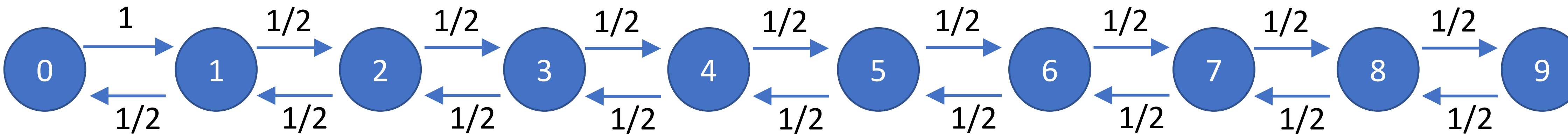
Class 13

Markov Chains I

Questions?

In the mini-lectures

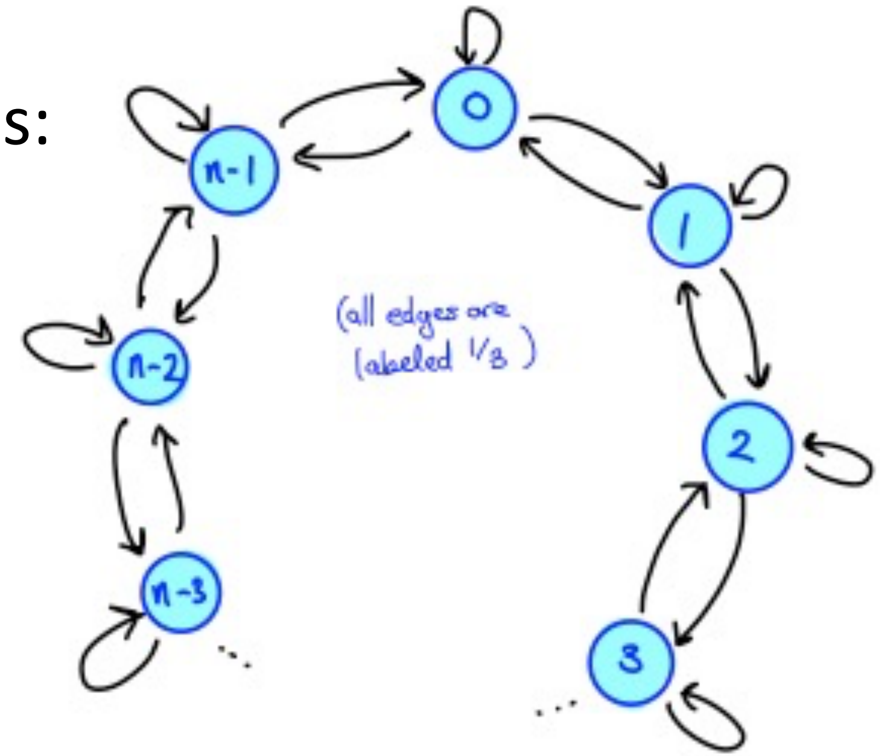
- We introduced Markov chains
- We saw how to manipulate transition matrices to do calculations
- We saw one way to analyze a chain that looked like this:



Today

- We will see another way to use the transition matrix to analyze (certain nice) Markov chains.

- We'll analyze a Markov chain that looks like this:



Group work!

1. How to figure out $\Pr[X_t = 2 \mid X_0 = 0]$?

$$\mathbb{P}[X_t = 2 \mid X_0 = 0] = (1\ 0\ 0\ 0)\, P^t \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

Whoops, everything on this slide should be divided by 3.

2. $P = F \cdot \Lambda \cdot F^*$

- Check that the columns of F are eigenvectors of P:

$$\begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = 3 \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

Eigenvector with eigenvalue 3

$$\begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \end{pmatrix}$$

Eigenvector with eigenvalue -1

$$\begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ -i \\ -1 \\ i \end{pmatrix} = \begin{pmatrix} 1 \\ -i \\ -1 \\ i \end{pmatrix}$$

Eigenvector with eigenvalue 1

$$\begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ i \\ -1 \\ -i \end{pmatrix} = \begin{pmatrix} 1 \\ i \\ -1 \\ -i \end{pmatrix}$$

Eigenvector with eigenvalue 1

$$\Rightarrow P \cdot F = F \cdot \begin{pmatrix} 3 & & & \\ & 1 & & \\ & & -1 & \\ & & & 1 \end{pmatrix} \quad \Rightarrow P = F \cdot \underbrace{\begin{pmatrix} 3 & & & \\ & 1 & & \\ & & -1 & \\ & & & 1 \end{pmatrix}}_{\text{call this } \Lambda} F^*$$

3. How to figure out $\Pr[X_t = 2 \mid X_0 = 0]$?

$$\mathbb{P}[X_t = 2 \mid X_0 = 0] = (1\ 0\ 0\ 0)\, P^t \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} = (1\ 0\ 0\ 0)\,(F \Lambda F^*)^t \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

$$= \underbrace{(1\ 0\ 0\ 0)\, F}_{\text{first row of } F}\ \Lambda^t\ \underbrace{F^* \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}}_{3^{rd} \text{ column of } F^*} = \frac{1}{4}(1\ 1\ 1\ 1) \begin{pmatrix} 1 & & & \\ & (1/3)^t & & \\ & & (-1/3)^t & \\ & & & (1/3)^t \end{pmatrix} \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$$

It's easy to take powers of a diagonal matrix!

$$= \frac{1}{4} + \frac{1}{4}\left(-2/3^t + \left(\frac{-1}{3}\right)^t \right)$$

$$= \frac{1}{4} + O(1/3^t)$$

More group work!

- Extending this logic to a bigger cycle!

1. $P = F \cdot \mathrm{D} \cdot F^*$

Once again, everything should be divided by 3.

- Check that the columns of F are eigenvectors of P:

$$\underbrace{\begin{pmatrix} 1 & 1 & 0 & 0 & \cdots & & 1 \\ 1 & 1 & 1 & 0 & \cdots & & 0 \\ 0 & 1 & 1 & 1 & 0 & \cdots & \\ \vdots & & & \ddots & & & \\ & & & 1 & 1 & 1 & 0 \\ 0 & & & & 1 & 1 & 1 \\ 1 & 0 & 0 & \cdots & & 1 & 1 \end{pmatrix}}_{P} \underbrace{\begin{pmatrix} \\ \bullet \\ \\ \end{pmatrix}}_{k^{th} \text{ column of } F} = \begin{pmatrix} \\ \bullet \\ \\ \end{pmatrix}$$

j^{th} entry is $\frac{1}{\sqrt{n}} e^{-2\pi i jk/n}$

j^{th} entry is

$$\frac{1}{\sqrt{n}} e^{-2\pi \boldsymbol{i} jk/n} + \frac{1}{\sqrt{n}} e^{-2\pi \boldsymbol{i}(j+1)k/n} + \frac{1}{\sqrt{n}} e^{-2\pi \boldsymbol{i}(j-1)\frac{k}{n}}$$

$$= \frac{1}{\sqrt{n}} e^{-2\pi \boldsymbol{i} jk/n} \left[1 + e^{-2\pi \boldsymbol{i} k/n} + e^{2\pi \boldsymbol{i} k/n} \right]$$

$$= \frac{1}{\sqrt{n}} e^{-2\pi \boldsymbol{i} jk/n} \left[1 + 2\cos\left(\frac{2\pi k}{n}\right) \right]$$

This is the k'th eigenvalue. Let D be a matrix with these values on the diagonal.

Thus, $PF = FD \Rightarrow P = FDF^*$, as desired.

2. How to figure out $\Pr[X_t = 0 \mid X_0 = 0]$?

$$\Pr[X_t = 0 \mid X_0 = 0] = (1\,0\,0\cdots 0)\,P^t \begin{pmatrix}1\\0\\0\\\vdots\\0\end{pmatrix} = (1\,0\,0\cdots 0)\,F_n D^t F_n^* \begin{pmatrix}1\\0\\0\\\vdots\\0\end{pmatrix}$$

$$= \frac{1}{n}\,\mathbb{1}^T \begin{array}{c c} & j \\ j & \begin{pmatrix} \ddots & & \\ & \left(\frac{1+2\cos(2\pi j/n)}{3}\right)^t & \\ & & \ddots \end{pmatrix} \end{array} \mathbb{1}$$

$$= \frac{1}{n}\sum_{j=0}^{n-1}\left(\frac{1+2\cos(2\pi j/n)}{3}\right)^t$$

$$= \frac{1}{n} + \frac{1}{n}\sum_{j=1}^{n-1}\left(\frac{1+2\cos(2\pi j/n)}{3}\right)^t$$

3. $\Pr[\, X_t = 0 \mid X_0 = 0 \,] \rightarrow \frac{1}{n}$ as $t \rightarrow \infty$

$$\mathbb{P}[X_t = 0 \mid X_0 = 0] = \frac{1}{n} + \frac{1}{n} \sum_{j=1}^{n-1} \underbrace{\left(\frac{1 + 2\cos(2\pi j / n)}{3} \right)^t}_{\text{this} \rightarrow 0 \text{ as } t \rightarrow \infty}$$

since $1 + 2\cos\left(\frac{2\pi j}{n}\right) < 3$

4. How fast does it converge?

$$\mathbb{P}[X_t = 0 \mid X_0 = 0] = \frac{1}{n} + \frac{1}{n}\sum_{j=1}^{n-1} \underbrace{\left(\frac{1 + 2\cos(2\pi j/n)}{3}\right)^t}$$

How big does t have to be before this is small?

CLAIM: t should be about n^2 before this starts to get small.

Suppose that $t = o(n^2)$ is too small.

$$\frac{1}{n}\sum_{j=1}^{n-1}\left(\frac{1+2\cos(2\pi j/n)}{3}\right)^t \geq \frac{1}{n}\sum_{j=1}^{n/\sqrt{t}}\left(\frac{1+2\cos(2\pi j/n)}{3}\right)^t$$

$$\geq \frac{1}{n}\sum_{j=1}^{n/\sqrt{t}}\left(1-\frac{Cj^2}{n^2}\right)^t \quad \text{for some constant } C$$

$$\approx \frac{1}{n}\sum_{j=1}^{n/\sqrt{t}} \exp\left(-Cj^2 t/n^2\right)$$

For $j \leq \frac{n}{1000}$, say, $\cos\left(\frac{2\pi j}{n}\right) = 1 - \Theta\left(\left(j/n\right)^2\right)$

assume t is big enough that $n/\sqrt{t}$ is a lot less than n.

Suppose that $t = o(n^2)$ is too small.

$$\frac{1}{n}\sum_{j=1}^{n-1}\left(\frac{1+2\cos(2\pi j/n)}{3}\right)^t \approx \frac{1}{n}\sum_{j=1}^{n/\sqrt{t}} \exp\left(-cj^2 t/n^2\right)$$

$$\geq \frac{1}{n}\sum_{j=1}^{n/\sqrt{t}} \exp\left(-c\left(\frac{n}{\sqrt{t}}\right)^2 \frac{t}{n^2}\right)$$

$$= \exp(-c)/\sqrt{t}.$$

$$\gg 1/n$$

$$t = o(n^2) \Rightarrow \sqrt{t} = o(n)$$

$$\Rightarrow 1/\sqrt{t} = \omega(1/n)$$

4. How fast does it converge?

$$\mathbb{P}[X_t = 0 \mid X_0 = 0] = \frac{1}{n} + \frac{1}{n}\sum_{j=1}^{n-1} \underbrace{\left(\frac{1 + 2\cos(2\pi j/n)}{3}\right)^t}$$

- We just saw: If $t = o(n^2)$, then this term is not small, it's $\omega\left(\frac{1}{n}\right)$.

How big does t have to be before this is small?

CLAIM: t should be about n^2 before this starts to get small.

Suppose that $t \geq Cn^2 \log n$ is big enough

Here, C is some constant TBD

$$\frac{1}{n}\sum_{j=1}^{n-1}\left(\frac{1+2\cos(2\pi j/n)}{3}\right)^t$$

this is biggest for $j=1$.

$$\leq \frac{1}{n}\sum_{j=1}^{n-1}\left(\frac{1+2\cos\left(\frac{2\pi}{n}\right)}{3}\right)^t$$

$$\approx \frac{1}{n}\sum_{j=1}^{n-1}\left(\frac{1+2(1-\Theta(1/n^2))}{3}\right)^t$$

$$= \frac{1}{n}\sum_{j=1}^{n-1}\left(1-\Theta(1/n^2)\right)^t$$

$$\leq \frac{1}{n}\sum_{j=1}^{n-1}\exp(-C't/n^2) \quad \text{for some constant } C'$$

Suppose that $t \geq Cn^2 \log n$ is big enough

Here, C is some constant TBD

$$\frac{1}{n}\sum_{j=1}^{n-1}\left(\frac{1+2\cos(2\pi j/n)}{3}\right)^t \leq \frac{1}{n}\sum_{j=1}^{n-1}\exp\left(-C't/n^2\right) \quad \text{for some constant } C'$$

this is biggest for $j=1$.

$$\leq \frac{1}{n}\sum_{j=1}^{n-1}\exp\left(-C'\cdot C\log(n)\right)$$

$$= o(1/n) \quad \text{if } C \geq 2/C', \text{ say.}$$

4. How fast does it converge?

$$\mathbb{P}[X_t = 0 \mid X_0 = 0] = \frac{1}{n} + \underbrace{\frac{1}{n}\sum_{j=1}^{n-1}\left(\frac{1+2\cos(2\pi j/n)}{3}\right)^t}$$

- We saw earlier: If $t = o(n^2)$, then this term is not small, it's $\omega\left(\frac{1}{n}\right)$.
- We just saw: If $t = \Omega(n^2 \log n)$, then this term is indeed small, $o\left(\frac{1}{n}\right)$.

How big does t have to be before this is small?

CLAIM: t should be about n^2 before this starts to get small.

Question to ponder

- Today, we saw that if we can diagonalize the transition matrix, it makes analyzing a Markov chain easier.
- How general can you make this? What can you say about the behavior of a (symmetric) Markov chain in terms of the eigenvalues of its transition matrix?