# Markov Chains and Random Walks

## 1 Markov Chains

Some definitions

$S = \{1, \ldots, n\}$: Set of states

$X_t$: Random variable: State of the system in step $t$. $X_0$ is choosen according to some probability distribution

P: matrix of transition probabilities. $\mathbb{P}[X_{t+1} = j \mid X_t = i] = P_{ij}$

$\mathbf{q}^{(t)} = (q_1^{(t)}, \ldots, q_n^{(t)})$: distribution of $X_t$

Aperiodic: $\forall T > 1, \mathbf{q}^{(0)}, a, i : \exists t : q_i^{(t)} > 0 \wedge t \bmod T \neq a$

Underlying graph: $(S, \{(i,j) : P_{ij} > 0\})$



Simple Observations:

A Markov chain describes a memoryless system.

$\mathbf{q}^{(t+1)} = \mathbf{q}^{(t)}\mathsf{P}$

Key Theorem

**Theorem 1.** *If a finite Markov chain is strongly connected and aperiodic then* $\exists$ *unique stationary distribution* $\pi$ *with* $\pi\mathsf{P} = \pi$ *and* $\forall i \in S : \pi_i > 0$ *that is achieved from any initial distribution.*



### 1.1 Applications

(ca. 90 000 Google hits ($10\times$ more than “randomized algorithm”), 24000 $\times$ ps or pdf)

- Queueing Theory. For example, consider the queue length in parallel disk buffered writing.
- Models of complex systems. For example Speech recognition.
- Understanding Simulated Annealing
- ...



## 2 Random Walks on Graphs

Let $G = (V = \{1, \ldots, n\}, E)$ denote a nonbipartite, connected, undirected graph, $m = |E|$.

A random walk on $G$ is the Markov chain with $P_{ij} = 1/d(i)$ for $(i,j) \in E$ and $P_{ij} = 0$ otherwise.

**Lemma 2.** $\pi_i = d(i)/2m$ *defines a unique stationary distribution.*

**Proof:** $j$-th component of $\pi\mathsf{P}$ is

$$\sum_{i=1}^{n} \pi_i P_{ij} = \sum_{i \in (i,j) \in E} \frac{1}{2m} = \frac{d(j)}{2m}$$

$G$ is aperiodic because it is nonbipartite

$\pi$ is unique by Theorem 1.

## 2.1 Some Definitions

Hitting Time

$h_{uv}$ := expected number of steps in a RW starting at $u$ and first reaching $v$.

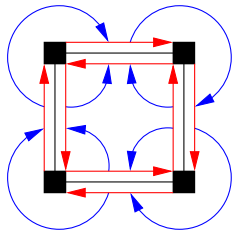
Commute time $C_{uv} = C_{vu} := h_{uv} + h_{vu}$

$\mathcal{C}_u$ = expected length of a walk starting at $u$ and visiting all $v \in V$.

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



## 2.2 Traversing Edges



**Lemma 3.** $\forall (u,v) \in E : C_{uv} \leq 2m$

*Proof.* Consider $2m$ state Markov chain $Q$ with
$Q_{(u,v),(v,w)} = P_{vw} = 1/d(v)$
that keeps track of visits of edges.

$Q$ is strongly connected and aperiodic.
The uniform distribution $\pi_{(u,v)} = 1/2m$ is stationary for $Q$:

$$(\pi Q)_{(u,v)} = \sum_{w:(v,w)\in E} \frac{1}{2md(v)} = \frac{1}{2m}$$

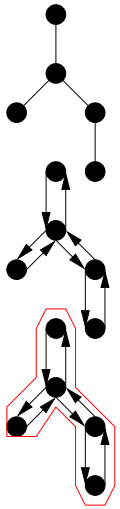
Hence, the expected time between traversals of $(u,v)$ is $2m$.

Expected time for the cycle $u \xrightarrow{(u,v)} v \rightsquigarrow u$ is also $2m$.

□



## 2.3 Cover Time



**Theorem 4.** $\mathcal{C}(G) \leq 2m(n-1)$

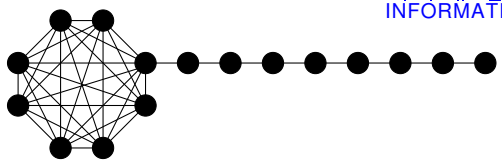
*Proof.* $T$ := spanning tree ($n-1$ edges)
$T'$ :=replace $\{u,v\}$ in $T$ by $(u,v),(v,u)$
find an Euler Tour $v_0, v_1, \ldots, v_{2n-2} = v_0$ in $T'$
Consider a random walk $v_0 \rightsquigarrow v_1 \rightsquigarrow \cdots \rightsquigarrow v_{2n-2}$.

$$\mathcal{C}_{v_0} \leq \sum_{j=0}^{2n-3} h_{v_j,v_{j+1}} = \sum_{(u,w)\in T} C_{uw} \leq 2m(n-1)$$

by Lemma 3 □



## 2.4 Lollipop



$E = \{\{i,j\} : 1 \leq i < j \leq n\} \cup \{\{i, i+i\} : n \leq i < 2n\}$

$h_{1,2n} = \Omega(n^3)$.

Proof outline:
$\approx n^2$ expected steps to reach node $n+1$ from the clique.

$\mathbb{P}[\approx] 1 - 1/n$ to fall back to the clique before reaching node $2n$.
"ruin probability"

[Feller: "Introduction to Probability Theory and its Applications, Vol. 1"]