Asmt 7: Graphs

Turn in through GradeScope by 5pm:
Wednesday, April 23
100 points

Overview

In this assignment you will explore different approaches to analyzing Graphs via Markov chains.

You will use one data sets for this assignment:

- http://www.cs.utah.edu/~jeffp/teaching/DM/A7/M.csv

As usual, it is recommended that you use LaTeX or another method which can properly display mathematical notation for this assignment. If you do not, you may lose points if your assignment is difficult to read or hard to follow. Find a sample form in this directory: http://www.cs.utah.edu/~jeffp/teaching/latex/

You could also utilize an LaTeX template specifically created for this assignment. Click here.

1 Finding q_* (100 points)

We will consider four ways to find $q_* = M^t q_0$ as $t \to \infty$.

Matrix Power: Choose some large enough value t, and create M^t. Then apply $q_* = (M^t)q_0$. There are two ways to create M^t, first we can just let $M^{i+1} = M^i * M$, repeating this process $t-1$ times. Alternatively, (for simplicity assume t is a power of 2), then in $\log_2 t$ steps create $M^{2i} = M^i * M^i$.

State Propagation: Iterate $q_{i+1} = M * q_i$ for some large enough number t iterations.

Random Walk: Starting with a fixed state $q_0 = [0, 0, \ldots, 1, \ldots, 0, 0]^T$ where there is only a 1 at the ith entry, and then transition to a new state with only a 1 in the jth entry by choosing a new location proportional to the values in the ith column of M. Iterate this some large number t_0 of steps to get state q'_0. (This is the *burn in period*.)

Now make t new step starting at q'_0 and record the location after each step. Keep track of how many times you have recorded each location and estimate q_* as the normalized version (recall $\|q_*\|_1 = 1$) of the vector of these counts.

Eigen-Analysis: Compute `LA.eig(M)` and take the first eigenvector after it has been L_1-normalized.

A (40 points): Run each method (with $t = 1024$, $q_0 = [1, 0, 0, \ldots, 0]^T$ and $t_0 = 100$ when needed) and report the answers.

B (20 points): Rerun the Matrix Power and State Propagation techniques with $q_0 = [0.1, 0.1, \ldots, 0.1]^T$. For what value of t is required to get as close to the true answer as the older initial state?

C (24 points): Explain at least one **Pro** and one **Con** of each approach. The **Pro** should explain a situation when it is the best option to use. The **Con** should explain why another approach may be better for some situation.

D (8 points): Is the Markov chain *ergodic*? Explain why or why not.

E (8 points): Rerun at least one of the algorithms using teleportation with $\beta = 0.15$ probability. Report q_* and discuss how it changed.

2 BONUS: Dozen Step Probability (1 point)

Consider the graph encoded in matrix M. Let the rows (and columns) index nodes 0 through 9 (from first to last). Start at node 3. Calculate the probability at after 12 steps you are in node 0.