| | |
|---|---|
| **6.842 Randomness and Computation** | March 5, 2020 |

# Lecture 9

*Lecturer: Ronitt Rubinfeld* *Scribe: Kuo-An Wei*

Lecture Preview: Random Walks

- Stationary Distributions
- Cover Times
- $s$-$t$ connectivity in undirected graphs (will be covered next lecture)

We first introduce some basic properties of markov chains.

**Definition 1** *A matrix $P$ is stochastic if every row sums to 1. $P$ is doubly stochastic if every row and column sum to 1.*

The transition matrix of a markov chain must be stochastic, but it does not have to doubly stochastic.

**Definition 2** *A markov chain is irreducible if for all $x, y$*

$$\exists t \text{ s.t. } P^t(x, y) > 0.$$

Irreducibility tell us that the markov chain is strongly connected in some sense.

**Definition 3** *A markov chain is ergodic if there exists $t_0$ such that for all $x, y$*

$$\forall t > t_0,\ P^t(x, y) > 0.$$

Note that ergodicity is naturally a stronger condition than irreducibility.

**Definition 4** *A markov chain is aperiodic if for all $x$*

$$\gcd\{t \mid P^t(x, x) > 0\} = 1.$$

**Theorem 5** *A markov chain is ergodic if and only if it is irreducible and aperiodic.*

We will not prove this formally, but it is not hard to see why both irreducibility and aperiodicity are necessary and sufficient conditions for ergodicity in a finite graph.

**Definition 6** *Given a stochastic matrix $P$, we say that a probability distribution $\pi$ over the nodes is stationary iff for all $y$,*

$$\pi(y) = \sum_x \pi(x) P(x, y).$$

Hence, $\pi$ is stationary if and only if $\pi P^t = \pi$ for all $t$.

For any given graph, we will assume a corresponding markov chain with uniform transition probabilities at each node, unless specified otheriwse.

Here is an example stationary distribution for an undirected graph:

$$\pi = \left(\frac{\deg(x_1)}{2|E|}, \ldots, \frac{\deg(x_n)}{2|E|}\right).$$

Also, if a markov chain has a doubly stochastic transition matrix, the uniform distribution over all nodes will be a stationary solution. Some examples include $d$-regular undirected and directed graphs.

**Theorem 7** *Every ergodic markov chain has a unique stationary distribution.*

This is a very important theorem on markov chains that we will come back to later.

**Definition 8** *We define the hitting time between nodes $i$ and $j$ in a markov chain to be*

$$h_{ij} = \text{Exp}[\textit{time starting at } i \textit{ to reach } j].$$

**Theorem 9** *Given an ergodic markov chain and its unique stationary distribution $\pi$, we have*

$$\forall i,\ h_{ii} = \frac{1}{\pi(i)}.$$

**Definition 10** *Define the cover time of a graph $G$ w.r.t. a node $i$ to be*

$$C_i(G) = \text{Exp}[\textit{time starting at } i \textit{ to reach all nodes in } G]$$

*Then, the cover time of $G$ is defined as*

$$C(G) = \max_u C_i(G).$$

For example, here are the cover times for some specific graphs

- $C(K_n^*) = \Theta(n \ln n)$
- $C(L_n^*) = \Theta(n^2)$
- $C(lollipop_n) = \Theta(n^3)$

Note that $K_n^*$ is the complete graph on $n$ vertices with self loops, $L_n^*$ is the line graph on $n$ vertices with self loops, and $lollipop_n$ is the graph on $n$ vertices where we attach $K_{n/2}$ to one end of $L_{n/2}$.

It is worthwhile to note that $lollipop_n$ gives an asymptotically worst case construction, as we shall see later. Our goal

**Definition 11** *Define the commute time between nodes $i$ and $j$ in a markov chain to be*

$$C_{ij} = \text{Exp}[\textit{time starting at } i \textit{ to reach } j \textit{ and return to } i].$$

By linearity of expectation, we have simply that $C_{ij} = h_{ij} + h_{ji}$.

**Lemma 12** *For all nodes $i, j$ where $(i, j) \in E$, we have*

$$C_{ij} = O(|E|).$$

**Proof** Without loss of generality, assume $G$ is irreducible. Otherwise, we can simply consider the irreducible component that contains nodes $i$ and $j$.

Construct the directed graph $G' = (V', E')$ from $G = (V, E)$ where $V' = V$ by replacing each undirected edge with two directed edges. In addition, for each node $i \in V$, add $\deg(i)$ self loops to the corresponding node $i \in V'$. This will guarantee the aperiodicty of $G'$, so $G'$ must be ergodic and have a unique stationary distribution.

At every node in $G'$, we now have $\frac{1}{2}$ probability of staying at the same node, and $\frac{1}{2}$ probability of following the original transition matrix of $G$. Hence, for all nodes $i$ and $j$, we have

$$C'_{ij} = 2C_{ij}.$$

Next, note that traversing the edge $(i, j)$ twice must give a commute between $i$ and $j$ in $G'$, which upper bounds the commute time in $G$. Hence, we could try taking the line graph of $G'$ and analyze the hitting time of the vertex corresponding to $(i, j)$.

In other words, construct $G'' = (V'', E'')$ where

$$V'' = E' \qquad \text{and} \qquad E'' = \{(e_1, e_2) \in E' \times E' \mid e_1 = (u, v), e_2 = (v, w) \text{ for some } u, v, w \in V'\}$$

Finally, one can check that $G''$ is doubly stochastic by construction. Hence, its stationary distribution must be the uniform distribution over all vertices. We can then bound the commute time $C_{ij}$ using Theorem 5

$$C_{ij} = \frac{1}{2} \cdot C'_{ij} \leq \frac{1}{2} \cdot h''_{(i,j),(i,j)} = O(|E|).$$

■

**Theorem 13** *For any graph $G$, we can bound its cover time by*

$$C(G) = O(mn) = O(n^3)$$

*where $n = |V|$ and $m = |E|$.*

**Proof** Without loss of generality, we can assume that $G$ is irreducible. Consider a spanning tree $T$ of $G$. For a given node $i$, consider a depth-first traversal of $T$ with $i$ as the root node, represented by the sequence of nodes $u_1, u_2, \ldots, u_{2n-2}$. Since each edge is traversed exactly twice in a depth-first traversal, we can bound the cover time of $G$ w.r.t. node $i$ using Lemma 12 by

$$C_i(G) \leq \sum_{j=1}^{2n-1} h_{u_j u_{j+1}} = \sum_{(u,v)\in T} (h_{uv} + h_{vu}) = \sum_{(u,v)\in T} C_{uv} = O(mn).$$

Since this is true for every node $i$, we have $C(G) = \max_i C_i(G) = O(mn)$, as desired. ■