**CS 388R: Randomized Algorithms** Fall 2015

Lecture 23 — November 25, 2015

*Prof. Eric Price* *Scribe: John Kallaugher, Surbhi Goel*

# 1 Overview

This lecture is about Markov Chains, a type of stochastic process where the distribution of the process at time $t$ depends only on the value of the process at time $t-1$.

# 2 Introduction

**Definition 2.0.1** (Markov Chain)**.** *A Markov Chain $(X_t)_{t\in\mathbb{N}}$ is a sequence of random variables on some state space $S$ which obeys the following property:*

$$\forall t > 0, (s_i)_{i=0}^t \in S, \mathbb{P}\left[X_t = s_t \middle| \bigcap_{i=0}^{t-1}(X_i = s_i)\right] = \mathbb{P}\left[X_1 = s_t | X_0 = s_{t-1}\right]$$

We will write these probabilities as a *transition matrix* $P$, where $P_{ij} = \mathbb{P}\left[X_1 = s_j | X_0 = s_i\right]$. Note that $\forall i, \sum_j P_{ij} = 1$ is necessary for $P$ to be a valid transition matrix.

If $q \in \mathbb{R}^{|S|}$ is the distribution of $X$ at time 0, the distribution of $X$ at time $t$ will then be $qP^t$.

## 2.1 Example: Random Walk on a Graph

Let our state space be the vertices of a graph $G = (V, E)$. Then we can define a Markov chain by a random walk on $G$, where at each step the walk jumps to a random neighbour of the current vertex. This gives us the following transition matrix:

$$P_{uv} = \begin{cases} \frac{1}{d(u)} & (u,v) \in E \\ 0 & \text{Otherwise.} \end{cases}$$

# 3 The Fundamental Theorem of Markov Chains

**Definition 3.0.1** (Ergodicity)**.** *A Markov Chain is ergodic if $\exists \Pi \in \mathbb{R}^{|S|}$ such that:*

$$\forall s \in S, \Pi_s > 0$$
$$\lim_{t\to\infty} qP^t = \Pi$$

We will call this $\Pi$ the *stationary distribution* of $X$. Note that when it exists, $\Pi$ is the unique vector $\Pi \in \mathbb{R}^{|S|}$ such that $\Pi P = \Pi$, with $\sum_{s \in S} \Pi_s = 1$ and $\Pi_s \in [0,1]$ for all $s$.

**Theorem 3.0.2** (The Fundamental Theorem of Markov Chains)**.** *Let $X$ be a Markov Chain on a finite state space $S = [n]$ satisfying the following conditions:*

**Irreducibility** *There is a path between any two states which will be followed with $> 0$ probability, i.e. $\forall i, j \in [n]$, $\exists t \mathbb{P}[X_t = j | X_0 = i] > 0$.*

**Aperiodicity** *Let the* period *of a pair of states $u, v$ be the GCD of the length of all paths between them in the Markov chain, i.e.* $\gcd\{t \in \mathbb{N}_{>0} | \mathbb{P}[X_t = v | X_0 = u] > 0\}$. *$X$ is aperiodic if this is 1 for all $u, v$.*

*Then $X$ is ergodic.*

Note that both these conditions are necessary as well as sufficient.

### 3.1 Further Definitions

$$N(i, t) = |\{t \in \mathbb{N} | X_t = i\}|$$

This obeys $\lim_{t \to \infty} \frac{N(i,t)}{t} = \Pi_i$ for an ergodic chain with stationary distribution $\Pi$.

$$h_{u,v} = \mathbb{E}[\min_t \{t | X_t = v\} | X_0 = u]$$

This is called the *hitting time* of $v$ from $u$, and it obeys $h_{i,i} = \frac{1}{\Pi_i}$ for an ergodic chain with stationary distribution $\Pi$.

## 4 Random Walks on Undirected Graphs

We consider a random walk $X$ on a graph $G$ as before, but now with the assumption that $G$ is undirected.

Clearly, $X$ will be irreducible iff $G$ is connected. It can also be shown that it will be aperiodic iff $G$ is not bipartite. The $\Rightarrow$ direction follows from the fact that paths between two sides of a bipartite graph are always of even length, whereas the $\Leftarrow$ direction follows from the fact that a non-bipartite graph always contains a cycle of odd length.

We can always make a walk on a connected graph ergodic simply by adding self-loops to one or more of the vertices.

### 4.1 Ergodic Random Walks on Undirected Graphs

**Theorem 4.1.1.** *If the random walk $X$ on $G$ is ergodic, then its stationary distribution $\Pi$ is given by $\forall v \in V, \Pi_v = \frac{d(v)}{2m}$.*

*Proof.* Let $\Pi$ be as defined above. Then:

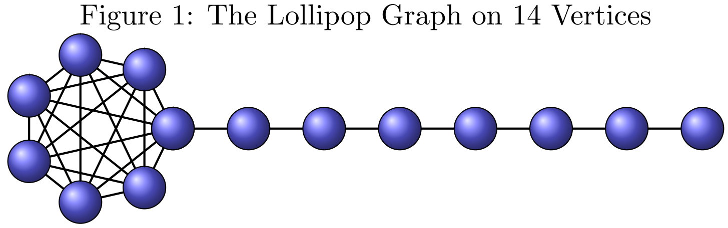
$$\begin{aligned}(\Pi P)_v &= \sum_{u,v \in E} \Pi_u \frac{1}{d(u)} \\ &= \sum_{u,u,v \in E} \frac{1}{2m} \\ &= \frac{d(v)}{2m} \\ &= \Pi_v\end{aligned}$$

So as $\sum_v \Pi_v = \frac{2m}{2m} = 1$, $\Pi$ is the stationary distribution of $X$. □

In general, even on this subset of random walks, the hitting time will not be symmetric, as will be shown in our next example. So we define the commute time $C_{u,v} = h_{u,v} + h_{v,u}$.

### 4.2 Example: The Lollipop Graph

Figure 1: The Lollipop Graph on 14 Vertices



The lollipop graph on $n$ vertices is a clique of $\frac{n}{2}$ vertices connected to a path of $\frac{n}{2}$ vertices. Let $u$ be any vertex in the clique that does not neighbour a vertex in the path, and $v$ be the vertex at the end of the path that does not neighbour the clique. Then $h_{u,v} = \theta(n^3)$ while $h_{v,u} = \theta(n^2)$. This is because it takes $\theta(n)$ time to go from one vertex in the clique to another, and $\theta(n^2)$ time to successfully proceed up the path, but when travelling from $u$ to $v$ the walk will fall back into the clique $\theta(1)$ times as often as it makes it a step along the path to the right, adding an extra factor of $n$ to the hitting time.

## 5 Electrical Resistance and Commute Time of a Graph

View graph $G$ as an electrical network with unit resistors as edges. Let $R_{u,v}$ be the effective resistance between vertices $u$ and $v$. The commute time between $u$ and $v$ in a graph is related to $R_{u,v}$ by $C_{u,v} = 2mR_{u,v}$. We get the following inequalities assuming this relation.

If $(u, v) \in E$,

$$R_{u,v} \leq 1 \therefore C_{u,v} \leq 2m$$

In general, $\forall\ u, v \in V$,

$$R_{u,v} \leq n - 1 \therefore C_{u,v} \leq 2m(n-1) < n^3$$

We inject $d(v)$ amperes of current into $\forall v \in V$. Subsequently we pick some vertex $u \in V$ and remove $2m$ current from $u$ leaving net $d(u) - 2m$ current at $u$. Now we get voltages $x_v$ $\forall v \in V$. Assume we have $x_v - x_u = h_{v,u}$ $\forall v \neq u \in V$ (will prove subsequently). Let $L$ be the Laplacian for $G$ and $D$ be the degree vector, then we have

$$Lx = i_u = D - 2m\mathbb{1}_u$$

$$\forall v \in V, \sum_{(u,v)\in E} x_v - x_u = d(v) \tag{1}$$

## 5.1 Lollipop Graph

Let us revisit the lollipop graph with the electrical network view and compute $h_{u,v}$ and $h_{v,u}$ with $u$ and $v$ as before. To compute $h_{u,v}$. Let $u'$ be the vertex common to the clique and the path. Clearly, the path has resistance $\theta(n)$. $\theta(n)$ current is injected in the path and $\theta(n^2)$ current is injected in the clique.

Consider draining current from $v$. The current in the path is $\theta(n^2)$ as $2m - 1 = \theta(n^2)$ current is drained from $v$ which enters $v$ through the path implying $x'_u - x_v = \theta(n^3)$ using Ohm's law ($V = IR$). Now consider draining current from $u$ instead. The current in the path is now $\theta(n)$ implying $x_v - x'_u = \theta(n^2)$ by the same argument.

Since the effective resistance between any edge in the clique is less than 1 and $\theta(n^2)$ current is injected, there can be only $\theta(n^2)$ voltage gap between any 2 vertices in the clique. We get $h_{u,v} = x_u - x_v = \theta(n^3)$ in the former case and $h_{v,u} = x_v - x_u = \theta(n^2)$ in the latter.

## 5.2 Proof of Relation

Define $h'_{v,u} = h_{v,u}$ when $v \neq u$ except $h'_{v,v} = 0$. By current conversion, $\forall u \neq v \in V$, we have

$$h'_{v,u} = \sum_{(v,w)\in E} \frac{1}{d(v)}(1 + h'_{w,u})$$

$$h'_{v,u} = 1 + \sum_{(v,w)\in E} \frac{1}{d(v)} h'_{w,u}$$

$$d(v) = \sum_{(v,w)\in E} h'_{v,u} - h'_{w,u} \tag{2}$$

Equations 1 and 2 are linear systems with unique solutions and are identical under $x_v - x_u = h'_{v,u}$ (up to same additive shift to each entry). $x_v = h'_{v,u}$ if $x_u = 0$.

We have shown that for $i_u = D - 2m\mathbb{1}_u$ with $x = L^+ i_u$ that $x_v - x_u = h_{v,u}$. For $u'$, we have $x' = L^+ i_{u'}$. Now, we have,

$$x - x' = L^+(i_u - i_{u'}) = 2mL^+(e_{u'} - e_u)$$

The above is equivalent to $2m$ times voltage obtained if you inject 1 ampere at $u'$ and remove 1 ampere from $u$. Using Kirchoff's law we get

$$\begin{aligned} 2mR_{u,u'} &= (x - x')_{u'} - (x - x')_u \\ &= (x_{u'} - x_u) - (x'_u - x'_{u'}) \\ &= h_{u',u} + h_{u,u'} = C_{u,u'} \end{aligned}$$

## 6 Cover Time of a Graph

We define $C_u(G)$ as the expected time for a random walk starting at $u$ to visit all vertices in a graph. $C(G)$ is the maximum of $C_u(G)$ over all $u \in V$.

We have $\forall u \in V$,

$$C_u(G) \le 2m(n-1)$$

Consider the spanning tree $T$ of graph $G$. The cover time is bounded by traversing the edges of the tree in both directions (as we could just do a DFS on the spanning tree), and hitting time gives the expected time of moving along an edge, we get

$$\begin{aligned} C_u(G) &\le \sum_{(u,v)\in E(T)} h_{u,v} + h_{v,u} \\ &= \sum_{(u,v)\in E(T)} C_{u,v} \\ &\le (n-1)\max_u C_{u,v} \\ &\le 2m(n-1) \end{aligned}$$

This above inequality is tight for lollipop ($\theta(n^3)$) but not for cliques which has $O(n \log n)$ as we can model it as a coupon collector problem.

Let $R_max = \max_{u,v\in V} R_{u,v}$. We give a tighter bound without proof on $C(G)$ as follows:

$$mR_{max} \le C(G) \lesssim mR_{max} \log n$$

## References

[MR] Rajeev Motwani, Prabhakar Raghavan Randomized Algorithms. *Cambridge University Press*, 0-521-47465-5, 1995.