COMPSCI 690RA: Randomized Algorithms and Probabilistic Data Analysis

Prof. Cameron Musco

University of Massachusetts Amherst. Spring 2022.

Lecture 10

Logistics

- Problem Set 3 is due this Friday 4/15 at 8pm.
- We have no class next Wednesday – it's a Monday at UMass.
- I will post a quiz due Tuesday 4/26 at 8pm.
- Remember that office hours are now Thursday at 4pm.

Summary

Last Week: Markov Chains.

- Finish spectral graph sparsification and physical interpretation
- 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. $2^\wedge$

Summary

Last Week: Markov Chains.

- Finish spectral graph sparsification and physical interpretation
- 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.
- Mixing time and its analysis via coupling.
- Markov Chain Monte Carlo (MCMC) methods.

Question 3

Not complete

Points out of 1.00

Flag question

Edit question

Consider a matrix $A \in \mathbb{R}^{5\times 3}$ such that $x = [0, 2, 2, 1, 1]$ is in the column span of A.

$$\underset{5}{\overset{3}{[A]}} [c] = \begin{bmatrix} 0 \\ 2 \\ 2 \\ 1 \\ 1 \end{bmatrix}$$

What can we say about the leverage score of the second row of A, i.e., τ_2?

$\tau_2 \in [0, 1]$

- a. $\tau_2 \le 2$
- b. $\tau_2 \le 0.4$
- c. $\tau_2 \ge 2$
- d. Nothing, without knowing A.
- e. $\tau_2 \ge 0.4$

$\tau_2 \in [0.4, 1]$

$$\tau_2 = \max_{y \in \mathbb{R}^3} \frac{[(Ay)(2)]^2}{\|Ay\|_2^2}$$

$$\frac{x(2)^2}{\|x\|_2^2} = \frac{4}{10} = .4$$

Check

$$\underset{5}{\overset{3}{\begin{bmatrix} \boxed{1\ \|u_2\|_2^2\ 1} \\ \\ \\ \end{bmatrix}}} \begin{bmatrix} \\ z_1 \\ \\ \end{bmatrix} \begin{bmatrix} \\ z_2 \\ \\ \end{bmatrix}$$

- norm 1

Question 4

Not complete

Points out of 1.00

Flag question

Edit question

Consider a matrix $A \in \mathbb{R}^{n\times d}$ with full column rank. Let $U \in \mathbb{R}^{n\times d}$ be its left singular vector matrix. Let $Q \in \mathbb{R}^{n\times d}$ be an orthonormal basis for its column span, computed e.g., using Gram-Schmidt.

True of False: for all $i \in [n]$, $\|Q_{i,:}\|_2 = \|U_{i,:}\|_2$.

Select one:

- ◯ True
- ◯ False

Check

$Q_{i,:} = U_{i,:} \cdot R$ $\qquad Q = UR$

$R \in \mathbb{R}^{d\times d}$ is orthonormal

$\|Q_{i,:}\|_2^2$
$= U_{i,:} \cdot \overbrace{R^T R}^{I} U_{i,:}^T$

$$\begin{array}{c} \\ \\ \end{array}\begin{bmatrix} | & | \\ u_1 & u_2 \\ | & | \end{bmatrix}\begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix} = \begin{array}{cc} q_1 \quad Q & q_2 \\ \begin{bmatrix} | & | \\ u_1 r_{11} + u_2 r_{12}, & u_1 r_{21} + u_2 r_{22} \\ | & | \end{bmatrix} \end{array}$$

$A = QR$ $\qquad A = U\Sigma V^T$

Question 5

Not complete

Points out of 1.00

Flag question

Edit question

Let $E_0, E_1, \ldots$ be independent, identically distributed random variables. Which of the following are Markov chains? Select all that apply.

- [ ] a. $X_0, X_1, \ldots$ where $X_0 = E_0$ and $X_{i+1} = X_i + E_i$
- [ ] b. $E_0, E_1, \ldots$ themselves.
- [ ] c. $X_0, X_1, \ldots$ where $X_0 = E_0$ and $X_{i+1} = X_i \cdot E_i$
- [ ] d. $X_0, X_1, \ldots$ where $X_0 = E_0$, and $X_{i+1} = X_i + E_{i-1} + E_i$
- [ ] e. $X_0, X_1, \ldots$ where $X_0 = E_0$, $X_1 = E_1$, and $X_{i+1} = X_i + X_{i-1} + E_i$

Check

$\Pr(X_{i+1} = v \mid X_i) \neq \Pr(X_{i+1} = v \mid X_i, X_{i-1})$

$X_0 = E_0$

$X_1 = \underset{X_0}{E_0} + E_1 + E_0$

$X_2 = X_1 + E_1 + E_0$

$X_2 = X_0 + X_1 + E_0$

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

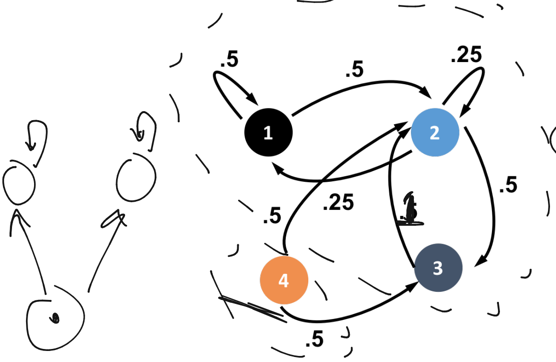
$$P_{i,j} = \Pr(X_{t+1} = j | X_t = i).$$

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

$$\underset{\mathbf{q_1}}{\begin{bmatrix} .5 & .5 & 0 & 0 \end{bmatrix}} \underset{\mathbf{P}}{\begin{bmatrix} .5 & .5 & 0 & 0 \\ .25 & .25 & .5 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & .5 & .5 & 0 \end{bmatrix}} = \underset{\mathbf{q_2}}{\begin{bmatrix} .375 & .375 & .25 & 0 \end{bmatrix}}$$

Markov Chain Review

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



$$P = \begin{bmatrix} .5 & .5 & 0 & 0 \\ .25 & .25 & .5 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & .5 & .5 & 0 \end{bmatrix}$$

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

Gambler's Ruin

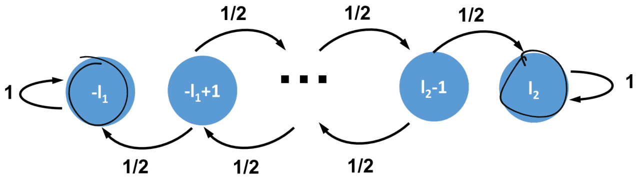
Gambler's Ruin

- 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 $\$\ell_2$?

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

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

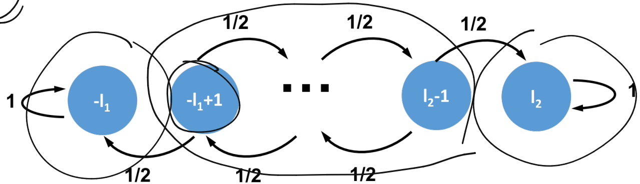
$$P_{i,i+1} = P_{i,i-1} = 1/2 \text{ for } -\ell_1 < i < \ell_2$$



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

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

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



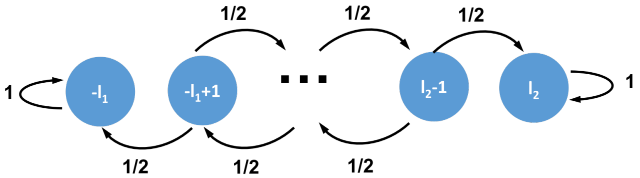
- ℓ_1 and ℓ_2 are absorbing states.
- All i with $-\ell_1 < i < \ell_2$ are transient states. I.e., $\Pr[X_{t'} = i \text{ for some } t' > t \mid X_t = i] < 1$.

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 $\mathbb{E}[X_{t+1}|X_t] = X_t$.

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

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

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

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

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,

$\lim_{t\to\infty} \Pr[X_t = i] = 0$, $i \neq -\ell_1, \ell_2$

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

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

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,

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

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

$q\ell_2 = \ell_1(1-q)$

$q(\ell_1+\ell_2) = \ell_1$

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?

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

$\ell_2 = 1$

$$\left(\frac{\ell_1}{\ell_1 + 1}\right) \cdot 1 + \left(1 - \frac{\ell_1}{\ell_1 + 1}\right) \cdot -\ell_1$$

$$= 0$$

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

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 $X_{t+1} \sim \pi P = \pi$.

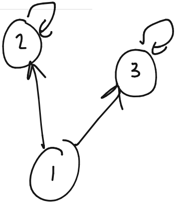
$$\underbrace{\begin{bmatrix} .4 & .4 & .2 & 0 \end{bmatrix}}_{\pi} \underbrace{\begin{bmatrix} .5 & .5 & 0 & 0 \\ .5 & 0 & .5 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & .5 & .5 & 0 \end{bmatrix}}_{\mathbf{P}} = \underbrace{\begin{bmatrix} .4 & .4 & .2 & 0 \end{bmatrix}}_{\pi}$$

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 $X_{t+1} \sim \pi P = \pi$.

$$\underset{\boldsymbol{\pi}}{\begin{bmatrix} .4 & .4 & .2 & 0 \end{bmatrix}} \; \underset{\mathbf{P}}{\begin{bmatrix} .5 & .5 & 0 & 0 \\ .5 & 0 & .5 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & .5 & .5 & 0 \end{bmatrix}} = \underset{\boldsymbol{\pi}}{\begin{bmatrix} .4 & .4 & .2 & 0 \end{bmatrix}}$$

yes if there are finite states.
we



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

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

$X_{i+1} = X_i \pm 1$ — infinite # states

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 $\pi = \pi P$.

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

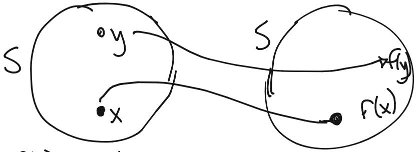
Perron-Frobenius theorem

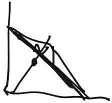
Claim (Existence of Stationary Distribution)

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 $\boxed{\pi = \pi P}$.

$qP = q'$ $P: S \to S$ S is the m-dim simplex

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 $f(x) = x$.





$f(x) = x + 1$

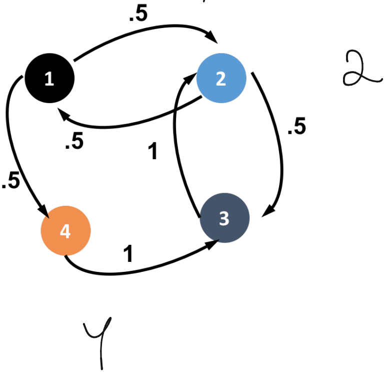
$[_{-\infty}$ $]_{10}$ $\to$ $[_{-9}$ $]_{11}$

Periodicity

The periodicity of a state i is defined as:

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

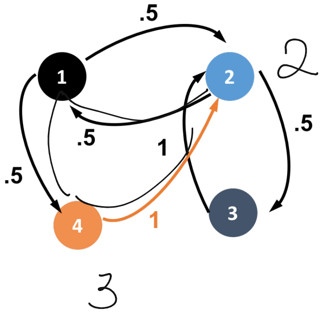
$T = 2$



The periodicity of a state i is defined as:

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

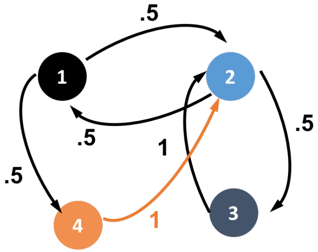
$T = 1$



Periodicity

The periodicity of a state i is defined as:

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

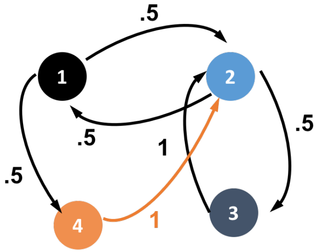


The state is aperiodic if it has periodicity $T = 1$.

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.

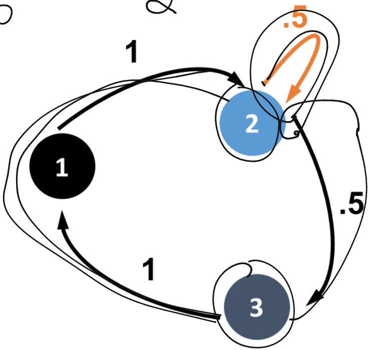
Periodicity

Claim

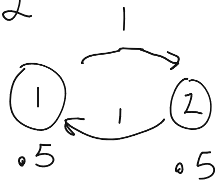
If a Markov chain is irreducible, and has at least one self-loop, then it is aperiodic.

3, 4, 5, 6

2



2



$q_1 = [1 \quad 0]$

$q_2 = [0 \quad 1]$

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.*

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,

, $\sum \pi(i) = 1$

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 a random card, place it on top, and repeat.

Stationary Distribution Example 1

Shuffling Markov Chain: Given a pack of c cards. At each step draw a random card, place it on top, and repeat.

- What is the state space of this chain?

$c!$ all possible permutations of c distinct cards

Stationary Distribution Example 1

Shuffling Markov Chain: Given a pack of c cards. At each step draw ~~a~~ two random card, ~~place it on top, and repeat.~~ swap them

swap with the top card 2

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

$P_{ij} = P_{ji}$

This slide has many errors

1	3
2	2
3	1
4	4
5	5
i	j

$P_{ii} = 1/c$

$P_{ij} = 1/c$ $\quad$ $P_{ij} = 0$

$P_{ji} = 1/c$ $\quad$ $P_{ji} = 0$

Stationary Distribution Example 1

Shuffling Markov Chain: Given a pack of c cards. At each step draw a random card, place it on top, and repeat.

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

  $P_{ij} = P_{ji}$
- 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 a random card, place it on top, 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,

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

$$[\ \pi\ ][P]\,[P_{:,i}] = [\ \pi \quad \pi(i)\ ]$$

Stationary Distribution Example 1

Shuffling Markov Chain: Given a pack of c cards. At each step draw a random card, place it on top, 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 a random card, place it on top, 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, $\pi(i) = \frac{1}{c!}$.

$m = c!$

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

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

$= 1$

Stationary Distribution Example 1

Shuffling Markov Chain: Given a pack of c cards. At each step draw a random card, place it on top, 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,

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

$$\left[\frac{1}{m} \quad \frac{1}{m} \quad \cdots \quad \frac{1}{m}\right]$$

Stationary Distribution Example 1

Shuffling Markov Chain: Given a pack of c cards. At each step draw a random card, place it on top, 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,

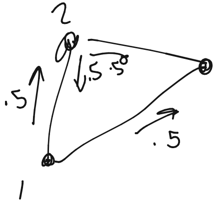
$$\pi P_{:,i} = \sum_j \pi(j) P_{j,i} = \sum_j \pi(j) P_{i,j} = \frac{1}{m} \sum_j P_{i,j} = \frac{1}{m} = \pi(i).$$

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 $\frac{1}{d_i}$.

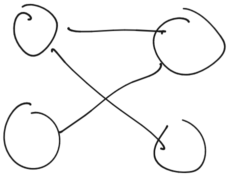
- What is the state space of this chain? V
- What is the transition probability $P_{i,j}$? $\frac{1}{d_i}$



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 $\frac{1}{d_i}$.

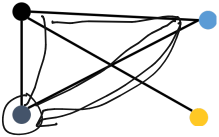
- What is the state space of this chain?
- What is the transition probability $P_{i,j}$?
- Is this chain aperiodic?



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 $\pi(i) = \frac{d_i}{|E|}$.

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 $\pi(i) = \frac{d_i}{2|E|}$.

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

$$\pi P = \pi$$

$$\pi P_{:,i} = \pi(i) = \frac{d_i}{2|E|}$$

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 $\pi(i) = \frac{d_i}{2|E|}$.

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

$$P_{j,i} = \begin{cases} \frac{1}{d_j} & \text{if } j \in N(i) \\ 0 & \text{if } j \notin N(i) \end{cases}$$

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 $\pi(i) = \frac{d_i}{|E|}$.

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

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 $\pi(i) = \frac{d_i}{2|E|}$.

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

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 $\pi(i) = \frac{d_i}{|E|}$.

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

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.

Mixing Times

~ [wasserstein]

Definition (Total Variation (TV) Distance)

For two distributions $p, q \in [0,1]^m$ over state space $[m]$, the total variation distance is given by:

$$\|p - q\|_{TV} = \frac{1}{2}\sum_{i \in [m]} |p(i) - q(i)| = \max_{A \subseteq [m]} |p(A) - q(A)|.$$

$$\begin{array}{cc} p & \\ \begin{bmatrix} \frac{1}{6} \\ \frac{1}{6} \\ \vdots \\ 1/6 \end{bmatrix} & \begin{bmatrix} 1/2 \\ 1/2 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \end{array}$$

$$\left(\frac{1}{2} + \frac{1}{2}\right) - \frac{1}{6} - \frac{1}{6} = \frac{2}{3}$$

$$\frac{1}{2}\left(4 \cdot \left|-\frac{1}{6}\right| + 2 \cdot \left(\frac{1}{2} - \frac{1}{6}\right)\right) = \frac{2}{3}$$

Total Variation Distance

Definition (Total Variation (TV) Distance)

For two distributions $p, q \in [0,1]^m$ over state space $[m]$, the total variation distance is given by:

$$\|p - q\|_{TV} = \frac{1}{2} \sum_{i \in [m]} |p(i) - q(i)| = \max_{A \subseteq [m]} |p(A) - q(A)|.$$

Kontorovich-Rubinstein duality: Let P, Q be possibility correlated random variables with marginal distributions p, q. Then

~~coupling~~

$$\|p - q\|_{TV} \leq \Pr[\mathrm{P} \neq \mathrm{Q}].$$

$P = Q$

$$\begin{array}{cc} p & q \\ \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \end{bmatrix} & \begin{bmatrix} 2/3 \\ 1/3 \end{bmatrix} \end{array}$$

$$\|p - q\|_{TV} = \min_{P,Q} \Pr[P \neq Q]$$

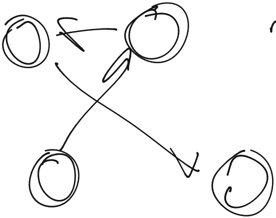
$p = q$ so $\|p - q\|_{TV} = 0$

Mixing Time

Definition (Mixing Time)

Consider a Markov chain $X_0, X_1, \ldots$ with unique stationary distribution π. Let $q_{i,t}$ be the distribution over states at time t assuming $X_0 = i$. The mixing time is defined as:

$$\tau(\epsilon) = \min\left\{ t : \max_{i \in [m]} \|q_{i,t} - \pi\|_{TV} \leq \epsilon \right\}.$$



$\min\{100, 101, \ldots\}$

Mixing Time

Definition (Mixing Time)

Consider a Markov chain $X_0, X_1, \ldots$ with unique stationary distribution π. Let $q_{i,t}$ be the distribution over states at time t assuming $X_0 = i$. The mixing time is defined as:

$$\tau(\epsilon) = \min \left\{ t : \max_{i \in [m]} \|q_{i,t} - \pi\|_{TV} \leq \epsilon \right\}.$$

I.e., what is the maximum time it takes the Markov chain to converge to within ϵ in TV distance of the stationary distribution?

Definition (Mixing Time)

Consider a Markov chain $X_0, X_1, \ldots$ with unique stationary distribution π. Let $q_{i,t}$ be the distribution over states at time t assuming $X_0 = i$. The mixing time is defined as:

$$\tau(\epsilon) = \min \left\{ t : \max_{i \in [m]} \|q_{i,t} - \pi\|_{TV} \leq \epsilon \right\}.$$

I.e., what is the maximum time it takes the Markov chain to converge to within ϵ in TV distance of the stationary distribution?

Claim: If $X_0, X_1, \ldots$ is finite, irreducible, and aperiodic, then $\tau(\epsilon) \leq \tau(1/2) \cdot c \log(1/\epsilon)$ for large enough constant c.

$q_0 P^t = q_t \approx \pi$

$q_t = \pi + e$

$q_{2t} = \pi P^t + e P^t$

Coupling Motivation

Claim: $\max_{i \in [m]} \|q_{i,t} - \pi\|_{TV} \leq \max_{i,j \in [m]} \|q_{i,t} - q_{j,t}\|_{TV}$.

Coupling Motivation

Claim: $\max_{i \in [m]} \|q_{i,t} - \pi\|_{TV} \leq \max_{i,j \in [m]} \|q_{i,t} - q_{j,t}\|_{TV}$.

$$\|q_{i,t} - \pi\|_{TV} = \|q_{i,t} - \pi P^t\|_{TV}$$

i, j starting states of markov chain.

Coupling Motivation

Claim: $\max_{i\in[m]} \|q_{i,t} - \pi\|_{TV} \leq \max_{i,j\in[m]} \|q_{i,t} - q_{j,t}\|_{TV}$.

$$\begin{aligned}\|q_{i,t} - \pi\|_{TV} &= \|q_{i,t} - \pi P^t\|_{TV} \\ &= \|q_{i,t} - \sum_j \pi(j)\underbrace{e_j P^t}_{q_{j,t}}\|_{TV}\end{aligned}$$

$\pi = \pi(j) e_j \qquad e_j = [0\ 0\ 1\ 0\ 0\ 0]$

Coupling Motivation

Claim: $\max_{i \in [m]} \|q_{i,t} - \pi\|_{TV} \le \max_{i,j \in [m]} \|q_{i,t} - q_{j,t}\|_{TV}$.

$$\begin{aligned}
\|q_{i,t} - \pi\|_{TV} &= \|q_{i,t} - \pi P^t\|_{TV} \\
&= \|q_{i,t} - \sum_j \pi(j) e_j P^t\|_{TV} \\
&= \|q_{i,t} - \sum_j \pi(j) q_{j,t}\|_{TV} \\
&= \|\sum_j \pi(j)(q_{i,t} - q_{j,t})\|_{TV}
\end{aligned}$$

Coupling Motivation

Claim: $\max_{i\in[m]} \|q_{i,t} - \pi\|_{TV} \le \max_{i,j\in[m]} \|q_{i,t} - q_{j,t}\|_{TV}$.

$$\begin{aligned}
\|q_{i,t} - \pi\|_{TV} &= \|q_{i,t} - \pi P^t\|_{TV} \\
&= \|q_{i,t} - \sum_j \pi(j) e_j P^t\|_{TV} \\
&= \|q_{i,t} - \sum_j \pi(j) q_{j,t}\|_{TV} \\
&\le \sum_j \|\pi(j) q_{i,t} - \pi(j) q_{j,t}\|_{TV}
\end{aligned}$$

Coupling Motivation

Claim: $\max_{i\in[m]} \|q_{i,t} - \pi\|_{TV} \leq \max_{i,j\in[m]} \|q_{i,t} - q_{j,t}\|_{TV}$.

$$\begin{aligned}
\|q_{i,t} - \pi\|_{TV} &= \|q_{i,t} - \pi P^t\|_{TV} \\
&= \|q_{i,t} - \sum_j \pi(j) e_j P^t\|_{TV} \\
&= \|q_{i,t} - \sum_j \pi(j) q_{j,t}\|_{TV} \\
&\leq \sum_j \|\pi(j) q_{i,t} - \pi(j) q_{j,t}\|_{TV} = \sum_j \pi(j) \underline{\|q_{i,t} - q_{j,t}\|_{TV}} \\
&\leq \max_{j\in[m]} \underline{\|q_{i,t} - q_{j,t}\|_{TV}}.
\end{aligned}$$

Coupling Motivation

Claim: $\max_{i\in[m]} \|q_{i,t} - \pi\|_{TV} \leq \max_{i,j\in[m]} \|q_{i,t} - q_{j,t}\|_{TV}$.

$$
\begin{aligned}
\|q_{i,t} - \pi\|_{TV} &= \|q_{i,t} - \pi P^t\|_{TV} \\
&= \|q_{i,t} - \sum_j \pi(j) e_j P^t\|_{TV} \\
&= \|q_{i,t} - \sum_j \pi(j) q_{j,t}\|_{TV} \\
&\leq \sum_j \|\pi(j) q_{i,t} - \pi(j) q_{j,t}\|_{TV} \\
&\leq \max_{j\in[m]} \|q_{i,t} - q_{j,t}\|_{TV}.
\end{aligned}
$$

Coupling: A common technique for bounding the mixing time by showing that $\max_{i,j\in[m]} \|q_{i,t} - q_{j,t}\|_{TV}$ is small.