COMPSCI 614: Randomized Algorithms with Applications to Data Science

Prof. Cameron Musco

University of Massachusetts Amherst. Spring 2024.

Lecture 19

Logistics

- I will send responses to project progress reports soon.

Summary

Last Week: Start on Markov Chains.

- Start on Markov chains and their analysis
- Markov chain based algorithms for satisfiability: $\approx n^2$ time for 2-SAT, and $\approx (4/3)^n$ for 3-SAT.

↳ improves on naive 2^n time algorithm

Summary

Last Week: Start on Markov Chains.

- Start on Markov chains and their analysis
- Markov chain based algorithms for satisfiability: $\approx n^2$ time for 2-SAT, and $\approx (4/3)^n$ for 3-SAT.

Today: Markov Chains Continued

- The gambler's ruin problem.
- Aperiodicity and stationary distribution of a Markov chain.
- The fundamental theorem of Markov chains.

Markov Chain Review

- A discrete time stochastic process is a Markov chain if is it memoryless:

$$\Pr(X_t = a_t | X_{t-1} = a_{t-1}, \ldots, X_0 = a_0) = \Pr(X_t = a_t | X_{t-1} = a_{t-1})$$

- If each X_t can take m possible values, the Markov chain is specified by the transition matrix $P \in [0,1]^{m \times m}$ with

$$P_{i,j} = \Pr(X_{t+1} = j | X_t = i).$$

- Let $q_t \in [0,1]^{1 \times m}$ be the distribution of X_i. Then $q_{t+1} = q_t P$.

q_1

.5	.5	0	0

P

.5	.5	0	0
.25	.25	.5	0
0	1	0	0
0	.5	.5	0

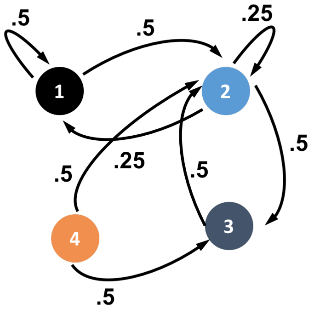
=

q_2

.375	.375	.25	0

Markov Chain Review

Often viewed as an underlying state transition graph. Nodes correspond to possible values that each X_t can take.



P

.5	.5	0	0
.25	.25	.5	0
0	1	0	0
0	.5	.5	0

The Markov chain is irreducible if the underlying graph consists of single strongly connected component.

Gambler’s Ruin

Gambler's Ruin

usually not fair

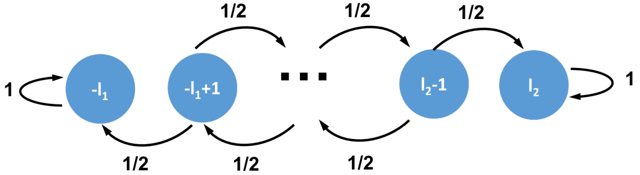
- You and 'a friend' repeatedly toss a fair coin. If it hits heads, you give your friend \$1. If it hits tails, they give you \$1.
- You start with $\$\ell_1$ and your friend starts with $\$\ell_2$. When either of you runs out of money the game terminates.
- What is the probability that you win $\$\ell_2$?

Gambler's Ruin Markov Chain

Let $X_0, X_1, \ldots$ be the Markov chain where X_t is your profit at step t. $X_0 = 0$ and:

$$P_{-\ell_1,-\ell_1} = P_{\ell_2,\ell_2} = 1$$

$$P_{i,i+1} = P_{i,i-1} = 1/2 \text{ for } -\ell_1 < i < \ell_2$$



$X_0 = 0$

$X_1 = \begin{cases} 1 & \text{w.p. } 1/2 \\ -1 & \text{w.p. } 1/2 \end{cases}$

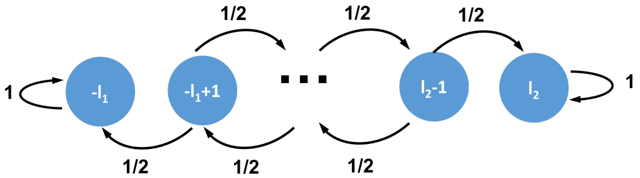
$$\begin{bmatrix} 1 & & & & & \\ \frac{1}{2} & 0 & \frac{1}{2} & & & \\ & \frac{1}{2} & 0 & \frac{1}{2} & & \\ & & \frac{1}{2} & 0 & \frac{1}{2} & \\ & & & & 0 & \\ & & & & 0 & 1 \end{bmatrix}$$

Gambler's Ruin Markov Chain

Let $X_0, X_1, \ldots$ be the Markov chain where X_t is your profit at step t. $X_0 = 0$ and:

$$P_{-\ell_1,-\ell_1} = P_{\ell_2,\ell_2} = 1$$

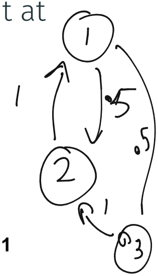
$$P_{i,i+1} = P_{i,i-1} = 1/2 \text{ for } -\ell_1 < i < \ell_2$$



- ℓ_1 and ℓ_2 are absorbing states.
- All i with $-\ell_1 < i < \ell_2$ are transient states. I.e., $\Pr[X_{t'} = i \text{ for some } t' > t \mid X_t = i] < 1$.

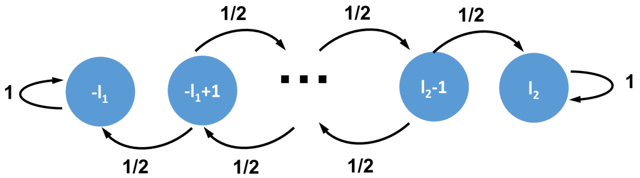
Gambler's Ruin Markov Chain

Let $X_0, X_1, \ldots$ be the Markov chain where X_t is your profit at step t. $X_0 = 0$ and:



$$P_{-\ell_1,-\ell_1} = P_{\ell_2,\ell_2} = 1$$

$$P_{i,i+1} = P_{i,i-1} = 1/2 \text{ for } -\ell_1 < i < \ell_2$$



- ℓ_1 and ℓ_2 are absorbing states.
- All i with $-\ell_1 < i < \ell_2$ are transient states. I.e., $\Pr[X_{t'} = i \text{ for some } t' > t \,|\, X_t = i] < 1$.

Observe that this Markov chain is also a Martingale since $\mathbb{E}[X_{t+1}|X_t] = X_t$.

Gambler's Ruin Analysis

Let $X_0, X_1, \ldots$ be the Markov chain where X_t is your profit at step t.
$X_0 = 0$ and:

$$P_{-\ell_1,-\ell_1} = P_{\ell_2,\ell_2} = 1$$
$$P_{i,i+1} = P_{i,i-1} = 1/2 \text{ for } -\ell_1 < i < \ell_2$$

We want to compute $q = \lim_{t\to\infty} \Pr[X_t = \ell_2]$.

Gambler's Ruin Analysis

Let $X_0, X_1, \ldots$ be the Markov chain where X_t is your profit at step t. $X_0 = 0$ and:

$$P_{-\ell_1,-\ell_1} = P_{\ell_2,\ell_2} = 1$$

$$P_{i,i+1} = P_{i,i-1} = 1/2 \text{ for } -\ell_1 < i < \ell_2$$

$$\mathbb{E}[X_0] = 0$$

$$\mathbb{E}[X_{i+1} | X_i] = 0$$

We want to compute $q = \lim_{t\to\infty} \Pr[X_t = \ell_2]$.

By linearity of expectation, for any i, $\mathbb{E}[X_i] = 0$. Further, for $q = \lim_{t\to\infty} \Pr[X_t = \ell_2]$, since $-\ell_1, \ell_2$ are the only non-transient states,

$$\lim_{t\to\infty} \mathbb{E}[X_t] = \ell_2 q + -\ell_1(1-q) = 0.$$

$$\lim_{t\to\infty} \Pr(X_t = i) = 0$$

$$i \neq -\ell_1, \ell_2$$

$$\frac{\ell_1}{\ell_1+\ell_2}$$

Gambler's Ruin Analysis

Let $X_0, X_1, \ldots$ be the Markov chain where X_t is your profit at step t. $X_0 = 0$ and:

$$P_{-\ell_1,-\ell_1} = P_{\ell_2,\ell_2} = 1$$

$$P_{i,i+1} = P_{i,i-1} = 1/2 \text{ for } -\ell_1 < i < \ell_2$$

We want to compute $q = \lim_{t\to\infty} \Pr[X_t = \ell_2]$.

By linearity of expectation, for any i, $\mathbb{E}[X_i] = 0$. Further, for $q = \lim_{t\to\infty} \Pr[X_t = \ell_2]$, since $-\ell_1, \ell_2$ are the only non-transient states,

$$\lim_{t\to\infty} \mathbb{E}[X_t] = \ell_2 q + -\ell_1(1-q) = 0.$$

Solving for q, we have $q = \frac{\ell_1}{\ell_1+\ell_2}$.

why don't we fix $\ell_2 = 1$

Gambler's Ruin Thought Exercise

What if you always walk away as soon as you win just $1. Then what is your probability of winning, and what are your expected winnings?

Lequivilant to setting $\ell_2 = 1$

LProb of winnig is $\frac{\ell_1}{\ell_1+1} \approx 1$

Lexpected winnigs? 0

$$\frac{\ell_1}{\ell_1+1} \cdot 1 + \frac{1}{\ell_1+1} \cdot -\ell_1 = 0$$

"optional stopping theorem"

Stationary Distributions

Stationary Distribution

A stationary distribution of a Markov chain with transition matrix $P \in [0,1]^{m \times m}$ is a distribution $\pi \in [0,1]^m$ such that $\pi = \pi P$.

I.e. if $X_t \sim \pi$, then $X_{t+1} \sim \pi P = \pi$.

π

.4	.4	.2	0

P

.5	.5	0	0
.5	0	.5	0
0	1	0	0
0	.5	.5	0

=

π

.4	.4	.2	0

Stationary Distribution

A stationary distribution of a Markov chain with transition matrix $P \in [0,1]^{m \times m}$ is a distribution $\pi \in [0,1]^m$ such that $\pi = \pi P$.

I.e. if $X_t \sim \pi$, then $X_{t+1} \sim \pi P = \pi$.

π

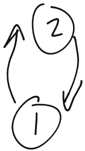
.4	.4	.2	0

P

.5	.5	0	0
.5	0	.5	0
0	1	0	0
0	.5	.5	0

= **π**

.4	.4	.2	0



Think-pair-share: Do all Markov chains have a stationary distribution?

$$[.5 \ .5]\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = [.5, .5]$$

Claim (Existence of Stationary Distribution)

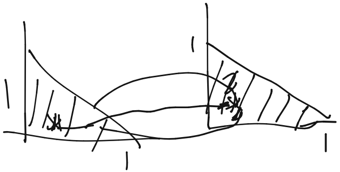
Any Markov chain with a finite state space, and transition matrix $P \in [0,1]^{m \times m}$ has a stationary distribution $\pi \in [0,1]^m$ with $\pi = \pi P$.

Claim (Existence of Stationary Distribution)

Any Markov chain with a finite state space, and transition matrix $P \in [0,1]^{m \times m}$ has a stationary distribution $\pi \in [0,1]^m$ with $\pi = \pi P$.

Follows from the Brouwer fixed point theorem: for any continuous function $f : \mathcal{S} \to \mathcal{S}$, where $\mathcal{S}$ is a compact convex set, there is some x such that $f(x) = x$. ↖ probability simplex

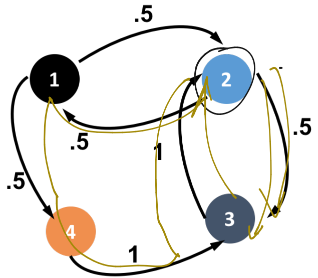
$f(x) = \pi P = \pi$



Periodicity

The periodicity of a state i is defined as:

$$T = \gcd\{t > 0 : \Pr(X_t = i \mid X_0 = i) > 0\}.$$



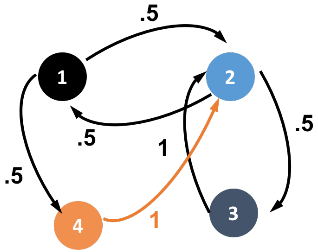
Periodicity of state 2 is 2

Periodicity

The periodicity of a state i is defined as:

3, 5

$$T = \gcd\{t > 0 : \Pr(X_t = i \mid X_0 = i) > 0\}.$$



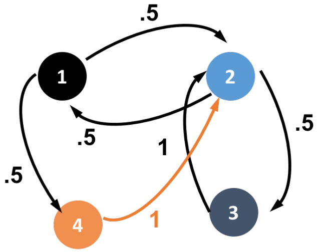
Periodicity of state 4?

↳ 1.

Periodicity

The periodicity of a state i is defined as:

$$T = \gcd\{t > 0 : \Pr(X_t = i \mid X_0 = i) > 0\}.$$

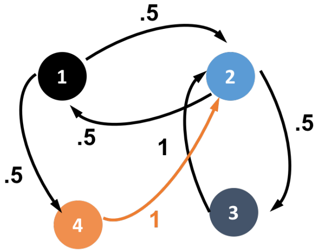


The state is aperiodic if it has periodicity $T = 1$.

Periodicity

The periodicity of a state i is defined as:

$$T = \gcd\{t > 0 : \Pr(X_t = i \mid X_0 = i) > 0\}.$$



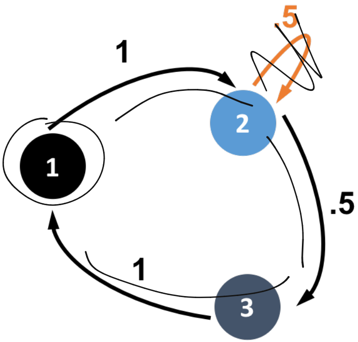
The state is aperiodic if it has periodicity $T = 1$.

A Markov chain is aperiodic if all states are aperiodic.

Periodicity

Claim

If a Markov chain is irreducible, and has at least one self-loop, then it is aperiodic.

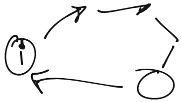


Fundamental Theorem

Theorem (The Fundamental Theorem of Markov Chains)

Let $X_0, X_1, \ldots$ *be a Markov chain with a finite state space and transition matrix* $P \in [0,1]^{m \times m}$*. If the chain is both irreducible and aperiodic,*

1. *There exists a unique stationary distribution* $\pi \in [0,1]^m$ *with* $\pi = \pi P$.
2. *For any states* i, j, $\lim_{t \to \infty} \Pr[X_t = i | X_0 = j] = \pi(i)$*. I.e., for any initial distribution* q_0, $\lim_{t\to\infty} q_t = \lim_{t \to \infty} q_0 P^t = \pi$.
3. $\pi(i) = \frac{1}{\mathbb{E}[\min(t : X_t = i) | X_0 = i]}$*. I.e.,* $\pi(i)$ *is the inverse of the average expected return time from state i back to i.*



Fundamental Theorem

Theorem (The Fundamental Theorem of Markov Chains)

Let $X_0, X_1, \ldots$ be a Markov chain with a finite state space and transition matrix $P \in [0,1]^{m \times m}$. If the chain is both irreducible and aperiodic,

1. *There exists a unique stationary distribution $\pi \in [0,1]^m$ with $\pi = \pi P$.*
2. *For any states i, j, $\lim_{t \to \infty} \Pr[X_t = i | X_0 = j] = \pi(i)$. I.e., for any initial distribution q_0, $\lim_{t \to \infty} q_t = \lim_{t \to \infty} q_0 P^t = \pi$.*
3. $\pi(i) = \frac{1}{\mathbb{E}[\min(t: X_t = i) | X_0 = i]}$. *I.e., $\pi(i)$ is the inverse of the average expected return time from state i back to i.*

In the limit, the probability of being at any state i is independent of the starting state.

Stationary Distribution Example 1

Shuffling Markov Chain: Given a pack of c cards. At each step draw two random cards, swap them and repeat.

Stationary Distribution Example 1

Shuffling Markov Chain: Given a pack of c cards. At each step draw two random cards, swap them and repeat.

- What is the state space of this chain?

↳ all permutations of c cards

↳ $c!$ states

Stationary Distribution Example 1

Shuffling Markov Chain: Given a pack of c cards. At each step draw two random cards, swap them and repeat. or. with prob 1/2 you do nothing.

- What is the state space of this chain?
- What is the transition probability $P_{i,j}$? How does it compare to $P_{j,i}$?

→ depends on adjacency of i, j

i	j
1	1
2	2
3	4
4	3

$$P_{ij} = \frac{1/2}{\binom{c}{2}}$$

$$P_{ji} = \frac{1/2}{\binom{c}{2}}$$

$$P_{ii} = 1/2$$

i	j
1	2
2	1
3	4
4	3

$$P_{ij} = 0$$

$$P_{ji} = 0$$

1 2 3 → 3 1 2 → 2 3 1 → 1 2 3

Stationary Distribution Example 1

Shuffling Markov Chain: Given a pack of c cards. At each step draw two random cards, swap them and repeat.

- What is the state space of this chain?
- What is the transition probability $P_{i,j}$? How does it compare to $P_{j,i}$?
- This Markov chain is symmetric and thus its stationary distribution is uniform, $\pi(i) = \frac{1}{c!}$.

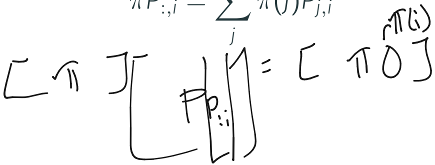
Stationary Distribution Example 1

Shuffling Markov Chain: Given a pack of c cards. At each step draw two random cards, swap them and repeat.

- What is the state space of this chain?
- What is the transition probability $P_{i,j}$? How does it compare to $P_{j,i}$?
- This Markov chain is symmetric and thus its stationary distribution is uniform, $\pi(i) = \frac{1}{c!}$.

Letting $m = c!$ denote the size of the state space,

$$\pi P_{:,i} = \sum_j \pi(j) P_{j,i}$$



Stationary Distribution Example 1

Shuffling Markov Chain: Given a pack of c cards. At each step draw two random cards, swap them and repeat.

- What is the state space of this chain?
- What is the transition probability $P_{i,j}$? How does it compare to $P_{j,i}$?
- This Markov chain is symmetric and thus its stationary distribution is uniform, $\pi(i) = \frac{1}{c!}$.

Letting $m = c!$ denote the size of the state space,

$$\pi P_{:,i} = \sum_j \pi(j) P_{j,i} = \sum_j \pi(j) P_{i,j}$$

Stationary Distribution Example 1

Shuffling Markov Chain: Given a pack of c cards. At each step draw two random cards, swap them and repeat.

- What is the state space of this chain?
- What is the transition probability $P_{i,j}$? How does it compare to $P_{j,i}$?
- This Markov chain is symmetric and thus its stationary distribution is uniform, $\pi(i) = \frac{1}{c!}$.

Letting $m = c!$ denote the size of the state space,

$$\pi P_{:,i} = \sum_j \pi(j) P_{j,i} = \sum_j \pi(j) P_{i,j} = \frac{1}{m} \sum_j P_{i,j}$$

$\pi = \left[\frac{1}{m}, \frac{1}{m}, \ldots \frac{1}{m}\right]$

$\frac{1}{m}$

Stationary Distribution Example 1

Shuffling Markov Chain: Given a pack of c cards. At each step draw two random cards, swap them and repeat.

- What is the state space of this chain?
- What is the transition probability $P_{i,j}$? How does it compare to $P_{j,i}$?
- This Markov chain is symmetric and thus its stationary distribution is uniform, $\pi(i) = \frac{1}{c!}$.

Letting $m = c!$ denote the size of the state space,

$$\pi P_{:,i} = \sum_j \pi(j) P_{j,i} = \sum_j \pi(j) P_{i,j} = \frac{1}{m} \sum_j P_{i,j} = \frac{1}{m} = \pi(i).$$

Stationary Distribution Example 1

Shuffling Markov Chain: Given a pack of c cards. At each step draw two random cards, swap them and repeat.

- What is the state space of this chain?
- What is the transition probability $P_{i,j}$? How does it compare to $P_{j,i}$?
- This Markov chain is symmetric and thus its stationary distribution is uniform, $\pi(i) = \frac{1}{c!}$.

Letting $m = c!$ denote the size of the state space,

$$\pi P_{:,i} = \sum_j \pi(j) P_{j,i} = \sum_j \pi(j) P_{i,j} = \frac{1}{m} \sum_j P_{i,j} = \frac{1}{m} = \pi(i).$$

Once we have exhibited a stationary distribution, we know that it is unique and that the chain converges to it in the limit!

Stationary Distribution Example 2

Random Walk on an Undirected Graph: Consider a random walk on an undirected graph. If it is at node i at step t, then it moves to any of i's neighbors at step $t+1$ with probability $\frac{1}{d_i}$.

- What is the state space of this chain?
- What is the transition probability $P_{i,j}$?

Stationary Distribution Example 2

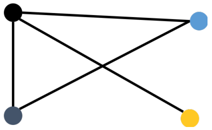
Random Walk on an Undirected Graph: Consider a random walk on an undirected graph. If it is at node i at step t, then it moves to any of i's neighbors at step $t+1$ with probability $\frac{1}{d_i}$.

- What is the state space of this chain?
- What is the transition probability $P_{i,j}$?
- Is this chain aperiodic?

Stationary Distribution Example 2

Random Walk on an Undirected Graph: Consider a random walk on an undirected graph. If it is at node i at step t, then it moves to any of i's neighbors at step $t+1$ with probability $\frac{1}{d_i}$.

- What is the state space of this chain?
- What is the transition probability $P_{i,j}$?
- Is this chain aperiodic?
- If the graph is not bipartite, then there is at least one odd cycle, making the chain aperiodic.



Stationary Distribution Example 2

Random Walk on an Undirected Graph: Consider a random walk on an undirected graph. If it is at node i at step t, then it moves to any of i's neighbors at step $t + 1$ with probability $\frac{1}{d_i}$.

Claim: When the graph is not bipartite, the unique stationary distribution of this Markov chain is given by $\pi(i) = \frac{d_i}{2|E|}$.

Stationary Distribution Example 2

Random Walk on an Undirected Graph: Consider a random walk on an undirected graph. If it is at node i at step t, then it moves to any of i's neighbors at step $t+1$ with probability $\frac{1}{d_i}$.

Claim: When the graph is not bipartite, the unique stationary distribution of this Markov chain is given by $\pi(i) = \frac{d_i}{2|E|}$.

$$\pi P_{:,i} = \sum_j \pi(j) P_{j,i}$$

Stationary Distribution Example 2

Random Walk on an Undirected Graph: Consider a random walk on an undirected graph. If it is at node i at step t, then it moves to any of i's neighbors at step $t+1$ with probability $\frac{1}{d_i}$.

Claim: When the graph is not bipartite, the unique stationary distribution of this Markov chain is given by $\pi(i) = \frac{d_i}{2|E|}$.

$$\pi P_{:,i} = \sum_j \pi(j) P_{j,i} = \sum_j \frac{d_j}{2|E|} \cdot \frac{1}{d_j}$$

Stationary Distribution Example 2

Random Walk on an Undirected Graph: Consider a random walk on an undirected graph. If it is at node i at step t, then it moves to any of i's neighbors at step $t+1$ with probability $\frac{1}{d_i}$.

Claim: When the graph is not bipartite, the unique stationary distribution of this Markov chain is given by $\pi(i) = \frac{d_i}{2|E|}$.

$$\pi P_{:,i} = \sum_j \pi(j) P_{j,i} = \sum_j \frac{d_j}{2|E|} \cdot \frac{1}{d_j} = \sum_j \frac{1}{2|E|}$$

Stationary Distribution Example 2

Random Walk on an Undirected Graph: Consider a random walk on an undirected graph. If it is at node i at step t, then it moves to any of i's neighbors at step $t+1$ with probability $\frac{1}{d_i}$.

Claim: When the graph is not bipartite, the unique stationary distribution of this Markov chain is given by $\pi(i) = \frac{d_i}{2|E|}$.

$$\pi P_{:,i} = \sum_j \pi(j) P_{j,i} = \sum_j \frac{d_j}{2|E|} \cdot \frac{1}{d_j} = \sum_j \frac{1}{2|E|} = \frac{d_i}{2|E|} = \pi(i).$$

Stationary Distribution Example 2

Random Walk on an Undirected Graph: Consider a random walk on an undirected graph. If it is at node i at step t, then it moves to any of i's neighbors at step $t+1$ with probability $\frac{1}{d_i}$.

Claim: When the graph is not bipartite, the unique stationary distribution of this Markov chain is given by $\pi(i) = \frac{d_i}{2|E|}$.

$$\pi P_{:,i} = \sum_j \pi(j) P_{j,i} = \sum_j \frac{d_j}{2|E|} \cdot \frac{1}{d_j} = \sum_j \frac{1}{2|E|} = \frac{d_i}{2|E|} = \pi(i).$$

I.e., the probability of being at a given node i is dependent only on the node's degree, not on the structure of the graph in any other way.