and any other parameters if needed) as input and writes to file the matrices $A^2, \ldots, A^k$ (one matrix to one file) [as given in task-1]
  - Function that accepts the adjacency matrix A, an integer k (and any other parameters if needed) as input and writes to file the matrices $X^2, \ldots, X^k$ (one matrix to one file). Refer to Task-3 [Slide-5] for details of this computation.
- You may code in any language of your choice.
- You may use any plotting tool of your choice.

Submission

- Adjacency matrices $A_{50}, A_{100}, A_{200}, B_{50}, B_{100}, B_{200}$ each in a separate text file. Values in the row can be space-separated/comma-separated. Keep the filename extension as .txt.
- A report. The report should contain your graphs and answers to the questions asked in different tasks. If you want to include any other point/observation that you may have made by looking at the results, feel free to include that. Make sure to mark your name and roll number in the report.
- Zip together the code files, adjacency matrix files and the report. Name the zip file as <your-roll-no>.zip.
- Department plagiarism policy: https://cse.iith.ac.in/?q=node/254