6.842 Lec 10

- Markov Chains
  - Random walks
  - Stationary Dist.
  - Cover Times
Markov Chain

set of states: \( \Omega \)

\( x_1 \ldots x_t \in \Omega^t \): sequence of visited states

Markovian Property:

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

Only current state matters NOT how we get there

Transitions independent of time

def: \( P(x,y) = P [X_{t+1} = y \mid X_t = x] \)

Represent w/ “transition matrix”
Example

Important special case:

Transition to uniformly random neighbor

**def**: Random Walk on $G = (V, E)$ is a sequence $S_0, S_1, \ldots$ of nodes $S_i$, chosen uniformly from $N(S_i)$ out edges.

Let $d_v = \# \text{ out edges of } v$

$$P(x, y) = \begin{cases} \frac{1}{d_x} & \text{if } (x, y) \in E \\ 0 & \text{o.w.} \end{cases}$$
\[ P^t(x, y) = \begin{cases} P(x, y) & \text{if } t = 1 \\ \sum_{z} P(x, z) P^{t-1}(z, y) & \text{if } t > 1 \end{cases} \]
Initial dist. \( \pi^{(0)} = \pi_1^{(0)} \pi_2^{(0)} \ldots \)

\[ \pi^{(0)} \xrightarrow{\text{one step}} \pi^{(1)} \xrightarrow{\text{one step}} \pi^{(2)} \ldots \]

\[ = \pi^{(0)} \rho \]

\[ = \pi^{(1)} \rho \]

\[ = \pi^{(2)} \rho^2 \]

\[ = \pi^{(0)} \rho^t \]

\[ t\text{-step distribution:} \quad \pi^{(0)} \rho^t \]

**Does this converge?**

**Properties**

Irreducible \(\Rightarrow\) (strongly connected)

\[ \forall x, y \exists t(x, y) \text{ s.t. } p^t(x, y) > 0 \]

Aperiodic: \[ \forall x \text{ gcd } \{ t : p^t(x, x) > 0 \} = 1 \]

(\( \text{gcd} \) of possible cycle lengths = 1)

Ergodic: \[ \exists t^* \text{ s.t. } \forall t > t^* \quad p^t(x, y) > 0 \]

\[ \text{Ergodic } \iff \text{Irreducible } + \text{Aperiodic} \]
Stationary Distribution

\[ \Pi \text{ s.t. } \forall x \quad \Pi(x) = \sum_y \Pi(y) P(y, x) \]

or \( \Pi = \Pi P \)

(consider \( P \) s.t. \( \Pi^* \) exists and unique)

\( i.e \) does not depend on \( \Pi^{(0)} \)

Periodic \( \circ \circ \circ \) Reducible \( \circ \circ \circ \)

\( (0,1) \rightarrow (1,0) \rightarrow (0,1) \rightarrow \cdots \rightarrow (1/2) \rightarrow (0,1) \rightarrow (1,0) \)

Thm: Ergodic M.C. \( \Rightarrow \) Unique \( \Pi^* \)

Undirected Graph \( G_i = (V, E) \)

\[ \Pi^* = \left( \frac{d_{v_1}}{2|E|}, \frac{d_{v_2}}{2|E|}, \ldots \right) \]

- \( \Pi^* \) uniform for \( d \)-reg graphs
  - Also for digraphs when \( \text{indeg} = \text{outdeg} = d \)
- Not true for general digraphs
Hitting Time

\[ \text{def: } h_{xy} = \mathbb{E}[\text{#steps to go } x \longrightarrow y] \]

\[ h_{xx} : \text{Recurrence time} \]

\[ \text{Thm: } h_{xx} = \frac{1}{\Pi^*(x)} \]

\[ \text{Pf: Consider a very long walk} \]

\[ \Pi^*(x) \text{ fraction of the positions are } x \]

\[ \Rightarrow \text{Average gap between occurrences} \]

\[ h_{xx} = \frac{1}{\Pi^*(x)} \]

Cover Time

\[ C_v(G) = \mathbb{E}[\text{#steps to visit all nodes in } G \text{ starting at } v] \]

\[ C(G) = \max_v C_v(G) \]
Cover Time Examples

- $C(K_n) = K_n$ is the complete graph on $n$ vertices
  $= \Theta(n \log n)$ w/self loops at each node

- $C(L_n) = L_n$ is the line graph w/self loops at each node
  $= \Theta(n^2)$

- $C$ (lollipop) $= \Theta(n^3)$

**Thm:**
$C(G) \leq O(m \log n)$

**Pf.**
$G \xrightarrow{add \ self \ loops \ (prob \ \frac{1}{2})} G'$

Worst start
$\Theta(n)$ time to reach
$\Theta(n^2)$ time to reach
$\Theta(n^3)$ time to reach
Claim: \[ C(G') = 2C(G) \]

Path in \( G' \) \( \xrightarrow{\text{remove self loops}} \) path in \( G \)

\[ \mathbb{E}[\# \text{ self loops}] = \frac{1}{2} \cdot \text{length of path} \]

Since \( G' \) is ergodic, it has an unique stationary distribution

Commute Time

\[ \text{def } \ C_{xy} = \mathbb{E}[\# \text{ steps for } x \xrightarrow{\cdot} y \xrightarrow{\cdot} x] \]

\[ = h_{xy} + h_{yx} \] (linearity of expectation)

Lemma: \( \forall (x, y) \in E \quad C_{xy} \leq O(m) \)

pf: Consider a long walk \( u_1, u_2, u_3, \ldots \)
where \( u_i \in V \) and \( (u_i, u_{i+1}) \in E \forall i \)
We look for commutes of the following form $x \rightarrow y \rightarrow o \rightarrow o \rightarrow x \rightarrow y$ commute

Prob of finding $(x, y)$

\[
\begin{align*}
\mathbb{P}[(u_i, u_{i+1}) = (x, y)] &= \mathbb{P}[u_i = x] \cdot \mathbb{P}[u_{i+1} = y | u_i = x] \\
&= \prod_i^*(x) \cdot \frac{1}{d_x} \\
&= \frac{dx}{2m} \cdot \frac{1}{d_x} = \frac{1}{2m}
\end{align*}
\]

So, expected gap between consecutive occurrences of $x-y$ is $2m$

\[C_{xy} \leq O(m)\]
Finally, consider \( T \subseteq G' \) where \( T \) is a spanning tree \((n-1 \text{ edges})\)

\[
V_0 \ V_1 \ V_2 \ 0 \ 0 \ 0 \ \ V_{2n-2}
\]

**DFS traversal of \( T \)**

\[\Rightarrow 1 \ 2 \ 1 \ 3 \ 4 \ 1 \ 5 \ 3 \ 1\]

Each edge \((u, v)\) appears twice, as \((u, v)\) & \((v, u)\)

Using the DFS traversal sequence

\[
C(a) \leq \sum_{j=0}^{2n-3} h_{V_j V_{j+1}}
\]

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

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