Today

Random walks
Stationary Distributions
Cover Times
UST Conn
Markov chains:

\( \Omega = \text{set of "states" (or nodes)} \) (here always \text{FINITE})

\( X_0, \ldots, X_t \in \Omega \) sequence of visited states

Markovian property:

\[
\Pr[X_{t+1} = y \mid X_0 = x_0, X_1 = x_1, \ldots, X_t = x_t] = \Pr[X_{t+1} = y \mid X_t = x_t]
\]

Wlog, assume transitions independent of time:

i.e. \( P(x,y) = \Pr[X_{t+1} = y \mid X_t = x] \)

so can use "transition matrix" to represent it

Important special case:

transitions uniform on subset corresponding to neighbors of node

**Def. random walk on** \( G=(\Omega,E) \)

is a sequence \( S_0, S_1, \ldots \) of nodes

where \( S_0 \) is a start node.

At each step \( i \), \( S_{i+1} \) picked uniformly from \( N(S_i) \) (outedges)
**Examples**

**Markov chain**

![Diagram of a Markov chain with states 1, 2, 3 and transition probabilities](image)

**Random walk on digraph**

![Diagram of a random walk on a digraph with states 1, 2, 3](image)

---

\[
\begin{align*}
\text{Out edges of node } i & = d_i \\
p(i, j) & = \frac{1}{d_i} \quad \text{if } (i, j) \in E \\
\text{otherwise} & = 0 \\
\end{align*}
\]

\[
\forall i \sum_j p(i, j) = 1
\]
Distributions after $t$ steps

Transition probabilities for $t$ steps: $P^t(x,y) = \begin{cases} P(x,y) & t=1 \\ \sum_{z} p(z|x) P^{t-1}(z,y) & t>1 \end{cases}$ matrix multiplication

Initial distribution: $\Pi^{(0)} = (\Pi_1^{(0)}, \ldots, \Pi_n^{(0)})$ where $\Pi_i^{(0)} = \Pr$[start at node $i$]

distribution after one step:

$\Pi^{(1)} = \Pi^{(0)} \cdot P = (\sum_{z} P(z|1) \cdot \Pi(z), \ldots, \sum_{z} P(z|x) \cdot \Pi(z), \ldots)$

$t$-step distribution: $\Pi^{(t)} = \Pi^{(0)} P^t$

---

Finite Markov Chain Properties

Stochastic matrix: rows of $P$ sum to 1

doubly stochastic matrix: rows $\&$ columns sum to 1

e.g. random walk on undirected graph

or digraph in which

indegree = outdegree = const for all nodes

irreducible: ("strongly connected")

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

ergodic: change of quantifier order

$\exists t_0$ st. $\forall t > t_0 \forall x, y \ P^t(x,y) > 0$ — stronger than irreducible! why?
Aperiodic:
\[ \forall x \quad \gcd \exists t : p^t(x,x) > 0 \implies \gcd \text{ of possible cycle length} = 1 \]

**Thm**
Ergodic \iff Irreducible + Aperiodic

---

**Stationary Distributions**

Does it depend on \( \Pi_0 \)?

\[ \exists \text{ stationary distribution} \quad \Pi \quad \quad \exists \quad \Pi(\cdot) = \sum_x \Pi(x) P(x, \cdot) \]

\[ \forall y \quad \Pi(y) = \sum_x \Pi(x) P(x, y) \]

Will consider if \( \Pi \) is unique \& exists \( \exists \) if, \( \Pi(\cdot) = \Pi(\cdot - 1) \)

- if periodic: could have no stat. dist. or several
- if reducible: could have lots of stat. dist.

Some stat. dists:
\[ \left( \frac{1}{2}, \frac{1}{2}, 0, 1 \right), \left( 0, 1, 1, 0 \right) \ldots \]

**Important Thm**
Every ergodic M.C. has unique stationary distribution
Stationary dist. of undirected graph:

\[ \pi = \left( \frac{\text{deg}(x_1)}{2\text{vol}}, \frac{\text{deg}(x_2)}{2\text{vol}}, \ldots \right) \]

- So d-regular graphs have \( \pi \) uniform.

(also \( \text{indegree} = \text{outdegree} = d \) digraphs

\( ) \) doubly stochastic P.M.C.'s \( \) this implies the others!

- Not true in general for digraphs

- Bipartite, periodic graphs may have other stat. dists.

Hitting time

\[ h_{ii} = E \left[ \text{time starting at } i \text{ to return to } i \right] \]

\[ = \frac{1}{\pi_i} \quad \text{Very useful theorem!} \]

\[ h_{ij} = E \left[ \text{time starting at } i \text{ to reach } j \right] \]

Cover time of undirected graph

\[ C_u(G) = E \left[ \# \text{ steps to reach all nodes in } G \text{ on walk starting from } u \right] \]

\[ C(G) = \max_u C_u(G) \]
Graph Theory Examples:

- $C^\circ(k_n)$ where $k_n$ = complete graph with self-loops at each node
  
  $= \Theta(n \ln n)$ by coupon collector argument

- $C(L_n)$ where $L_n$ = $n$ node line
  
  $= \Theta(n^2)$

- $C(lollipop)$

  $= \Theta(n^3)$

**Thm:** $C^\circ(G) \leq 8m(n-1)$

**Proof:**

First - transform $G$ into $G'$ (see example on pg 8) to make $G$ aperiodic, add self-loops to each $u$ (i.e. take self-loop with prob $\frac{1}{2}$).

Claim: $C^\circ(G') = 2 C(G)$

Why? can transform paths in $G'$ by removing self-loops, expected # self-loops $= \frac{1}{2} \cdot$ (length of path)

So that we have unique stationary dist.
Next, commute time + a lemma:

**Definition:** $C_{ij} = E[\#\text{steps for r.w. starting at } i \text{ to hit } j \text{ and return to } i]$

**Claim:** $C_{ij} = h_{ij} + h_{ji}$ (linearity of expectation)

**Lemma:** $\forall (u,v) \in E \Rightarrow C_{uv} \in O(m)$

**Proof of Lemma**:

Key idea: (actually will show $C_{uv} \in O(m)$, but it's symmetric)

if traverse $(u,v)$ twice

have performed commute from $v \rightarrow u \rightarrow v$

Plan: show $E[\text{time between visits to } (u,v)]$ is $O(m)$

$\Rightarrow C_{uv}$ is $O(m)$

Given $G' = (V,E)$ (G with added self-loops)

Construct "G" representing walks on directed edges of $G'$

Line graph $E \rightarrow V$

$(u,v)(v,w) \rightarrow E''$

new nodes $\{v\}$ are edges $(u,v)$ in $G'$

new edges are length 2 paths in $G'$

Visit edge in $G'$ twice $\Rightarrow$ visit node in $G''$ twice
Note: $G''$ is doubly stochastic:

$$Q_{(uv) (vw)} = P_{vw} = \frac{1}{d(v)} \quad \text{if} \quad (uv) (vw) \in E$$

$$\forall (vw) \in E \quad \sum_{(uv) \in \text{E}'} Q_{(uv) (vw)} = \sum_{(uv) \in \text{E}} \frac{1}{d(v)} = 1$$

Column sum

:. $\Pi$ of $G''$ is uniform

$$\Pi = \frac{1}{|V''|} = \frac{1}{q_m}$$

\[ h_{u,v} = \frac{1}{\Pi_u} = q_m \quad \text{for all nodes } u \text{ in } G'' \]

\[ (u,v) \text{ in } G' \]

if start at $v$ conditioned on coming from edge $(uv)$

expect $\leq 4m$ steps to see $(uv)$ again.

But, it's an M.G. so conditioning doesn't affect.

$\Rightarrow$ if start at $v$, expect to see $(uv)$ in $\leq 4m$ steps.

$\Rightarrow C(u,v) = C(v,u) = O(m)$

$$\text{Note: valid only for } (uv) \in E$$
Lemma \( C(G) = O(nm) = O(n^3) \)

**Pf.**

1. Start vertex \( v_0 \)
2. \( T \) is a spanning tree rooted at \( v_0 \)
3. \( T = n-1 \) edges
4. \( v_0 \rightarrow v_1 \rightarrow \cdots \rightarrow v_{2n-2} \rightarrow v_0 \) is a depth-first traversal
   - St. each edge appears twice, once in each direction
   - \((a,b) + (b,a)\)

\[
C(G) \leq \sum_{j=1}^{2n-3} h_{v_j v_{j+1}}
\]

\[
= \sum_{(u,v) \in T} C_{uv}
\]

\[
= O\left( \sum_{(u,v) \in T} m \right)
\]

\[
= O(nm)
\]
$S-T$ connectivity (UST-Conn)

Input: Undirected $G$, nodes $s,t$
Output: "Yes" if $s,t$ connected
"No" o.w.

Can solve in poly time, in many ways.
What about small space?

$RL$ = class of problems solvable by randomized log-space computations

[no change for input space (read only), but can only have const #ptrs ...]

Thm  $UST$-Conn $\in$ $RL$

Algorithm:

- start at $s$
- take random walk for $\Theta(n^3)$ steps
- if ever see $t$, output "Yes"
- o.w. output "No"

Complexity:
- Keep track of $\#$ steps so far
- $\#$ edges at each node + toss coin to pick one randomly
Behavior:

If \( s,t \) not connected, never output "yes"

If \( s,t \) connected

\[
1_{st} \leq C_s(G_s) \leq n^3
\]

\( \uparrow \)

connected component of \( s \)

\[
\Pr \left[ \text{output "no"} \right] = \Pr \left[ \text{start at } s, \text{ walk } \geq c \cdot \mathbb{E}[C_s(G_s)] \text{ steps} \right.
\]

\[+ \text{ don't see } t ] \]

\[\leq \frac{1}{c} \quad \text{by Markov's } \]

Comments

• Actually \( \text{USTCONN} \in L \) !!!

• Open is \( RL=L \)?

we know \( RL \in L^{3/2} \)