

Randomized Algorithms

Markov Chains and Random Walks

Joseph Chuang-Chieh Lin
Dept. CSIE, Tamkang University

Outline

- Markov Chains: definitions and representations
- Application: Random Walk
- Classification of States
- Stationary Distribution

Stochastic Process

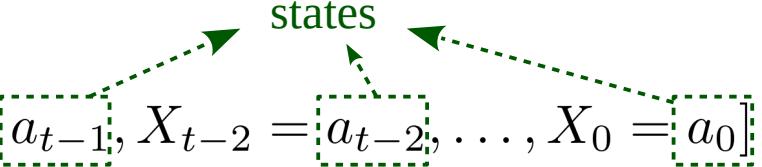
- A stochastic process $\mathbf{X} = \{X(t) : t \in T\}$ is a collection of random variables.
 - t : time
 - $X(t)$: state of the process at time t .
- If T is a countably infinite set, we say \mathbf{X} is a discrete time process.

Markov Chain

- A discrete time stochastic process X_0, X_1, X_2, \dots is a **Markov chain** if

$$\Pr[X_t = a_t \mid X_{t-1} = a_{t-1}, X_{t-2} = a_{t-2}, \dots, X_0 = a_0] = \Pr[X_t = a_t \mid X_{t-1} = a_{t-1}] = P_{a_{t-1}, a_t}.$$

states



- Markov property.

Markov Chain

- A discrete time stochastic process X_0, X_1, X_2, \dots is a **Markov chain** if

$$\begin{aligned}\Pr[X_t = a_t \mid X_{t-1} = a_{t-1}, X_{t-2} = a_{t-2}, \dots, X_0 = a_0] &= \Pr[X_t = a_t \mid X_{t-1} = a_{t-1}] \\ &= P_{a_{t-1}, a_t}.\end{aligned}$$

- Markov property.
- ✓ This does **NOT** imply that X_t is **independent** of X_0, X_1, \dots, X_{t-2} ,
 - ✓ The dependency of X_t on the past is captured in X_{t-1} .

Markov Chain

- Markov property implies:
→ The Markov chain is uniquely defined by the one-step transition matrix.

$$\mathbf{P} = \left\{ \begin{array}{cccccc} P_{0,0} & P_{0,1} & \cdots & P_{0,j} & \cdots \\ P_{1,0} & P_{1,1} & \cdots & P_{1,j} & \cdots \\ \vdots & \vdots & \ddots & \vdots & \ddots \\ P_{i,0} & P_{i,1} & \cdots & P_{i,j} & \cdots \\ \vdots & \vdots & \ddots & \vdots & \ddots \end{array} \right\}$$

for all i , $\sum_{j \geq 0} P_{i,j} = 1$.

Transition Probabilities

- $\bar{p}(t) = (p_0(t), p_1(t), p_2(t), \dots)$.
 - $p_i(t)$: the probability that the process is at state i at time t .

$$p_i(t) = \sum_{j \geq 0} p_j(t-1) P_{j,i}.$$

$$\bar{p}(t) = \bar{p}(t-1) \mathbf{P}.$$

Transition Probabilities

- $\bar{p}(t) = (p_0(t), p_1(t), p_2(t), \dots)$.
 - $p_i(t)$: the probability that the process is at state i at time t .

$$p_i(t) = \sum_{j \geq 0} p_j(t-1) P_{j,i}.$$

$$\bar{p}(t) = \bar{p}(t-1) \mathbf{P}.$$

- m step transition probability:

$$P_{i,j}^m = \Pr[X_{t+m} = j \mid X_t = i].$$

$$P_{i,j}^m = \sum_{k \geq 0} P_{i,k} P_{k,j}^{m-1}.$$

Transition Probabilities

- $\bar{p}(t) = (p_0(t), p_1(t), p_2(t), \dots)$.
 - $p_i(t)$: the probability that the process is at state i at time t .

$$p_i(t) = \sum_{j \geq 0} p_j(t-1) P_{j,i}.$$

$$\bar{p}(t) = \bar{p}(t-1) \mathbf{P}.$$

$$\mathbf{P}^{(m)} = \mathbf{P} \cdot \mathbf{P}^{(m-1)}$$

$\mathbf{P}^{(m)} = \mathbf{P}^m$ (by induction on m)

- m step transition probability:

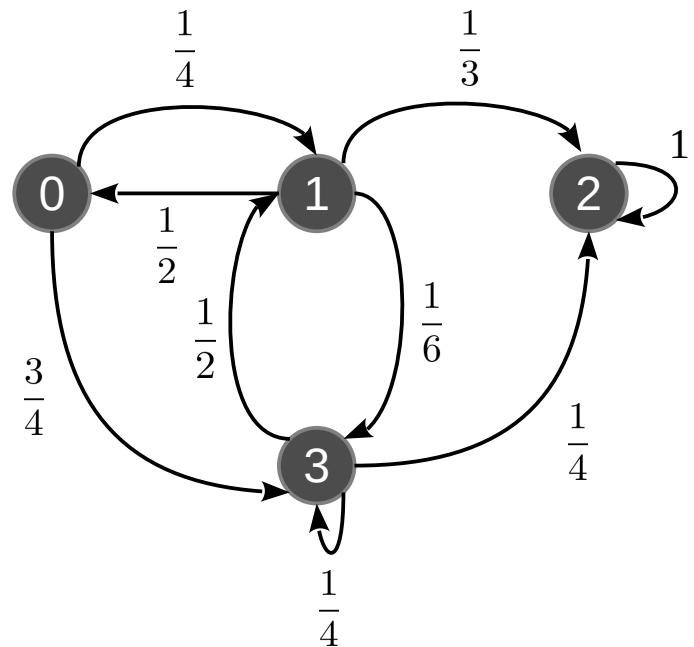
$$P_{i,j}^m = \Pr[X_{t+m} = j \mid X_t = i].$$

$$P_{i,j}^m = \sum_{k \geq 0} P_{i,k} P_{k,j}^{m-1}.$$

for any $t \geq 0$ and $m \geq 1$,

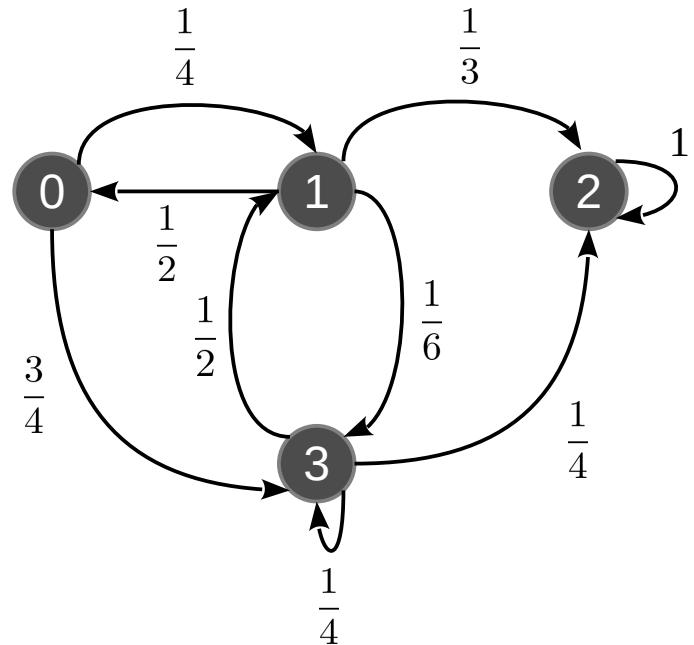
$$\bar{p}(t+m) = \bar{p}(t) \mathbf{P}^m.$$

Transition Probabilities



$$\mathbf{P} = \begin{bmatrix} 0 & 1/4 & 0 & 3/4 \\ 1/2 & 0 & 1/3 & 1/6 \\ 0 & 0 & 1 & 0 \\ 0 & 1/2 & 1/4 & 1/4 \end{bmatrix}$$

Transition Probabilities



$$\mathbf{P} = \begin{bmatrix} 0 & 1/4 & 0 & 3/4 \\ 1/2 & 0 & 1/3 & 1/6 \\ 0 & 0 & 1 & 0 \\ 0 & 1/2 & 1/4 & 1/4 \end{bmatrix}$$

$$\mathbf{P}^3 = \begin{bmatrix} 3/16 & 7/48 & 29/64 & 41/192 \\ 5/48 & 5/24 & 79/144 & 5/36 \\ 0 & 0 & 1 & 0 \\ 1/16 & 13/96 & 107/192 & 47/192 \end{bmatrix}$$

Transition Probabilities

- If we begin in a state chosen uniformly at random: $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$, what is the probability distribution after three steps?

Transition Probabilities

- If we begin in a state chosen uniformly at random: $(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$, what is the probability distribution after three steps?

$$(1/4, 1/4, 1/4, 1/4) \mathbf{P}^3 = (17/192, 47/384, 737/1152, 43/288).$$

$$\mathbf{P}^3 = \begin{bmatrix} 3/16 & 7/48 & 29/64 & 41/192 \\ 5/48 & 5/24 & 79/144 & 5/36 \\ 0 & 0 & 1 & 0 \\ 1/16 & 13/96 & 107/192 & 47/192 \end{bmatrix}$$

Exercise

- Consider the two-state Markov chain with the following transition matrix. Find a simple expression for $P_{0,0}^t$.

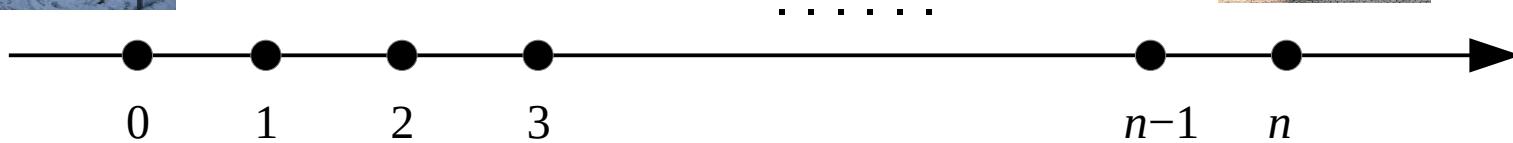
$$\mathbf{P} = \begin{bmatrix} p & 1-p \\ 1-p & p \end{bmatrix}$$

Application: Random Walks

Steppenberglallee Aachen



Aachener Dom



Application: Random Walks

Steppenberglallee Aachen



Aachener Dom



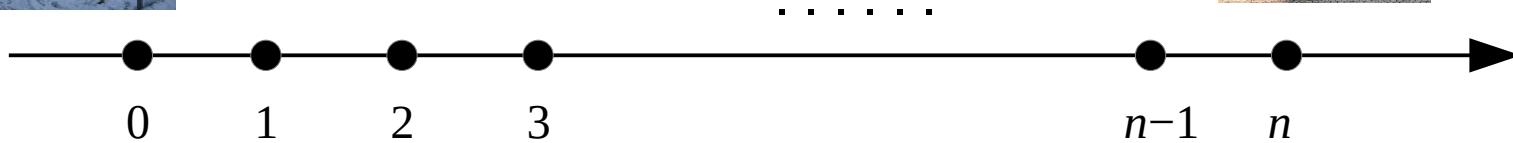
- X_i : the position after the i th step you've walked.

Application: Random Walks

Steppenberglallee Aachen



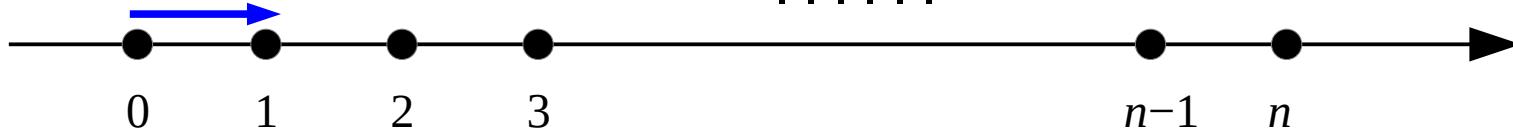
Aachener Dom



- Only at the position 0 (my home) we know how to make a right step towards the destination (cathedral).

Application: Random Walks

Steppenberglallee Aachen



Aachener Dom



- Only at the position 0 (my home) we know how to make a right step towards the destination (cathedral).

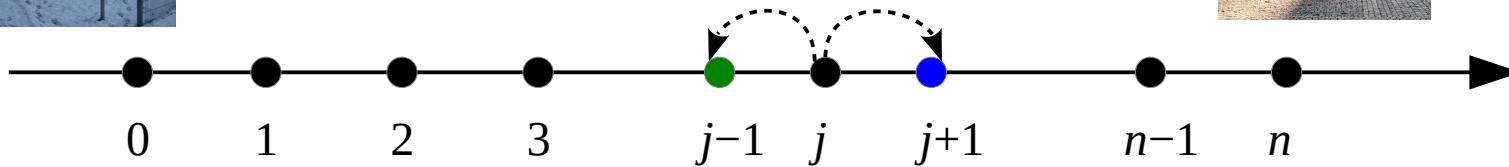
$$\Pr[X_{i+1} = 1 \mid X_i = 0] = 1.$$

Application: Random Walks

Steppenbergsallee Aachen



Aachener Dom



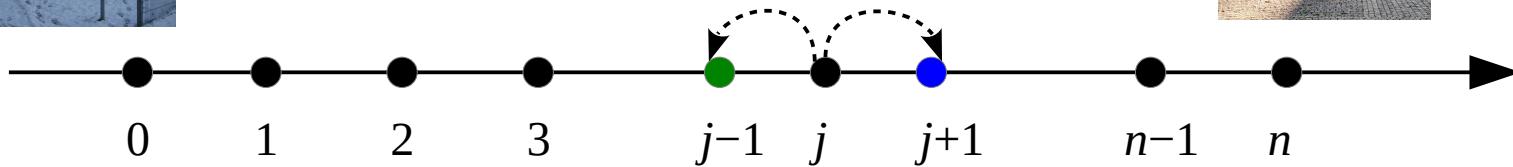
- If we are at positions $1, 2, \dots, n-1$, we have no idea about the direction to go.
- Suppose then we have chance of 50% to get one step **closer** to the destination and 50% to get one step **backward**...
- How many steps we expect to walk...?

Application: Random Walks

Steppenbergsallee Aachen



Aachener Dom



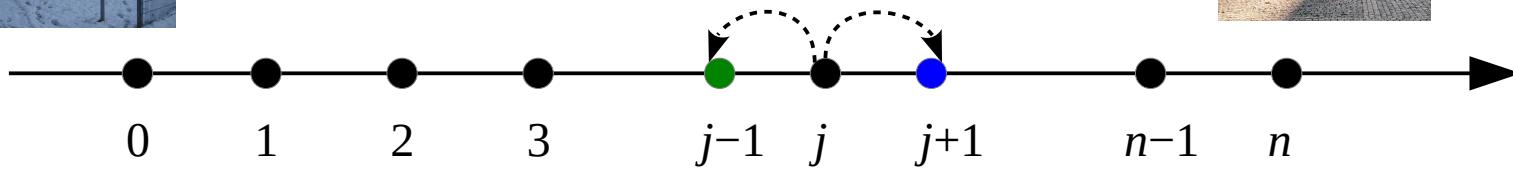
- Markov chain X_0, X_1, X_2, \dots
- Z_j : random variable; the number of steps to reach n from j .
- h_j : the expected steps to reach n when starting from j .
 - $\mathbb{E}[Z_j] = h_j$.
- $h_n = 0$, $h_0 = h_1 + 1$.

Application: Random Walks

Steppenbergsallee Aachen



Aachener Dom



- $\mathbf{E}[Z_j] = \mathbf{E} \left[\frac{1}{2} \cdot (1 + Z_{j-1}) + \frac{1}{2} \cdot (1 + Z_{j+1}) \right].$

$$h_j = \frac{h_{j-1}}{2} + \frac{h_{j+1}}{2} = \frac{h_{j-1} + h_{j+1}}{2} + 1. \quad (\text{for } 1 \leq j \leq n-1)$$

Application: Random Walks

$$h_{j+1} = 2h_j - h_{j-1} - 2$$

$$h_j = 2h_{j-1} - h_{j-2} - 2$$

$$h_{j-1} = 2h_{j-2} - h_{j-3} - 2$$

$$h_{j-2} = 2h_{j-3} - h_{j-4} - 2$$

...

$$h_2 = 2h_1 - h_0 - 2$$

$$h_1 = h_0 - 1$$

Application: Random Walks

$$h_{j+1} = 2h_j - h_{j-1} - 2$$

$$h_j = 2h_{j-1} - h_{j-2} - 2$$

$$h_{j-1} = 2h_{j-2} - h_{j-3} - 2$$

$$h_{j-2} = 2h_{j-3} - h_{j-4} - 2$$

...

$$h_2 = 2h_1 - h_0 - 2$$

$$h_1 = h_0 - 1$$

$$\Rightarrow h_j = h_{j+1} + 2j + 1.$$

Application: Random Walks

$$h_{j+1} = 2h_j - h_{j-1} - 2$$

$$h_j = 2h_{j-1} - h_{j-2} - 2$$

$$h_{j-1} = 2h_{j-2} - h_{j-3} - 2$$

$$h_{j-2} = 2h_{j-3} - h_{j-4} - 2$$

$$\dots \quad \dots \quad \dots$$

$$h_2 = 2h_1 - h_0 - 2$$

$$h_1 = h_0 - 1$$

$$h_0 = h_1 + 1 = h_2 + 1 + 3 = \dots = \sum_{i=0}^{n-1} (2i + 1) = n^2.$$

$$\Rightarrow h_j = h_{j+1} + 2j + 1.$$

Application: Random Walks

$$\begin{array}{lll} h_{j+1} & = 2h_j & -h_{j-1} - 2 \\ h_j & = 2h_{j-1} & -h_{j-2} - 2 \\ h_{j-1} & = 2h_{j-2} & -h_{j-3} - 2 \\ h_{j-2} & = 2h_{j-3} & -h_{j-4} - 2 \\ \dots & \dots & \dots \\ h_2 & = 2h_1 & -h_0 - 2 \\ h_1 & = h_0 & -1 \end{array}$$

$$h_0 = h_1 + 1 = h_2 + 1 + 3 = \dots = \sum_{i=0}^{n-1} (2i + 1) = n^2.$$

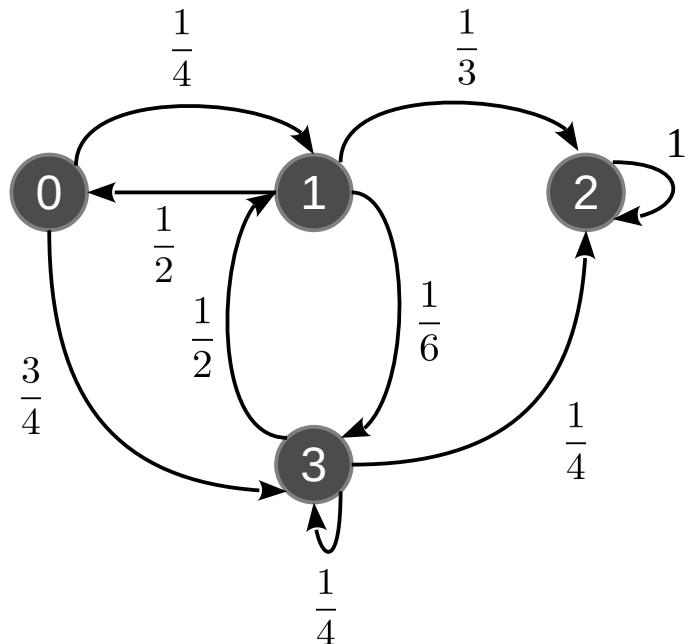
$$\Pr[\text{walking steps } > 2n^2] \leq \frac{n^2}{2n^2} = \frac{1}{2}.$$

$$\Rightarrow h_j = h_{j+1} + 2j + 1.$$

Exercise

- Consider the random walk we just discussed. Now we assume that whenever position 0 is reached, with probability $\frac{1}{2}$ the walk moves to position 1 and with probability $\frac{1}{2}$ the walk stays at 0. What is the expected number of steps to reach n starting from position 0?

Classification of States

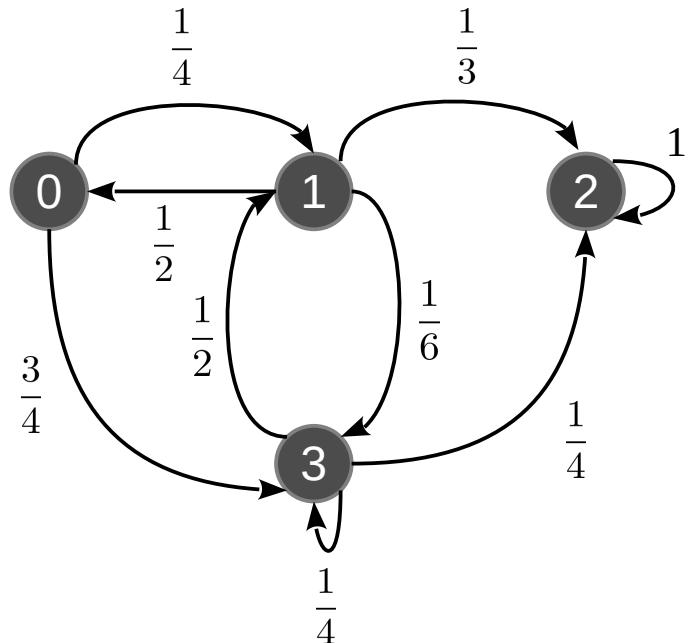


- $i \rightarrow j$ **accessible**:

For some integer $n \geq 0$, $P_{i,j}^n > 0$

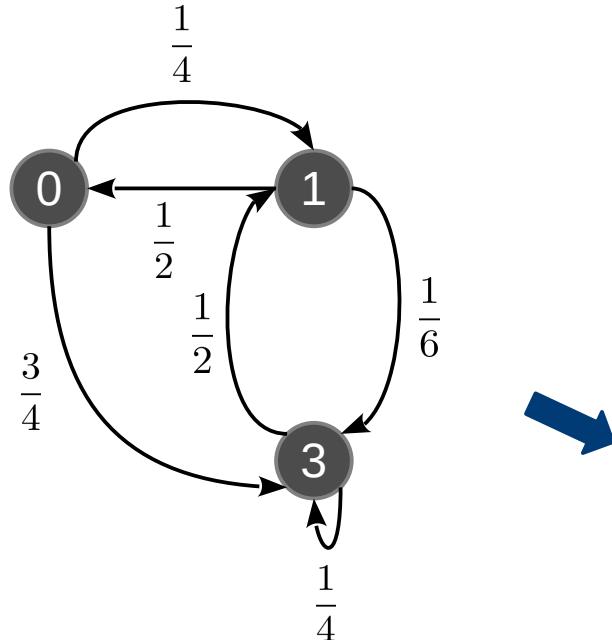
- How about $2 \rightarrow 0$? $2 \rightarrow 1$?

Classification of States



- $i \rightarrow j$ **accessible**:
For some integer $n \geq 0$, $P_{i,j}^n > 0$
 - ◆ How about $2 \rightarrow 0$? $2 \rightarrow 1$?
- $i \leftrightarrow j$: i and j **communicate**.

Classification of States

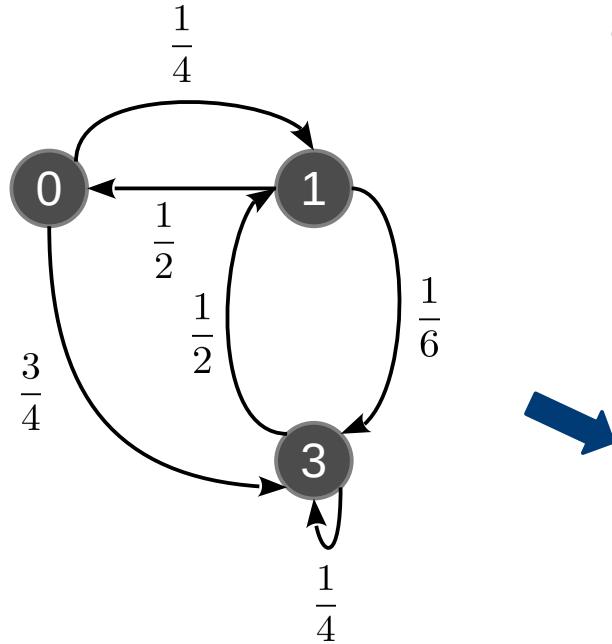


- $i \rightarrow j$ **accessible**:
For some integer $n \geq 0$, $P_{i,j}^n > 0$
 - ◆ How about $2 \rightarrow 0$? $2 \rightarrow 1$?
- $i \leftrightarrow j$: i and j **communicate**.

The Markov chain is **irreducible**.

- Any two states communicate.

Classification of States



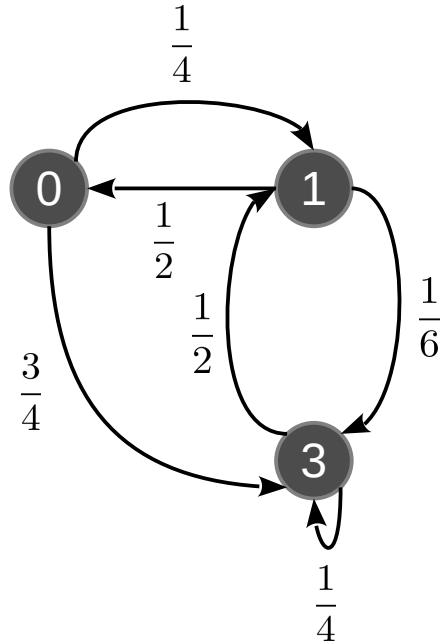
- $r_{i,j}^t$: the probability that starting at state i , the first transition to state j occurs at time t .

$$r_{i,j}^t = \Pr[X_t = j \text{ and, for } 1 \leq s \leq t-1, X_s \neq j \mid X_0 = i]$$

The Markov chain is **recurrent**.

- $\sum_{t \geq 1} r_{i,i}^t = 1$ for every state i

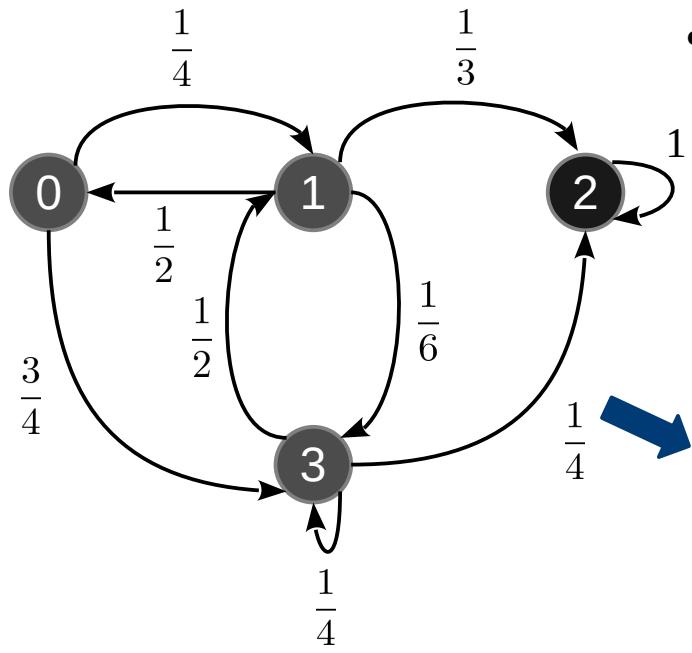
Classification of States



- $r_{i,j}^t$: the probability that starting at state i , the first transition to state j occurs at time t .
$$r_{i,j}^t = \Pr[X_t = j \text{ and, for } 1 \leq s \leq t-1, X_s \neq j \mid X_0 = i]$$
 - $h_{i,i}$: the expected time to return state i when starting from state i .
$$h_{i,i} = \sum_{t \geq 1} t \cdot r_{i,i}^t.$$
- The Markov chain is **recurrent**.
- $\sum_{t \geq 1} r_{i,i}^t = 1$ for every state i
 - Each state i is **positive recurrent**.

$$h_{i,i} < \infty$$

Classification of States



- $r_{i,j}^t$: the probability that starting at state i , the first transition to state j occurs at time t .

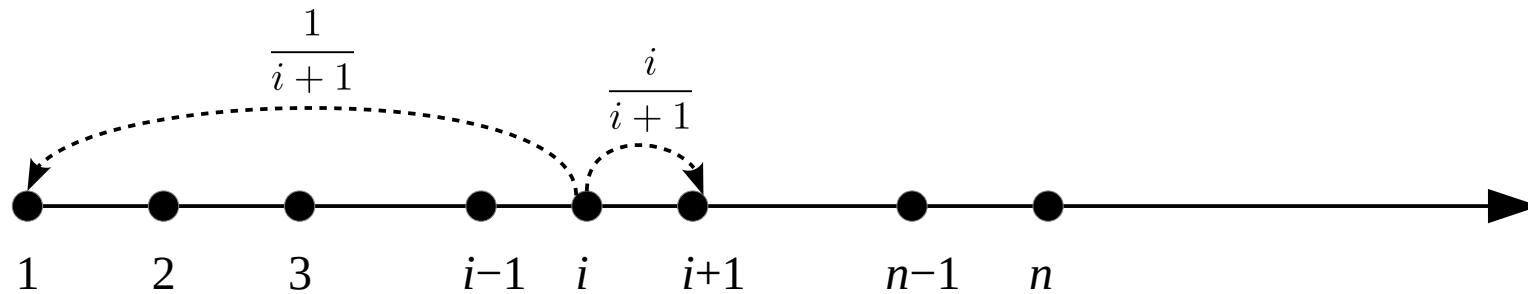
$$r_{i,j}^t = \Pr[X_t = j \text{ and, for } 1 \leq s \leq t-1, X_s \neq j \mid X_0 = i]$$

State 0, 1, 3 are **transient**.

- $\sum_{t \geq 1} r_{1,1}^t < 1$ for state $i \in \{0, 1, 3\}$

Classification of States

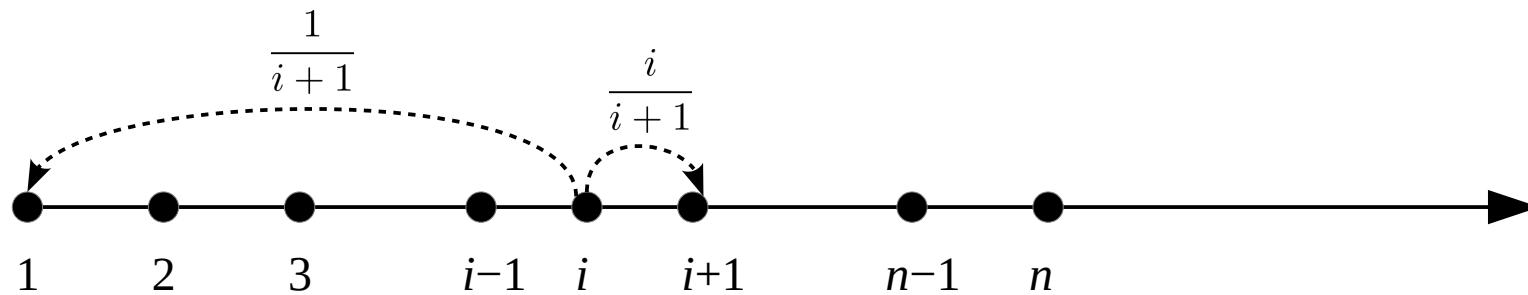
- An example of *null* recurrent:



$$r_{1,1}^t = \left(\prod_{j=1}^{t-1} \frac{j}{j+1} \right) \cdot \frac{1}{t+1} = \frac{1}{2} \cdot \frac{2}{3} \cdots \frac{t-1}{t} \cdot \frac{1}{t+1} = \frac{1}{t(t+1)}.$$

Classification of States

- An example of *null* recurrent:

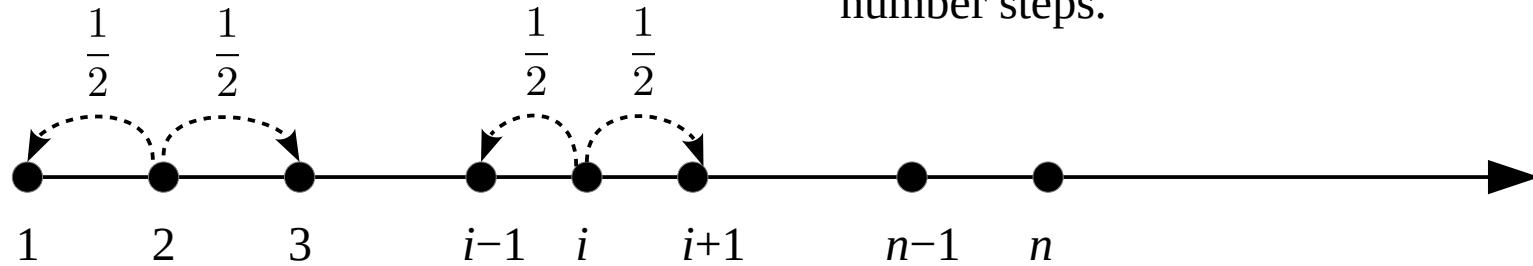


$$r_{1,1}^t = \left(\prod_{j=1}^{t-1} \frac{j}{j+1} \right) \cdot \frac{1}{t+1} = \frac{1}{2} \cdot \frac{2}{3} \cdots \frac{t-1}{t} \cdot \frac{1}{t+1} = \frac{1}{t(t+1)}.$$

$$h_{1,1} = \sum_{t \geq 1} t \cdot r_{1,1}^t = \sum_{t \geq 1} \frac{1}{t+1} \rightarrow \infty.$$

Classification of States

- **periodic** states.
 - Suppose the chain starts at 2.
 - It can be at **even** number states only after **even** number steps.

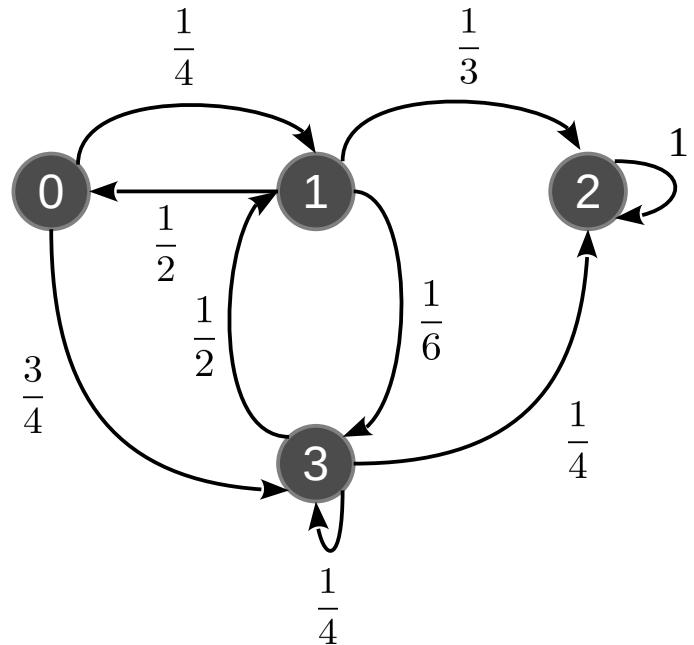


j is **periodic** :

$\exists \Delta > 1$ such that $\Pr[X_{t+s} = j \mid X_t = j] = 0$ unless s is divisible by Δ .

- *aperiodic* = not periodic

Classification of States



- An aperiodic, positive recurrent state is an **ergodic** state.
- **Ergodic Markov chain:** every state is ergodic.

Stationary Distributions

- Recall that

$$\bar{p}(t) = \bar{p}(t-1)\mathbf{P}.$$

- Consider $\bar{p}(t) = \bar{p}(t-1)$

That is, $\bar{\pi} = \bar{\pi}\mathbf{P}$.

$\bar{\pi}$: a probability distribution over the states.

Stationary Distributions

- Recall that

$$\bar{p}(t) = \bar{p}(t-1)\mathbf{P}.$$

- Consider $\bar{p}(t) = \bar{p}(t-1)$

That is, $\bar{\pi} = \bar{\pi}\mathbf{P}$.

$\bar{\pi}$: a probability distribution over the states.

- We call it a **stationary distribution** of the Markov chain.

Stationary Distributions

- **Theorem.** Any finite, irreducible, and ergodic Markov chain has the following properties:

1. The chain has a unique stationary distribution
2. for all j and i ,

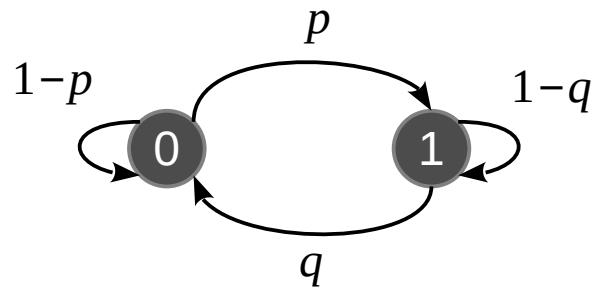
$$\lim_{t \rightarrow \infty} P_{j,i}^t \text{ exists and it's independent of } j$$

$$3. \pi_i = \lim_{t \rightarrow \infty} P_{j,i}^t = \frac{1}{h_{i,i}}.$$

Computing the Stationary Distribution

- Method 1: Solve the system of linear equations.
- Example:

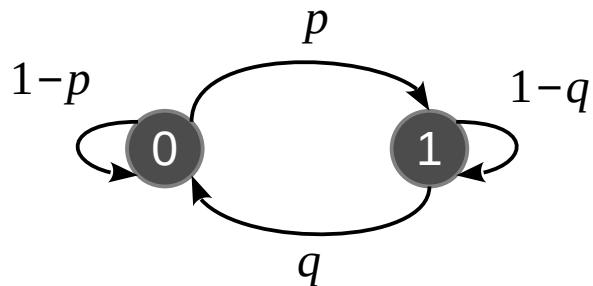
$$\mathbf{P} = \begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix}.$$



Computing the Stationary Distribution

- Method 1: Solve the system of linear equations.
- Example:

$$\mathbf{P} = \begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix}.$$



$$\bar{\pi}\mathbf{P} = \bar{\pi} \Leftrightarrow [\pi_0, \pi_1] \cdot \begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix} = [\pi_0, \pi_1].$$

$$\pi_0(1-p) + \pi_1 q = \pi_0;$$

$$\pi_0 p + \pi_1(1-q) = \pi_1;$$

$$\pi_0 + \pi_1 = 1$$

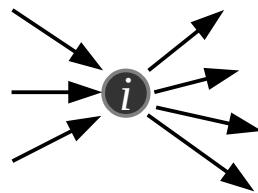
$$\pi_0 = \frac{q}{p+q}.$$

$$\pi_1 = \frac{p}{p+q}.$$

Computing the Stationary Distribution

- Method 2: Cut-sets of the Markov chain.
- The idea:
 - For any state i of the chain,

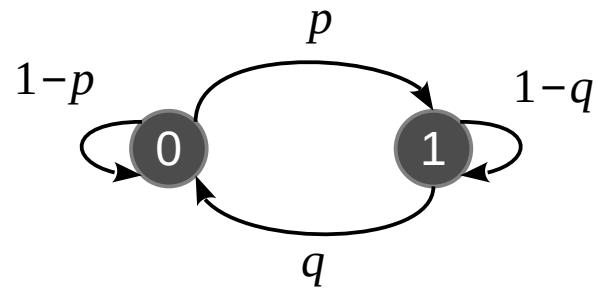
$$\sum_{j=0}^n \pi_j P_{j,i} = \pi_i = \pi_i \cdot \sum_{j=0}^n P_{i,j}$$



Computing the Stationary Distribution

- Method 2: Cut-sets of the Markov chain.
- Example:

$$\mathbf{P} = \begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix}.$$



The probability of leaving state 0 must equal the probability of entering state 0

$$\pi_0 p = \pi_1 q$$

$$\begin{aligned}\pi_0 &= \frac{q}{p+q}. \\ \pi_1 &= \frac{p}{p+q}.\end{aligned}$$

Exercise

- Consider a Markov chain with state space $\{0, 1, 2, 3\}$ and a transition matrix

$$\mathbf{P} = \begin{bmatrix} 0 & \frac{3}{10} & \frac{1}{10} & \frac{3}{5} \\ \frac{1}{10} & \frac{1}{10} & \frac{7}{10} & \frac{1}{10} \\ \frac{1}{10} & \frac{7}{10} & \frac{1}{10} & \frac{1}{10} \\ \frac{9}{10} & \frac{1}{10} & 0 & 0 \end{bmatrix}$$

Find the stationary distribution of the Markov chain.