Randomness in Computing

LECTURE 25

CS
537

Last time

- Randomized algorithm for 3SAT
- Gambler’s ruin
- Classification of Markov chains
- Stationary distributions

Today

- Random walks on graphs
- Algorithm for s-t-PATH

CS 537

Classification of Markov chains

- We want to study Markov chains that ``mix'' well.
- We will define Markov chains that avoid some problematic behaviors: irreducible and aperiodic.
- A finite Markov chain is irreducible if its graph representation consists of one strongly connected component.

Periodicity

- Example: a Markov chain whose states are integers and it moves to each neighboring state with probability ½.

If the chain starts at 0, when can it be in an even-numbered state?

- A state is periodic if there exists an integer $\Delta > 1$ such that $\Pr[X_{t+s} = j \mid X_t = j] = 0$ unless s is divisible by Δ ; otherwise, it is aperiodic.
- A Markov chain is aperiodic if *all* its states are aperiodic.

Fundamental theorem

- A stationary distribution of a Markov chain is a probability distribution $\bar{\pi}$ such that $\bar{\pi} = \bar{\pi}\boldsymbol{P}$.

(Describes steady state behavior of a Markov chain.)

Fundamental Theorem of Markov Chains (selected items)

Every finite, irreducible and aperiodic Markov chain satisfies the following:

1. There is a unique stationary distribution $\bar{\pi} = (\pi_0, \pi_1, \dots, \pi_n)$, where $\pi_i > 0$ for all $i \in \{0,1,\dots,n\}$.
2. For all $i \in \{0,1,\dots,n\}$, the hitting time $h_{ii} = 1/\pi_i$.

- The hitting time from u to v, denoted $h_{u,v}$, is the expected time to reach state v from state u.

CS 537

Random walks on undirected graphs

Given a connected, undirected graph $G = (V, E)$,
define the following Markov chain

- states = vertices of the graph
- from each state v, the chain moves to a uniformly random neighbor of v

$$P_{uv} = \begin{cases} \dfrac{1}{d(u)} & \text{if } (u, v) \in E \\ 0 & \text{otherwise} \end{cases}$$

- Observation: This Markov chain is aperiodic iff G isn't bipartite.

Stationary distribution

- Assume G is not bipartite.

Theorem

A random walk on G has stationary distribution $\bar{\pi}$, where, for all nodes v,

$$\pi_v = \frac{d(v)}{2|E|}$$

Proof: 1. Check probabilities sum to 1:

- Since $\sum_{v \in V} d(v) = 2|E|$,

$$\sum_{v \in V} \pi_v = \sum_{v \in V} \frac{d(v)}{2|E|} = 1$$

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

2. Check that $\bar{\pi} = \bar{\pi} \cdot P$

Hitting time, commute time, cover time

- The hitting time from u to v, denoted $h_{u,v}$, is the expected time to reach state v from state u.

- The commute time between u and v is $h_{u,v} + h_{v,u}$.

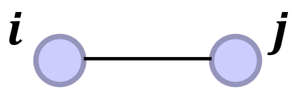
- The cover time of a graph $G = (V, E)$ is the maximum over $v \in V$ of the expected time for a random walk starting at v to visit all nodes in V.

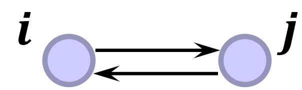
Bound on commute time

Commute Time Lemma

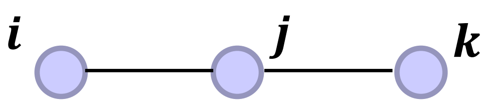
If $(u, v) \in E$, the commute time $h_{u,v} + h_{v,u}$ is at most $2|E|$.

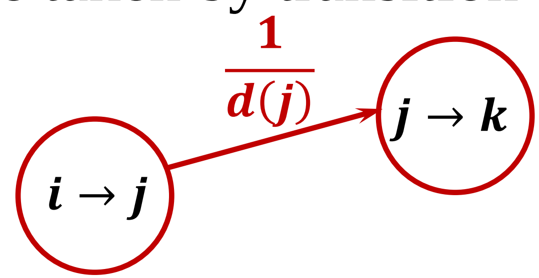
Proof: Let $D =$ set of $2|E|$ directed edges $\{\boldsymbol{i \to j} \mid \{i, j\} \in E\}$





- Random walk on G corresponds to Markov chain with states D, where state at time t is the directed edge taken by transition t.





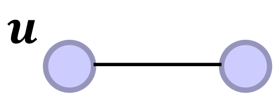
- This Markov Chain has uniform stationary distribution.

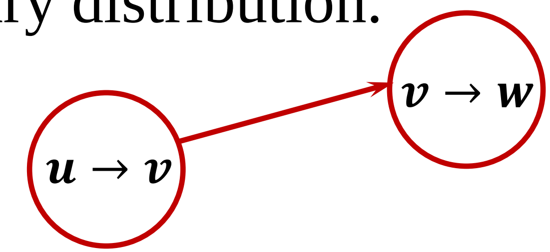
Bound on commute time

Commute Time Lemma

If $(u, v) \in E$, the commute time $h_{u,v} + h_{v,u}$ is at most $2|E|$.

Proof: This Markov Chain has uniform stationary distribution.





- By Fundamental Thm of Markov Chains,

$$h_{u\to v,u\to v} = \frac{1}{\pi_{u\to v}} = 2|E|$$

$=$ expected time to traverse $u \to v$ starting at $u \to v$

$=$ expected time to go from v to u and then traverse (u, v)

- But this is only one way to go from v to u to v:

$$h_{v,u} + h_{u,v} \le 2|E|$$

Bound on cover time

Cover Time Lemma

The cover time of G with n nodes and m edges is at most $2m(n-1)$.

Proof: Choose a spanning tree T of G and a starting vertex v_0.

- Consider the directed tour starting at v_0 that traverses each edge of T once in each direction

  (by visiting nodes in the order of DFS if T from v_0)
- Let $v_0, v_1, \dots, v_{2n-3}$ be the sequence of nodes (with repetitions) in the order visited by the tour

$\mathbb{E}$[# of steps to cover V starting from v_0]

$\leq \mathbb{E}$[# of steps, starting from v_0, to visit $v_1, \dots, v_{2n-3}$ in that order]

by linearity of expectation

$$= \sum_{i=0}^{n-3} h_{v_{i,} v_{i+1}}$$

Application: s-t-PATH

Problem: Given an undirected graph G with n nodes and m edges and two nodes, s and t, determine if G contains a path from s to t.

- Can be solved by BFS in $O(m + n)$ time
- This approach requires $\Omega(n)$ space.
- Today: a randomized algorithm that uses $O(\log n)$ space.

  Less space than it takes to store a path!

Algorithm for s-t-PATH

1. Start a random walk from s.
2. If the walk reaches t in $2n^3$ steps, **accept**; otherwise, **reject**.

Correctness of s-t-PATH algorithm

Theorem

The algorithm uses $O(\log n)$ bits and has error probability $\leq 1/2$.

Proof: If there is no path, the algorithm correctly rejects.

- Suppose there is an s-t path.
- The expected time to reach t from s is at most the expected cover time of the connected component, which is, by Cover Lemma is
$$\leq 2mn \leq n^3.$$
- By Markov's inequality, the probability that the walk takes more than $2n^3$ steps to reach t is at most 1/2.

Space analysis: Need to keep

- current position: $O(\log n)$ bits
- counter for the number of steps: $O(\log n)$ bits

What do we change for bipartite graphs?