Lecture 27

Markov Chains

and a Few Paradoxes.
Summary of Lecture 26

- A Markov Chain is a process that moves from state to state randomly and only remembers its current state.

  - \( \mathcal{X} \) - Finite state space, usually \( \mathcal{X} = \{1, 2, \ldots, k\} \).
  - \( \pi_0 \) - the initial distribution.
  - \( P(i,j) \) - Prob. to move from state \( i \) to \( j \).

This defines a Markov Chain: sequence of r.v.s \( X_0, X_1, X_2 \ldots \).

\[
\Pr[X_0 = i] = \pi_0(i)
\]

\[
\Pr[X_n = j \mid X_{n-1} = i, \ldots, X_1, X_0] = P(i,j).
\]

\( \Pi_n \) : Prob. dist of \( X_n \).

\( \Pi_n = \pi_0 \cdot P^n \)
Let \( \{X_n\}_{n=0}^{\infty} \) be a MC on \( \mathcal{X} \), \( A \subseteq \mathcal{X} \).

- \( \beta(i) \) is the expected time to reach \( A \) starting from \( i \).
- \( \beta(i) = 0 \) for \( i \in A \).
- \( \beta(i) = 1 + \sum_j P(i,j) \beta(j) \) for \( i \notin A \).
First Step Equations

Let \( \{X_n\}_{n=0}^\infty \) be a MC on \( X \) with \( A, B \subseteq X \) and \( A \cap B \) disjoint.

\[
\alpha(i) = \Pr \left[ \text{reaching } A \text{ before } B, \text{starting from } i \right]
\]

\[
\begin{align*}
\alpha(i) &= 0 & \text{for } i \in B \\
\alpha(i) &= 1 & \text{for } i \in A \\
\alpha(i) &= \sum_{j} P(i,j) \cdot \alpha(j) & \text{for } i \notin A \cup B.
\end{align*}
\]
**Definition:**

A distribution \( \pi \) over \( \mathcal{X} \) is stationary (aka invariant) if \( \pi = \pi P \).

If \( \pi_0 \) is stationary then \( \pi_n = \pi_0 \).
Stationary Distribution - Example

\[ \Pi \text{ is stationary iff } \begin{pmatrix} \pi(1) \\ \pi(2) \end{pmatrix} = \begin{pmatrix} \pi(1) \\ \pi(2) \end{pmatrix} \cdot \begin{pmatrix} 1-p & p \\ q & 1-q \end{pmatrix} \]

\[ \pi(1) = \pi(1) \cdot (1-p) + \pi(2) \cdot q \quad \iff \quad \pi(1) \cdot p = \pi(2) \cdot q \]

\[ \pi(2) = \pi(1) \cdot p + \pi(2) \cdot (1-q) \quad \iff \quad \pi(1) \cdot p = \pi(2) \cdot q \]

\[ \pi(1) + \pi(2) = 1 \]

Solution: \[ \Pi = \left[ \frac{q}{p+q}, \frac{p}{p+q} \right] \]
Stationary Distributions - Example 2

Which distributions are stationary?

all of them.

\[ \forall \pi \quad \pi = \pi \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \]
Irreducible Markov Chains

A MC is irreducible if you can go from every state \( i \) to every state \( j \) (possibly in multiple steps).

Which MC are irreducible?

(A)  

(B)  

(C)  

(D)
**Theorem:** Any finite irreducible MC has one and only one stationary distribution.

**Theorem 2:** (Long Term Fraction of Time in States)

If \((X_n)_{n=0}^{\infty}\) is an irreducible MC on \(\{1, \ldots, k\}\) with stationary distribution \(\pi\).

Then, for any start dist. \(\pi_0\), for all \(i \in \{1, \ldots, k\}\)

\[
\lim_{n \to \infty} \frac{1}{n} \sum_{m=0}^{n-1} 1_{\{X_m = i\}} = \pi(i).
\]
Intuition: Start at a dist. \( x_0 \) and suppose the limits exist. Denote by

\[ f(i) = \lim_{n \to \infty} \frac{1}{n} \sum_{m=0}^{n-1} 1_{X_m = i} \]

What’s the frac. of times we visit \( i \)?

What’s the frac. of times we visit \( i \) and then move to \( j \)?

In time \( m \), the MC is at state \( i \).
Intuition: Start at a dist. $\pi_0$ and suppose the limits exist. Denote by

$$f(i) = \lim_{n \to \infty} \frac{1}{n} \sum_{m=0}^{n-1} 1_{x_m = i}$$

What's the frac. of times we visit $i$?

What's the frac. of times we visit $i$ and then move to $j$?

$$f(i) \cdot p(i,j)$$

Frac. of times we're at $j = \sum_{i} f(i) \cdot p(i,j)$$

$$\forall j \quad f(j) = \sum_{i} f(i) \cdot p(i,j)$$

In matrix-vector form $\pi = \pi \cdot \mathbf{P}$.
Let's see a simulation:

Recall this MC from lecture 26.

We'll run it for many steps and count how many times we've been in each step.
Converges to the Stationary Distribution

Example:

\[
\begin{pmatrix}
1 & \uparrow \\
\downarrow & 2
\end{pmatrix}
\]

The MC is irreducible.

It's stationary dist satisfies

\[
\begin{pmatrix}
\pi(1) \\
\pi(2)
\end{pmatrix} =
\begin{pmatrix}
\pi(1) \\
\pi(2)
\end{pmatrix} 
\begin{pmatrix}
0 & 1 \\
1 & 0
\end{pmatrix} 
\]

\[
\Rightarrow \pi(1) = \pi(2) = \frac{1}{2}
\]

But starting from 1: \( \pi_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \)

\( \pi_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \)

\( \pi_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \) ...

\( \pi_{2n} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \)

\( \pi_{2n+1} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \)
Periodicity

**Defn.** The periodicity of a Markov Chain with transition matrix $P$ is the $\gcd$ of lengths of all closed walks in the chain $\gcd(n > 0 \mid \exists i \text{ s.t. } P^n(i,i) > 0)$.

A Markov chain is **aperiodic** if this $\gcd = 1$. 
Which Markov Chains are Aperiodic?

A. 

B. 

period: 2

C. 

gcd(2, 2, 1, ...) = 1
Theorem: Let $X_n$ be an irreducible aperiodic Markov chain with stationary dist $\pi$. Then, no matter what's the starting dist. To $\forall i: \pi_n(i) \xrightarrow{n \to \infty} \pi(i)$. 
Some Paradoxes
St. Petersburg Paradox

A casino is offering you to play the following game:

- Start with a stake of 2 £.
- At each point flip a fair coin

\[ \begin{align*}
H &: \text{double the stake.} \\
T &: \text{stop and give stake to player.}
\end{align*} \]

How much are you willing to pay to play this game?

What's the expected winning stake?

\[ X \text{ r.v. capturing winning stake} \]

\[ \Pr[X=2] = \frac{1}{2}, \quad \Pr[X=4] = \frac{1}{4}, \quad \ldots \quad \Pr[X=2^k] = \frac{1}{2^k}. \]

\[ \mathbb{E}X = \sum_{i=1}^{\infty} \Pr[X=2^i] \cdot 2^i = \sum_{i=1}^{\infty} \frac{1}{2^i} \cdot 2^i = \sum_{i=1}^{\infty} 1 = \infty. \]
St. Petersburg Paradox

A casino is offering you to play the following game:

- Start with a stake of 2 £.
- At each point flip a fair coin

\[ H: \text{double the stake.} \]
\[ T: \text{stop and give stake to player.} \]

How much are you willing to pay to play this game?

If the casino has only \( n=2^k \) dollars. What's the expected win?

\[
E[X] = 2 \cdot \frac{1}{2} + 4 \cdot \frac{1}{4} + \ldots + 2^{k-1} \cdot \frac{1}{2^{k-1}} + 2^k \cdot \frac{1}{2^k} = k+1 = \log_2 n + 1
\]
Double-or-Nothing

Suppose you go to a casino and you want a betting strategy that will guarantee you win 1 

They have a simple game. You choose how much to bet $x$. 

They flip a fair coin $\begin{array}{c} H \quad \text{you win } x \text{ dollars} \\ T \quad \text{lose } x \text{ dollars.} \end{array}$

Strategy: Start with betting 1 

else, bet 2 

else, bet $-v$ 

$
\cdots$

Eventually a $T$ will be flipped and you will win overall 1 

Double-or-Nothing

Bet $X$ dollars; w.p. \( \frac{1}{2} \) win $X$; w.p. \( \frac{1}{2} \) lose $X$.

**Strategy:**

Start with betting $1\$, if successful, stop.
else, bet $2\$, if successful stop.
else, bet $4\$, if " stop.

Eventually a H will be flipped and you will win overall $1\$.

What happens if you have a limited budget, say $1023\$:

You’ll lose $1023\$ with prob. \( \frac{1}{1024} \)
and win $1\$ with prob. \( 1 - \frac{1}{1024} \).

On expectation, you’ll even out.

\[
\frac{1}{1024} \cdot (-1023) + \left(-\frac{1}{1024}\right) = 0.
\]
**Confusing Statistics:**

**Simpson's Paradox**

Results from real-life medical study:

<table>
<thead>
<tr>
<th></th>
<th>Treatment A</th>
<th>Treatment B</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small Stones</td>
<td>Group 1: 93% (81/87)</td>
<td>Group 2: 87% (234/270)</td>
</tr>
<tr>
<td>Large Stones</td>
<td>Group 3: 73% (192/263)</td>
<td>Group 4: 69% (55/80)</td>
</tr>
<tr>
<td>Both</td>
<td>78% (273/350)</td>
<td>83% (289/350)</td>
</tr>
</tbody>
</table>

Which treatment is better?
What's the best linear predictor of \( Y \) given \( X \)?
What's the best linear predictor of $Y$ given $X$?

Trend is down: as $X$ increases, we predict that $Y$ decreases.
But if the data is coming from a mixture of two distributions red and blue:

In each subgroup trend is up: as \( X \) increases we predict that \( Y \) increases.
Special Thanks to the CS70 staff:

- Prof. Rao.
- TAS: Alec, Christine, Gavin, Shreyas
  Caroline, Dominic, Evelyn, Hongxun, Josh, Michelle, Nathan, Sriram, Xin, Yinxuo.
- Readers: Allen, Andrew, Arjun, Ashton, Ayush, Bryan, Carolyn, Casey, Charming, Evan, Haroon, James, Jason, Jeff, Jennifer, Matthew, Nathan, Owen, Pranesh, Raymond, Rohan, Sena, Sina, Thomas, Youngmin.

Academic Interns: Gabriel, Michael, Shaswat, Sriram, Tejasv, Yash, Zixun.
Thank you and good luck in the final!