1 Preamble

This lecture introduces random walks. For the purposes of these notes, we assume that all graphs are finite and that time is discrete.

2 Random Walks

2.1 Introduction

Suppose we have $n$ nodes labeled $1, 2, \ldots, n$ and some object that moves around these nodes. Let $X_t$ be a random variable indicating the location of the object at time $t$.

We define $p_{ij} = \Pr(X_{t+1} = j | X_t = i)$. That is, $p_{ij}$ is the probability that if the object is at node $i$ it will move to node $j$ in the next time step.

For example, consider the graph below:

The probability that the object moves from node 1 to node 2 is $\alpha$, the probability the it stays at node 1 is $1 - \alpha$, the probability it moves from node 2 to node 1 is $\beta$, and the probability that it stays at node 2 is $1 - \beta$.

This can also be summarized in a transition matrix:

$$A = \begin{bmatrix} 1 - \alpha & \alpha \\ \beta & 1 - \beta \end{bmatrix}, \text{ where each entry } a_{ij} = p_{ij}$$

2.2 The position probability vector

Define $p_t$ to be a vector of size $1 \times n$ (i.e. 1 row, $n$ columns), where the $i$th entry is the probability that the object is at node $i$ at time $t$. Thus, we get the following equations:

$$p_{t+1} = p_t A$$  \hfill (1)
For example, consider the graph below:

The associated transition matrix is

\[
A = \begin{bmatrix}
0 & 1/4 & 0 & 3/4 \\
1/2 & 0 & 1/3 & 1/6 \\
0 & 0 & 1 & 0 \\
0 & 1/2 & 1/4 & 1/4 \\
\end{bmatrix}
\]

Suppose the object starts at node 1. Thus, \( p_0 = [1 \ 0 \ 0 \ 0] \)

Suppose we want to find the probability that the object ends up at node 4 after exactly 3 time steps. One option is to find all paths of length 3 from node 1 to node 4 and find the probability of each path.

- \( 1 \rightarrow 2 \rightarrow 1 \rightarrow 4 \rightarrow 1/2 \times 1/4 = 3/32 \)
- \( 1 \rightarrow 4 \rightarrow 4 \rightarrow 1/2 \times 3/4 \times 1 = 3/64 \)
- \( 1 \rightarrow 2 \rightarrow 4 \rightarrow 1/2 \times 1/6 \times 1/4 = 1/96 \)
- \( 1 \rightarrow 4 \rightarrow 1 \rightarrow 4 \rightarrow 1/2 \times 3/4 \times 1/6 = 3/48 \)

The sum of these possibilities is 41/192, which is the probability that the object ends up at node 4 (starting from node 1) after exactly 3 time steps.
However, for a larger number of nodes, this strategy becomes impractical very quickly. Fortunately, we can simply use Equation (2) from above.

\[
\therefore p_{t+m} = p_t A^m \Rightarrow p_{0+3} = p_3 = p_0 A^3 = \begin{bmatrix} 3/16 & 7/48 & 29/64 & 41/192 \end{bmatrix}
\]

We can see that the 4\textsuperscript{th} entry of position vector for time step 3 (which corresponds to the probability that the object is at node 4 at time step 3) is exactly the same as in our previous calculation.

### 3 Definitions

- **We say that a node** $i$ **is** accessible **from node** $j$ **if there exists** $n \geq 0$ **such that** $p_{n,j}^i > 0$. That is, there exists some $n \geq 0$ such that an object starting at node $i$ has a probability greater than 0 of reaching node $j$ after $n$ time steps.

- **We say that nodes** $i$ **and** $j$ **communicate** if they are both accessible from each other. If $i$ and $j$ communicate, we write $i \leftrightarrow j$.

- **Note that** $i \leftrightarrow j$ **forms an equivalence relation**. That is, the following properties are true for all nodes $i$ and $j$.
  
  1. Reflexive: $i \leftrightarrow i$
  2. Symmetric: $i \leftrightarrow j \Leftrightarrow j \leftrightarrow i$
  3. Transitive: $(i \leftrightarrow j$ and $j \leftrightarrow k) \Rightarrow i \leftrightarrow k$

Thus, we can partition the nodes into equivalence classes called communication classes (i.e. disjoint sets of nodes that all communicate with each other). Note that each node belongs to exactly one communication class.

- **We say a graph is** irreducible **if there is only one communication class**. This will happen if and only if the graph is strongly connected.

- **Define** $r_{i,j}^t = \Pr((X_t = j \land X_s \neq j, \forall 1 \leq s \leq t - 1) | X_0 = i)$. That is, $r_{i,j}^t$ is the probability that a random walk starts at node $i$ and reaches node $j$ for the first time in exactly $t$ time steps.

Thus, $\sum_{t=1}^{\infty} r_{i,j}^t$ is the probability that we start at node $i$ and eventually get back to node $i$.

We say a node is **recurrent** if this sum equals 1, and **transient** if this sum is less than 1.

- **We say that a node** $i$ **is periodic** if there exists an integer $\Delta > 1$ such that $\Pr(X_{t+s} = j | X_t = j) = 0$ unless $s$ is divisible by $\Delta$. That is, a node $i$ is periodic if all paths from $i$ to $i$ have a length that is divisible by $\Delta$.

For example, consider a directed cycle with 3 nodes:
Clearly all paths from any node \( i \) back to itself have a length that is divisible by 3. Thus, all nodes \( i \) in this graph are periodic with period 3.

- Lastly, we notice the following the two properties. If any node in a communication class is:
  1. recurrent \( \Rightarrow \) all nodes in that communication class are recurrent.
  2. periodic \( \Rightarrow \) all nodes in that communication class are periodic with the same period.