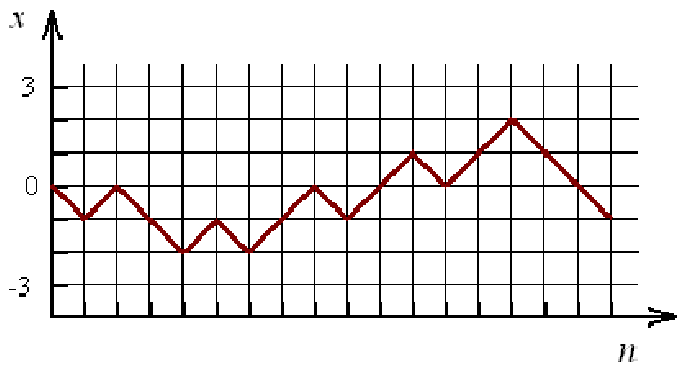
Random Walks DRP Summary

Noah Andrew McMahon

March 17, 2023

1 Part 1:

A random walk depicts seemingly random movement from one point in space to any other location. We started by talking about simple symmetric random walks, in which you start at position x = 0 and there is a equal probability moving up or down one position. A graph of a 1D simple symmetric walk looks something akin to this:

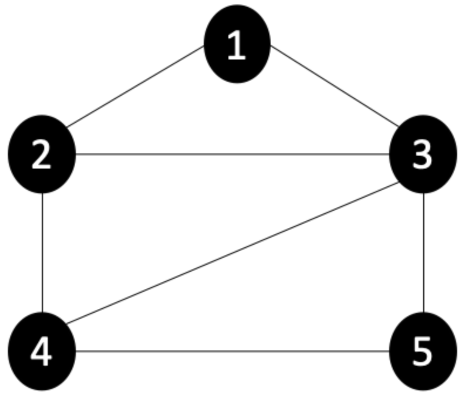


In this example, the probability of moving one unit down is equal to the probability of moving one unit up [P(down) = P(up) = $\frac{1}{2}$]. We further discussed the low probabilities of moving all the way up solely [P(up only) = $\frac{1}{2}^n$] or moving all the way down solely [P(down only) = $\frac{1}{2}^n$]. We also discussed all of the many ways to start at x = 0 and end back up at x = 0, using combinations.

2 Part 2:

We then continued on by learning about random walks visually and through basic matrices.

In the example, we have a random graph with five nodes upon which a random walk can be performed. In order to create the transition matrix M that



shows the probability of going from node i to node j with every direction having equal probability, we abide by the following formula:

$$\mathrm{M}_{ij} = \begin{cases} \frac{1}{deg(v_j)} & if\ (v_i, v_j)\ is\ an\ edge\ in\ the\ graph\ G \\ 0 & otherwise \end{cases}$$

This creates the following matrix:

$$\mathrm{M} = \begin{bmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ \frac{1}{3} & 0 & \frac{1}{3} & \frac{1}{3} & 0 \\ \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{1}{3} & \frac{1}{3} & 0 & \frac{1}{3} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 \end{bmatrix}$$

This matrix shows all of the probabilities of moving from one node to the next node. For example, the probability of performing one random step from node 5 to node 4 is $\mathrm{M}(5,\ 4) = \frac{1}{2}$.

Furthermore, if we want to find the probability of starting in a certain node, performing multiple random steps, and ending up in another certain node, all we have to do is raise the M matrix to the power of number of random steps performed. For example, if we wanted to know all of the probabilities after three random steps, we just raise the matrix M to the third power:

$$\mathrm{M}^3 = \begin{bmatrix} \frac{1}{12} & \frac{41}{144} & \frac{25}{72} & \frac{7}{48} & \frac{5}{36} \\ \frac{41}{216} & \frac{5}{36} & \frac{31}{108} & \frac{31}{108} & \frac{7}{72} \\ \frac{25}{144} & \frac{31}{144} & \frac{2}{9} & \frac{31}{144} & \frac{25}{144} \\ \frac{7}{72} & \frac{31}{108} & \frac{31}{108} & \frac{5}{36} & \frac{41}{216} \\ \frac{5}{36} & \frac{7}{48} & \frac{25}{72} & \frac{41}{144} & \frac{1}{12} \end{bmatrix}$$

So, the probability of starting at node 5 and ending up at node 1 after three random steps is $\mathrm{M}^3(5,\ 1) = \frac{5}{36}$.

This leads us into our main focus of random walks, which is convergence. Convergence is attained when the probability of ending up at node j is the same regardless of what node i you started at. After running this example through some R code that I coded up for convergence, we get that after 9 random steps, the transition matrix M converges to:

$$M^9 = \begin{bmatrix} 0.1428571 & 0.2142857 & 0.2857143 & 0.2142857 & 0.1428571 \\ 0.1428571 & 0.2142857 & 0.2857143 & 0.2142857 & 0.1428571 \\ 0.1428571 & 0.2142857 & 0.2857143 & 0.2142857 & 0.1428571 \\ 0.1428571 & 0.2142857 & 0.2857143 & 0.2142857 & 0.1428571 \\ 0.1428571 & 0.2142857 & 0.2857143 & 0.2142857 & 0.1428571 \end{bmatrix}$$

This goes to show how regardless of what node you start in, the probability of converging to the first node is 0.1428571.

3 Part 3:

We then got more technical with the matrix M and breaking it down into simpler components. The matrix M can be broken down into the product of the matrices D^{-1} and W, in which D^{-1} is the inverse of the matrix D that is the diagonal matrix where D_{ii} = deg(i) and W is the matrix of weights under the assumption that not every direction is equally probable. Assuming every direction is no longer equally probable, the elements of the matrix M are now:

$$M_{ij} = \frac{w_{ij}}{deg(i)},$$

where deg(i) $= \sum_j w_{ij}$ now. We then defined a probability cloud, which is conceptually the idea of starting at node i and all of the possible probabilities for each node after t random walks:

$$i \rightarrow M^t[i,:],$$

This row vector is each individual probability starting at node i and then ending at another certain node after t random walks.

Next, we moved on to rephrasing the M matrix as the symmetric matrix S:

$$M = D^{-\frac{1}{2}}SD^{\frac{1}{2}} = D^{-\frac{1}{2}}V\Lambda V^T D^{\frac{1}{2}} = (D^{-\frac{1}{2}}V)\Lambda(D^{\frac{1}{2}}V)^T = \Phi\Lambda\Psi^T,$$

where $\Phi = D^{-\frac{1}{2}}V$, Λ is a diagonal matrix with the eigenvalues, and $\Psi = D^{\frac{1}{2}}V$. Φ essentially is the product of the diagonal matrix of deg(i) D to the $-\frac{1}{2}$ power and the eigenvector matrix V. Also, Ψ is the product of the diagonal matrix of deg(i) D to the $\frac{1}{2}$ power and the eigenvector matrix V. Φ and Ψ form a biorthogonal system in that $\Phi\Psi^T = I_{nxn}$ and are the right and left eigenvectors of M, respectively. In addition:

$$M = \sum_{k=1}^{n} \lambda_k \varphi_k \psi_k^T$$

Moving forward, the power rule now works like:

$$\mathrm{M}^t = \sum_{k=1}^{n} \lambda_k^t \varphi_k \psi_k^T$$

which is much simpler than raising the whole matrix to a power t for the number of random walks t. In addition, the probability cloud now expands to:

$$\mathrm{M}^t[\mathrm{i},:] = \begin{bmatrix} \lambda_1^t \varphi_1(i) \\ \lambda_2^t \varphi_2(i) \\ \cdots \\ \lambda_n^t \varphi_n(i) \end{bmatrix}$$

We then looked through a proposition proving that the first element of the above matrix is redundant since it is a multiple of 1.

We then looked at a side note about disconnected and bipartite graphs, which refers to graphs being in two separate, distinct groups. In order to solve this issue, we make the walk lazy, such that there is a 50% chance to move from the node or a 50% chance to stay at the same node:

$$\mathrm{M}' = \frac{1}{2}\mathrm{M} + \frac{1}{2}\mathrm{I}$$

Using the new breakdown of the matrix M, we now get two new diffusion maps:

Diffusion Map:

$$\varphi_t(v_i) = \begin{bmatrix} \lambda_2^t \varphi_2(i) \\ \lambda_3^t \varphi_3(i) \\ \cdots \\ \lambda_n^t \varphi_n(i) \end{bmatrix}$$

4 Part 4:

The lazy walk matrix can be denoted as follows:

$$\mathrm{W}_G = (\tfrac{1}{2})(\mathrm{I} + \mathrm{M}_G \mathrm{D}_G^{-1}).$$

In addition, the Laplacian matrix can be defined as:

$$\mathrm{L} = \mathrm{D} - \mathrm{A},$$

where D is the diagonal degree matrix and A is the adjacency matrix where A(i, j) represents node i being adjacent to node j, and the rest of the adjacency matrix being zero.
This matrix can be normalized so that:

$$\mathrm{N} = \mathrm{D}^{-\frac{1}{2}} \mathrm{L} \mathrm{D}^{-\frac{1}{2}} = \mathrm{I} - \mathrm{D}^{-\frac{1}{2}} \mathrm{M} \mathrm{D}^{-\frac{1}{2}}.$$

Doing some matrix algebra, we get that:

$$W = (\frac{1}{2})(I + M_G D_G^{-1})$$
$$W = \frac{1}{2}I + \frac{1}{2}M_G D_G^{-1}$$
$$W = I - \frac{1}{2}I + \frac{1}{2}M_G D_G^{-1}$$
$$W = I - \frac{1}{2}D^{\frac{1}{2}}(I - {D_G}^{-\frac{1}{2}}M_G {D_G}^{-\frac{1}{2}}){D_G}^{-\frac{1}{2}}$$
$$W = I - \frac{1}{2}({D_G}^{\frac{1}{2}}I{D_G}^{-\frac{1}{2}} - {D_G}^{\frac{1}{2}}{D_G}^{-\frac{1}{2}}M_G{D_G}^{-\frac{1}{2}}{D_G}^{-\frac{1}{2}})$$
$$W = I - \frac{1}{2}(I - M_G{D_G}^{-1})$$

This new rewriting of the lazy walk matrix is done in order to make it symmetric as well as have nicer values through its normalization. This new walk matrix has eigenvectors of $D^{\frac{1}{2}}$ times the eigenvectors of N & the eigenvalues of 1 - (eigenvalues of N)/2.
We then, through more linear algebra, arrive at a stable distribution starting with the π vector:

$$\pi = d/(1^T d)$$

Using linear algebra, we can show that this π vector is, in fact, a right eigenvector of our lazy walk matrix W with eigenvalue 1:

$$MD^{-1}\pi = MD^{-1}d/(1^T d) = M1/(1^T d) = d/(1^T d) = \pi$$
$$W\pi = (1/2)I\pi + (1/2)MD^{-1}\pi = (1/2)\pi + (1/2)\pi = \pi$$

And after even more complex linear algebra, we find that the stable distribution is the π vector, meaning that whatever distribution of nodes you have in your random graph, you will eventually converge to the π vector.

$$D^{\frac{1}{2}}c_1\psi_1 = D^{\frac{1}{2}}\frac{1}{||d^{\frac{1}{2}}||}\frac{d^{\frac{1}{2}}}{||d^{\frac{1}{2}}||} = \frac{d}{||d^{\frac{1}{2}}||^2} = \frac{d}{\sum_j d(j)} = \pi$$

5 Part 5:

An additional application of random walks is dimensional reduction. After creating a diffusion map like the one below, we can remove the first element as it is redundant (its eigenvalue is one).

$$\varphi_t^{(d)}(v_i) = \begin{bmatrix} \lambda_2^t\varphi_2(i) \\ \lambda_3^t\varphi_3(i) \\ \ldots \\ \lambda_{d+1}^t\varphi_{d+1}(i) \end{bmatrix}$$

As we take more and more random steps and the power we raise our eigenvalues to gets larger and larger, we can start to ignore the later elements of this diffusion map as raising the smaller eigenvalues to high powers essentially makes them zero since the eigenvalues are in descending order. This knowledge allows us to reduce the dimensions to only the numerically essential elements.

6 Resources:

First Random Walks Text:
https://people.math.osu.edu/husen.1/teaching/571/random_walks.pdf

Second Random Walks Text:
https://people.math.ethz.ch/ abandeira/BandeiraSingerStrohmer-MDS-draft.pdf

Third Random Walks Text:
lect10-18_rwG%20(5).pdf

My Simulation Code in R:
https://github.com/NoahMcMahon1414/STAT_DRP_2023/blob/main/STAT_499_DRP_Simulation.R