6.895 Randomness and Computation March 13, 2006

Lecture 10

Lecturer: Ronitt Rubinfeld *Scribe: Krzysztof Onak*

1 Preliminaries

1.1 Markov chains

We say that a finite set Ω of states, and random variables $X_t \in \Omega$ constitute a *Markov chain* if they meet the following *Markovian property*:

$$\forall t\ \forall x_0, x_1, \ldots, x_t, y \in \Omega,\ \Pr[X_{t+1} = y | X_0 = x_0, X_1 = x_1, \ldots, X_t = x_t] = \Pr[X_{t+1} = y | X_t = x_t].$$

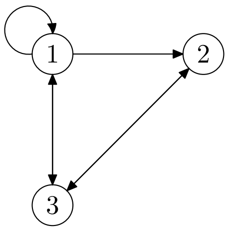
Without lost of generality we may assume that transitions are independent of time, and under this assumption we have the following *transition matrix* P:

$$P(x, y) = \Pr[X_{t+1} = y | X_t = x].$$

1.2 Random walks on graphs

A *random walk* on $G = (V, E)$ is a sequence s_0, s_1, ... of nodes where s_0 is a *start node*, and at each i, s_{i+1} is picked uniformly from $N(s_i)$, the neighborhood of s_i. If $d(i)$ is the number of neighbors of i, then

$$P(i,j) = \begin{cases} \frac{1}{d(i)}, & \text{if } (i,j) \in E; \\ 0, & \text{otherwise.} \end{cases}$$



$$P:\ \begin{array}{c|ccc} & 1 & 2 & 3 \\ \hline 1 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ 2 & 0 & 0 & 1 \\ 3 & \frac{1}{2} & \frac{1}{2} & 0 \end{array}$$

Random walks on graphs are special case of Markov chains.

1.3 More useful terminology

We say that a matrix is:

- *stochastic* if all rows sum to 1,
- *doubly stochastic* if all rows and all columns sum to 1.

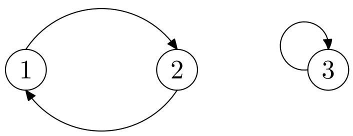
One may also find useful a *t-step distribution*, that is a distribution that we achieve moving t time steps into the future:

$$P^t(x, y) = \begin{cases} P(x, y), & \text{if } t = 1; \\ \sum_{z \in \Omega} P(x, z) P^{t-1}(z, y), & \text{if } t > 1. \end{cases}$$

As one can see the matrix of this transition really equals the t-th power of P.

We also consider distributions of states, and therefore, if π_i is a distribution at moment i, then $\pi_t = \pi_0 P^t$. We will be interested especially in stationary distributions. A distribution $\widetilde{\pi}$ is a *stationary distribution* if $\widetilde{\pi}(y) = \widetilde{\pi}(x) P(x, y)$.

Example It is not true that any initial distribution converges to a stationary distribution, nor is it true that a stationary distribution is unique. Consider the following graph:



The distribution $(1, 0, 0)$ becomes in next steps consecutively $(0, 1, 0)$, $(1, 0, 0)$, $(0, 1, 0)$, and so forth. Hence it doesn't converge. On the other hand, there are two different stationary distributions: $(0, 0, 1)$, and $(\frac{1}{2}, \frac{1}{2}, 0)$.

2 Ergodic Markov chains

Say a Markov chain is *ergodic* if

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

Theorem 1 *A Markov chain is ergodic if and only if it is*

- irreducible *(or* connected*), i.e.*

$$\forall x, y \ \exists t = t(x, y) : \ \ P^t(x, y) > 0,$$

- *and* aperiodic*, i.e.*

$$\forall x \in \Omega, \ \ \gcd\{t : P^t(x, x) > 0\} = 1.$$

Theorem 2 *Any strongly connected aperiodic Markov chain with a doubly stochastic transition matrix is ergodic*

Theorem 3 (important!) *Let M be an ergodic Markov chain. M has a unique stationary distribution $\widetilde{\pi}$. The expected time for a random walk starting at i to return to i is $1/\widetilde{\pi}_i$.*

3 Random walks on graphs

A stationary distribution of a random walk on an undirected graph is

$$\widetilde{\pi} = \left(\frac{\deg(x_1)}{2m}, \frac{\deg(x_2)}{2m}, \dots \right).$$

3.1 Cover time of a graph

Let $\mathcal{C}_u(G)$ be the expected time to visit each node in graph G, starting at vertex u, and let

$$\mathcal{C}(G) = \max_u \mathcal{C}_u(G).$$

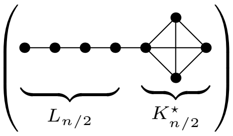
Examples Let $K_n^\star$ be the complete graph on n nodes with self loops. We have

$$\mathcal{C}(K_n^\star) = \Theta(n \log n)$$

by the coupon collector's problem. If L_n is a line on n nodes, then

$$\mathcal{C}(L_n) = \Theta(n^2).$$

Furthermore,

$\mathcal{C}(\text{lollipop}) = \mathcal{C}\Bigg($  $\Bigg) = \Theta(n^3).$

3.2 Expected numbers of steps

Let h_{ij} be the expected number of steps to reach j, starting at i, and c_{ij} be the expected number of steps to reach j, and return to i, starting at i. Value h_{ij} is called a "hitting time", and c_{ij} a "commute time". By linearity of expectation we have

$$c_{ij} = h_{ij} + h_{ji},$$

and therefore,

$$c_{ij} = c_{ji}.$$

Lemma 4 $\forall (u,v) \in E[c_{uv}] \leq 2m$

Proof If one traverses (u,v) twice, then commutes $u, v \to u$, and therefore, it suffices to show that

$$E[\text{time between visits to edge } (v,u)] \leq 2(m+1),$$

since we can assume that in the beginning we go from u to v, and ask how much time it takes to cross (u,v) once again.

Assume graph G has a self-loop.

Consider graph $G' = (V', E')$ such that V' is the set of directed edges from E (i.e. we distinguish (a,b) and (b,a) unless $a = b$), and

$$E' = \{((u,v)(v,w)) \,|\, (u,v) \in E, (v,w) \in E\}.$$

$$P: \begin{array}{c|c|c|} & 1 & 2 \\ \hline 1 & \frac{1}{2} & \frac{1}{2} \\ \hline 2 & 1 & 0 \\ \hline \end{array} \qquad Q: \begin{array}{c|c|c|c|} & (1,2) & (1,1) & (2,1) \\ \hline (1,2) & 0 & 0 & 1 \\ \hline (1,1) & \frac{1}{2} & \frac{1}{2} & 0 \\ \hline (2,1) & \frac{1}{2} & \frac{1}{2} & 0 \\ \hline \end{array}$$

In general, if Q is the transition matrix for G', then

$$Q((u,v),(v,w)) = P(v,w) = \frac{1}{d(v)}, \quad \text{if } (v,w) \in E,$$

and summing over columns,

$$\forall (u,v) \in E = V' \qquad \sum_{\substack{(u,v) \text{ s.t.} \\ ((u,v),(v,w)) \in E'}} Q_{(u,v)(v,w)} = \sum_{(u,v) \in E} \frac{1}{d(v)} = 1.$$

Hence Q is doubly stochastic. It is easy to see that G' is strongly connected. As G had a self-loop at some node v, (v,v) in G' also has a self-loop. Therefore, G' is egodic and has the following stationary distribution:

$$\widetilde{\pi}(x,y) = \frac{1}{2|E|}, \qquad \forall (x,y) \in E.$$

Hence,

$$E[\text{time starting at } (u,v) \text{ to reach } (u,v) \text{ again}] = 2|E|.$$

If G does not have a self-loop, then add one to get a new graph G''. With the exception of the node with a new self-loop, the transition probabilities are the same for all nodes in G'' and G. For that node, there is a chance of not moving for a cycle. Given that we move to a new node, the transition probabilities are the same for that node as well. Therefore we have:

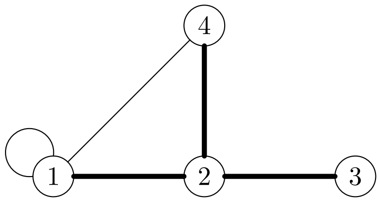
$$E[\text{time between visits to edge } (v,u) \in G] \leq E[\text{time between visits to edge } (v'',u'') \in G''] = 2(|E|+1)$$

■

Lemma 5 *For an undirected graph G,*

$$\mathcal{C}(G) \le 2(m+1)(n-1) < n^3.$$

Proof We start at an arbitrary vertex v_0, and let T be a spanning tree of G rooted at v_0. T has $n-1$ edges. Let $v_0, v_1, \ldots, v_{2n-2} = v_0$ be a sequence of vertices in the depth first traversal of T.



$$v_0 = 1 \to 2 \to 3 \to 2 \to 4 \to 2 \to 1$$

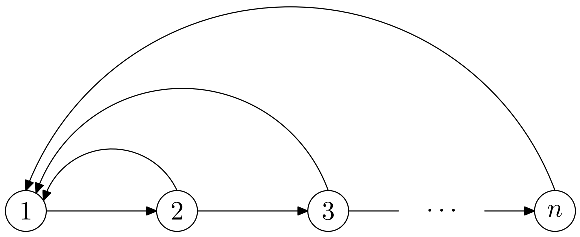
We visit each edge (a,b) exactly once in each direction. Therefore,

$$\mathcal{C}(G) \le \sum_{j=0}^{2n-3} h_{v_j v_{j+1}} = \sum_{(u,v)\in T} c_{uv} \le \sum_{(u,v)\in T} 2m = 2m(n-1).$$

■

3.3 Directed graphs

The following graph shows that many of the results that we have seen for undirected graphs do not hold for directed graphs.



To get from vertex 1 to vertex n we need on average a number of steps exponential in n.

4 Next time

Next time, we will see an application of random walks to the *st*-connectivity problem.