A Random Walk Through North America

Clark Alexander, Ivonne Garcia, Abdul Alanazi,
Ross Nelson, Vincent Chan

August 29, 2017

Abstract

We build a discrete graph of the contiguous states of The United States, Mexico, and Canada including Washington, D.C. and The Federal District of Mexico City. From this graph we determine its chromatic number and consider two types of random walks on the graph. The First is the classical random walk where we show that the walk converges to a stable vector in time $O(\lambda_2^t)$ where λ_2 is the second largest eigenvalue of the walking matrix. Second, we consider a continuous time quantum walk on North America and show that the walk does not converge, but that we can increase hitting time.

1 Introduction

This project grew out of several ideas bandied about in a group of undergraduate math and computer science classes, and a desire to compute a nontrivial example. Originally we were talking about Eulerian and Hamiltonian circuits on connected graphs and showing that these cannot exist on the stand alone graph of the United States or Mexico since they each have vertices of degree one (Maine and Baja California Sur respectively). Additionally, we discovered that a large amount of the Western United States is three-colorable, and so we set out to find the obstructions to three-colorability in the entire graph. Interestingly,but perhaps not surprisingly the addition of Canada breaks the three-colorability of the Western United States. Additionally, the addition of Mexico to the southern border of the United States prevents the graph from having chromatic number 3. We also know that all planar graphs have chromatic number less or equal to 4, so we set out to present a visually pleasing planar four-colored graph of North America. In section 2, we present our graph with and without chromaticity. In a class on quantum algorithms we were discussing discrete time quantum walks and realized that continuous time quantum walks are somewhat easier to compute. Thus we set up the rest of our paper as such: in section 3 we present some of the early notions of classical random walks on graph, including convergence in undirected graphs to a stable vector, hitting and mixing time. In section 4 we present some of these calculations. In section 5 we present the difference between classical and continuous quantum walks and show that the

quantum walk does not converge. Finally in the appendix we present our open source code for computing random walks and continuous time quantum walks in octave. The graph and the code will be freely available on github at (give address, probably put this in the references)

2   The Graph of North America

include graphs, scale them, explain what's going on. Show the coloration.

3   The Chromatic Number of North America

State the four-color theorem, give a reference. Use Ross's proof that this graph has chromatic number > 3. There are a couple of other instances, for example and vertex 74, there is a very clear copy of K_4. See if we can find a couple more.

4   Generalities on Classical Random Walks

To begin our discussion of classical random walks on graphs let's begin with the easy example of walking along a line. Suppose we have a graph of vertices labeled $\{-n, -n+1, \ldots, -1, 0, 1, \ldots, n-1, n\}$. Our random walker will begin at the vertex labeled zero and flip a coin. If this coin lands on "heads" then the walker moves to vertex 1, if "tails" then the walker moves to vertex -1. Continuing the process at any given vertex, the walker will flip a coin and "heads" mean the walker will move from v to $v+1$. If the coin lands "tails" then the walker moves from v to $v-1$. After some number of coin flips we see where the walker is standing. This process gives us a probability distribution identical to the balanced binomial distribution.
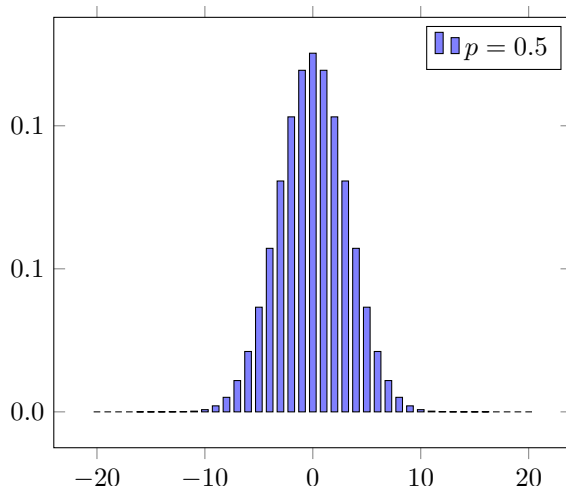
In order to formalize this we consider a general (undirected) graph G with vertex set V and edge set E. We need not require the graph to be connected, but in all the examples that follow, we will only discuss the case of connected graphs. Consider the walking matrix W which we define by
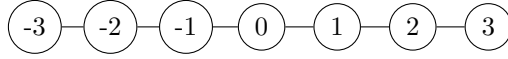
$$W_{i,j} = \begin{cases} 1/deg(v_i) & \text{if } (v_i, v_j) \in E \\ 0 & \text{else} \end{cases} \tag{4.1}$$

Consider then our graph from above with 7 vertices

$$V = \{-3, -2, -1, 0, 1, 2, 3\}$$

and 6 edges

$$E = \{(-3, -2), (-2, -1), \ldots, (2, 3)\}$$

We have the following relevant matrices

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \tag{4.2}$$

$$D = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{4.3}$$

$$W = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \tag{4.4}$$

$$L = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 2 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 \end{bmatrix} \tag{4.5}$$

These are the adjacency matrix A, the degree matrix D, the stochastic walking matrix $W = AD^{-1}$, and the Laplacian matrix $L = D - A$.

Since W is a stochastic matrix is takes probability distributions to probability distributions. Then we define the probability distribution of a random walk starting at vertex v and walking for k steps by

$$p^{(k)}(v) = (vW^k) \qquad (4.6)$$

In particular the probability of finding the walker starting from vertex i and walking k steps at vertex j is $p_{i,j}^{(k)}$.

Notice that while we tend to use column vectors in linear algebra class, we use row vectors to perform our random walk. We reason through this on graphs which are not regular as we check all the adjacencies first and then apply the probabilities. Equivalently we could consider the random walk $p^{(k)} = (W^t)^k v^t$.

4.1 Other Notions of Random Walks on Graphs

Given a graph G with relevant matrices A, D, W, L we define the following notions.

Definition 1. The *hitting time* $H(i, j)$ from vertex i to vertex j is the expectation of the number of steps to the first time a random walker is at vertex j starting at vertex i.

Definition 2. The *commute time* $\kappa(i, j)$ is the expectation time for a random to return to vertex i after having reached vertex j.

$$\kappa(i, j) = H(i, j) + H(j, i)$$

Definition 3. The *cover time* from vertex v is the expected time it takes to reach every vertex in the graph.

In a regular graph this is a constant for any given vertex. If the graph is not regular and no particular vertex v is given the cover time is the maximum amongst all vertices in the graph.
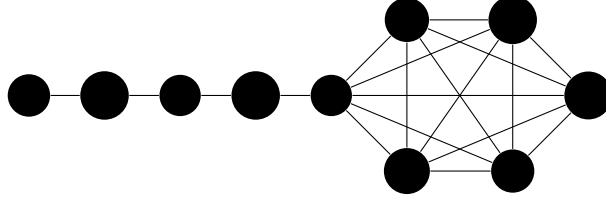
Definition 4. The *mixing rate* is how fast a walk converges to a stationary distribution. For a connected graph the stationary distribution π is given by

$$\pi(j) = \frac{deg(j)}{2|E|} \qquad (4.7)$$

and the mixing rate is defined by

$$\mu = \limsup_{k \to \infty} \max_{i,j} \left| p_{ij}^{(k)} - \pi(j) \right| \qquad (4.8)$$

4

There are several important ideas to consider with the above definitions. First, hitting time is not symmetric. Consider the lollipop graph

with a complete graph on the right hand side and a simple 2-regular graph glued to it on the left. The hitting time from the leftmost vertex to the rightmost vertex is significantly smaller than the hitting time from the rightmost vertex to the leftmost vertex.

$$H(\text{left}, \text{right}) = < H(\text{right}, \text{left}) =$$

5 The Walk Through North America

Here we'll show our explicit computations, via Octave code. present multiple graphs, use starting vectors in places like Missouri, Tennessee, Texas. Linear combinations of such. Compute the constant $C(v_0)$ for each starting vector v_0 where

$$\|v_0 W^k - \pi\| \sim C(v_0)\lambda_2^t$$

where π is the stable vector in this case. We'll do this for the lazy walk and the parametrized lazy walk. Let's also consider probabilities of starting in USA by population, Mexico by population, Canada by population, and also in the entirety of North America by population. What is the smallest possible C? given that we do not start at the stable vector? What is the difference in the values of $C(v_0)$ for the standard walks and what about C for the parametrized lazy walk with matrix

$$\alpha I + (1 - \alpha)W$$

Generally $\alpha = \frac{1}{2}$ for the lazy walk, but I don't see any reason why we need exactly $\frac{1}{2}$. Let's play around with this. We can automate some code that will pick this info.

6 A Quantum Mechanical Extension

This again is my (Clark's) section. I'll explain how continuous time quantum walks work, and then we'll compute some. We'll show that this walk does not converge to the stable vector. The easiest way to do this is to use π as our starting and stopping vector and show that the probability is significantly smaller than 1 for all times $t > 0$.

Appendix: Code in Octave

Present the code in octave

References

[1] Add some things here

[2] github page with codes.