Markov Chain I

July 28, 2022

- HW8 is optional
- HW7 is the last graded hw

Example 1a

Toss an unfair coin $\mathbb{P}(\text{Head}) = p$ for N times. What's the fraction of time for observing heads out of all outcomes?

$\mathbb{E}(\#H) = N \cdot p$

$\#H \sim \text{Binomial}(N, p)$

fraction of H $= \dfrac{Np}{N} = p$

Example 1b

Now we have two unfair coins, each is biased to either head or tail.

Coin 1: $\mathbb{P}(\mathrm{Head}) = p$;   Coin 2: $\mathbb{P}(\mathrm{Head}) = 1 - p$.

If seeing head, then use coin 1 for next toss; if seeing tail, then use coin 2 for the next toss.

Toss N times, what's the fraction of time for observing heads out of all outcomes?

Markov Chain (Discrete Time Finite MC):

State Space:

r.v.

At each time step n, the "state" is denoted by $\boldsymbol{X_n}$. The collection of all possible value a state can take is called the state space $S = \{1,2,\ldots,K\}$ for a finite number K.

Transition Probability:

next step current

$$0 \le P_{ij} = \mathbb{P}(\boldsymbol{X_{n+1}} = j \mid \boldsymbol{X_n} = i) \ \forall i,j \in S$$

$\sum_{j=1}^{K} P_{ij} = 1 \ , \forall i$; $P_{ij} \ \forall i,j$. forms a matrix $P \in \mathbb{R}^{K \times K}$

$P(i,j) = P_{ij}$

Markov Property

$$\mathbb{P}(\boldsymbol{X}_{\boldsymbol{n+1}} = j \mid \boldsymbol{X}_{\boldsymbol{n}} = i, \boldsymbol{X}_{\boldsymbol{n-1}} = i_{n-1}, \ldots, \boldsymbol{X}_{\boldsymbol{0}} = i_0)$$

$$= \mathbb{P}(\boldsymbol{X}_{\boldsymbol{n+1}} = j \mid \boldsymbol{X}_{\boldsymbol{n}} = i) = P_{ij}$$

Example 1b

Now we have two unfair coins, each is biased to either head or tail.

$$\text{Coin 1: } \mathbb{P}(\text{Head}) = p; \quad \text{Coin 2: } \mathbb{P}(\text{Head}) = 1 - p.$$

If seeing head, then use coin 1 for next toss; if seeing tail, then use coin 2 for the next toss.

Toss N times, what's the fraction of time for observing heads out of all outcomes?

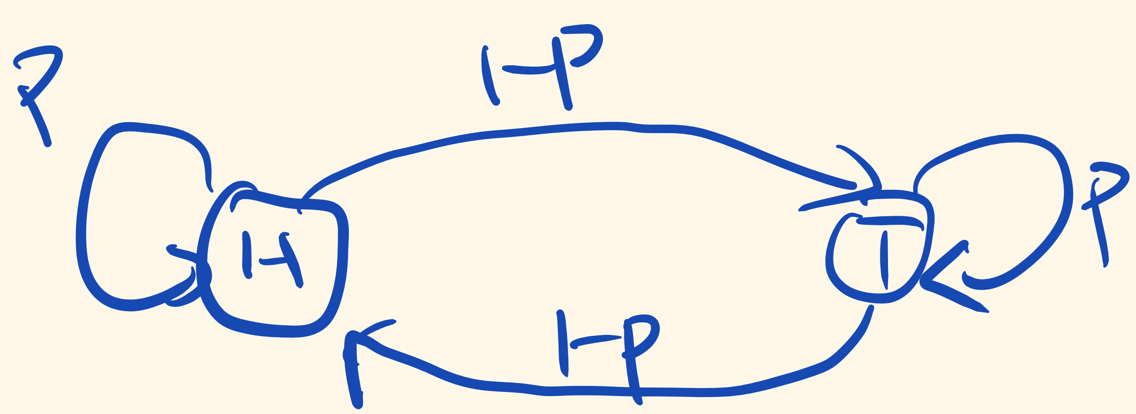
1: Head. 2: Tail

$S = \{1, 2\}$

or $S = \{H, T\}$

$P_{11} = p \quad P_{12} = 1-p$

$P_{21} = 1-p \quad P_{22} = p$



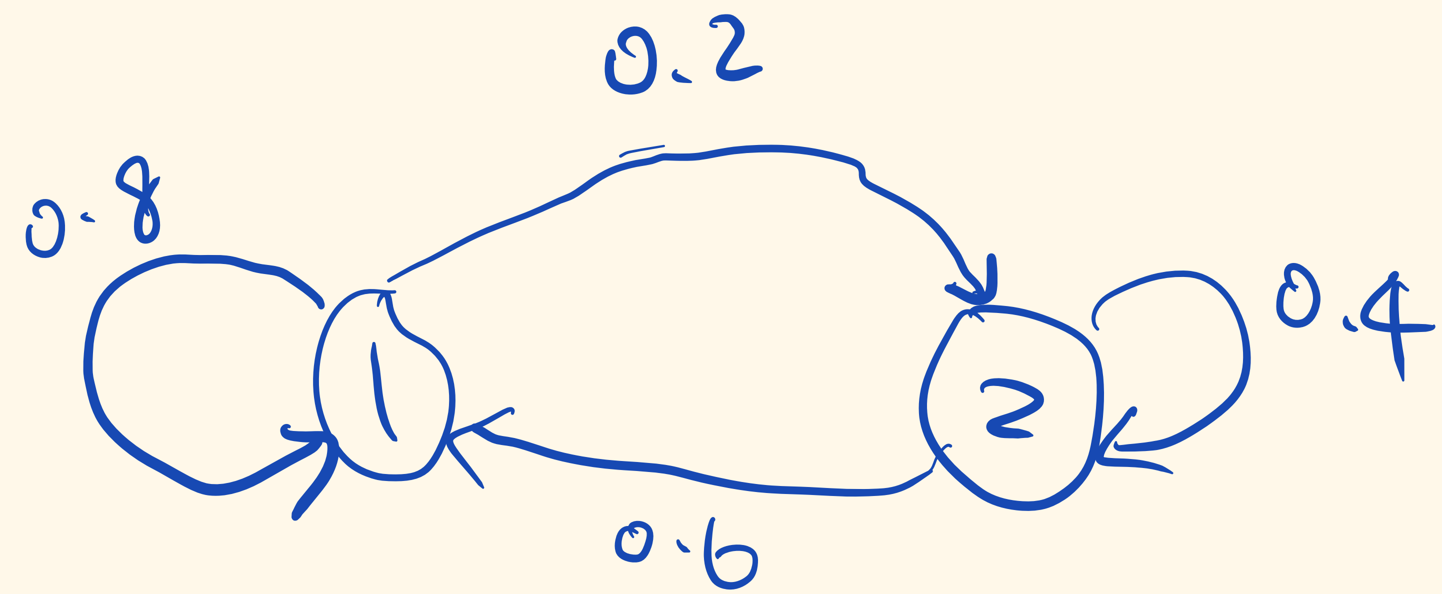
$$P = \begin{bmatrix} p & 1-p \\ 1-p & p \end{bmatrix}$$

Example 2: Alice takes probability class.

Alice is either (1) up-to-date or (2) fall behind

$$P_{11} = 0.8, \underline{P_{12} = 0.2}, P_{21} = 0.6\ P_{22} = 0.4$$

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{bmatrix}$$



Probability of being in state j, at time step n

$$\pi_n(j) := \mathbb{P}(X_n = j) \qquad \text{e.g. } \pi_n = [\pi_n(1), \pi_n(2)]$$

$$\pi_n(j) = \mathbb{P}(X_n = j) = \sum_{i=1}^{K} \mathbb{P}(X_n = j \mid X_{n-1} = i)\, \mathbb{P}(X_{n-1} = i)$$

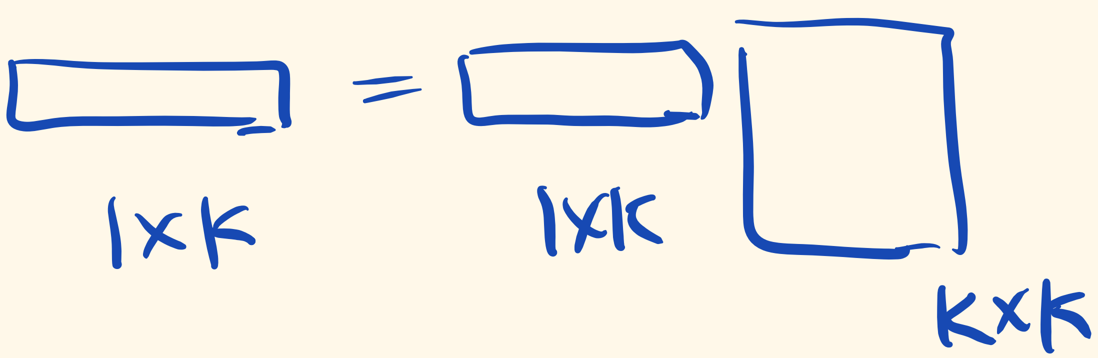
$$= \sum_{i=1}^{K} P_{ij}\, \pi_{n-1}(i) \qquad \forall j \in S$$

π_0 : initial distribution

$$\pi_0 = [\pi_0(1), \pi_0(2), \ldots, \pi_0(K)]$$

for $S = \{1, \ldots, K\}$.

$$\pi_n = \pi_{n-1} P$$



$$\pi_n = \pi_{n-1} P = \pi_{n-2} P^2 = \cdots = \pi_0 P^n$$

Balance Equation

does not depend on a specific time point.

A distribution π is invariant for the transition probability P if it satisfies the balance equation

has many names

$$\pi P = \pi \qquad \sum_j \pi(j) = 1$$

π: invariant dist, stationary dist, steady-state prob.

Thm: if $\pi_n = \pi_0$, $\forall n \geq 0$ $\Longleftrightarrow$ π_0 is invariant.

π is a row vector.

$\pi_n(i) := P(X_n = i)$

Example 2: Alice takes probability class.

Alice is either (1) up-to-date or (2) fall behind. Find the stationary distribution.

$$P_{11} = 0.8, P_{12} = 0.2, P_{21} = 0.6\ P_{22} = 0.4$$

$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{bmatrix} \qquad \pi = \pi P \Rightarrow \begin{array}{l} \pi_1 = \pi_1 P_{11} + \pi_2 P_{21} \\ \pi_2 = \pi_1 P_{12} + \pi_2 P_{22} \end{array}$$

$$\left.\begin{array}{l} \pi_1 = 0.8\pi_1 + 0.6\pi_2 \\ \pi_2 = 0.2\pi_1 + 0.4\pi_2 \end{array}\right\} \Rightarrow \begin{array}{l} \pi_1 = 3\pi_2 \\ \pi_1 + \pi_2 = 1 \end{array} \Rightarrow \begin{array}{l} \pi_1 = 0.75 \\ \pi_2 = 0.25 \end{array}$$

$$* \ \pi = \pi P, \ \pi = \pi P^2, \ \pi = \pi P^3, \ \pi(i) \geq 0, \ \sum_i \pi(i) = 1$$

Properties of Markov Chain

π

Q1: Does an invariant distribution always exist?

Q2: Is it unique?

Q3: Does π_n approach an invariant distribution?

π

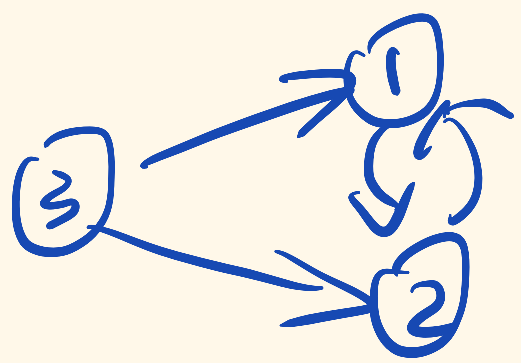
$\pi_n(j) = \mathbb{P}(X_n = j)$

Properties of Markov Chain

Irreducibility

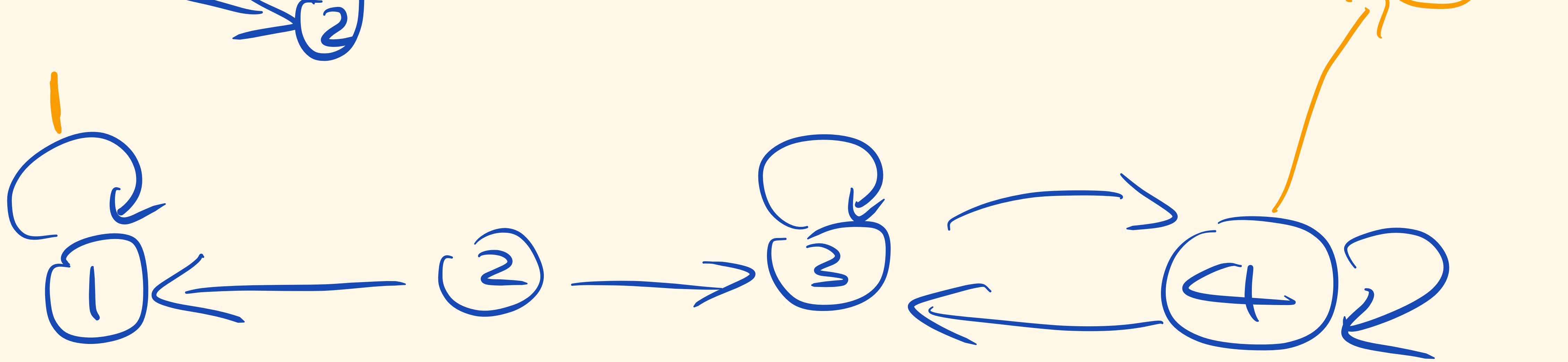
A Markov chain is irreducible if it can go from any state to any other state, possibly after many steps.

Ex.



reducible

Ex:



↓ reducible.

Properties of Markov Chain

(a)periodicity

for an irreducible Markov Chain defined on state space S with transition probability P.

Let

$$d(i) := \gcd\{n \geq 1 \mid P^n(i,i) > 0\}$$

Then, $d(i)$ has some value for all i, $d(i) = d$. ← see proof in the end

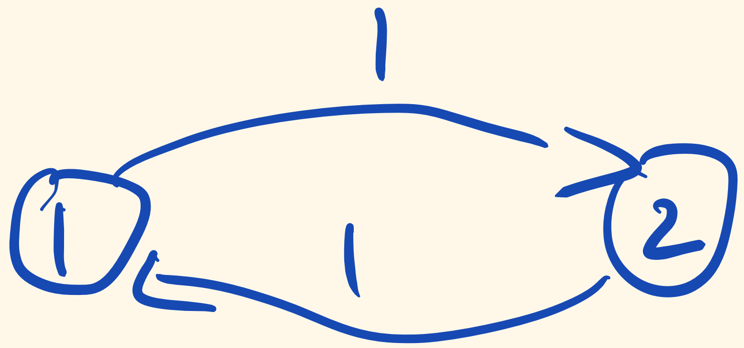
* the period of states is the same as the period of this MC. if MC is irreducible.

If d=1, MC is aperiodic

If d>1, MC is periodic with period d.

Example 3a

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



- irreducible.
- periodic. $d=2$

$$P^1 = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix} \qquad P^2 = \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} \Rightarrow P^3 = P$$

$$\pi = [0.5, 0.5]$$ ← invariant distribution

$$\pi_n = (1, 0) \qquad \pi_{n+1} = (0, 1), \quad \pi_{n+2} = (1, 0)$$

$$\lim_{n\to\infty} \pi_n \quad \times \qquad \pi_n \to \pi \quad n \to \infty$$

Example 3b: Alice studies Markov Chain

~~Find the invariant distribution of Alice's study status.~~

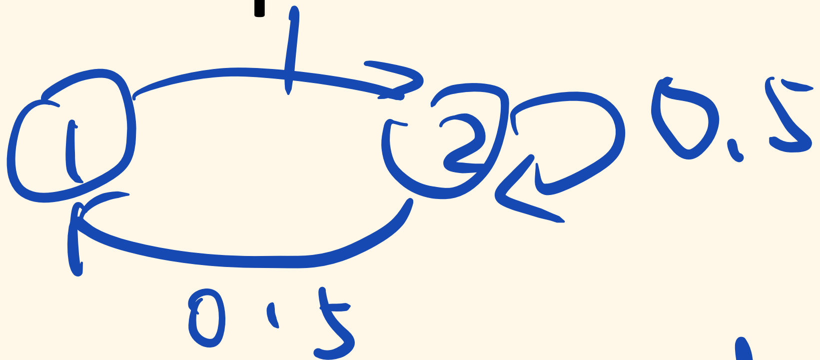
$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{bmatrix} \qquad P^n(i,i) > 0 \quad \forall n$$

$d=1$, aperiodic

$P, P^2, P^3 \ldots$ $$P^n = \begin{bmatrix} 0.75 & 0.25 \\ 0.75 & 0.25 \end{bmatrix}$$

$$\pi P = \pi \Rightarrow \pi = [0.75, 0.25]$$

Example 3c: Find the invariant distribution.



$$P = \begin{bmatrix} 0 & 1 \\ 0.5 & 0.5 \end{bmatrix}, \quad P^2 = \begin{bmatrix} 0.5 & 0.5 \\ 0.25 & 0.75 \end{bmatrix} \cdots$$

$$d = 1$$

$$\pi = \pi P \Rightarrow \pi = [0.5, 0.5]$$

A(n) irreducible Markov Chain with self-loop is aperiodic.

note: the reverse is not true,
counter example: random walk on a triangle.

Theorem for Markov Chain

does not matter if the MC is periodic or not

1) If a Markov Chain is finite and irreducible:

it has $\boxed{\text{unique}}$ invariant distribution π and

$\pi(i)$ long ~~term~~ fraction of time.

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{I}\{\{X_n = i\}\} = \pi(i)$$

$\leftarrow$ invariant distribution

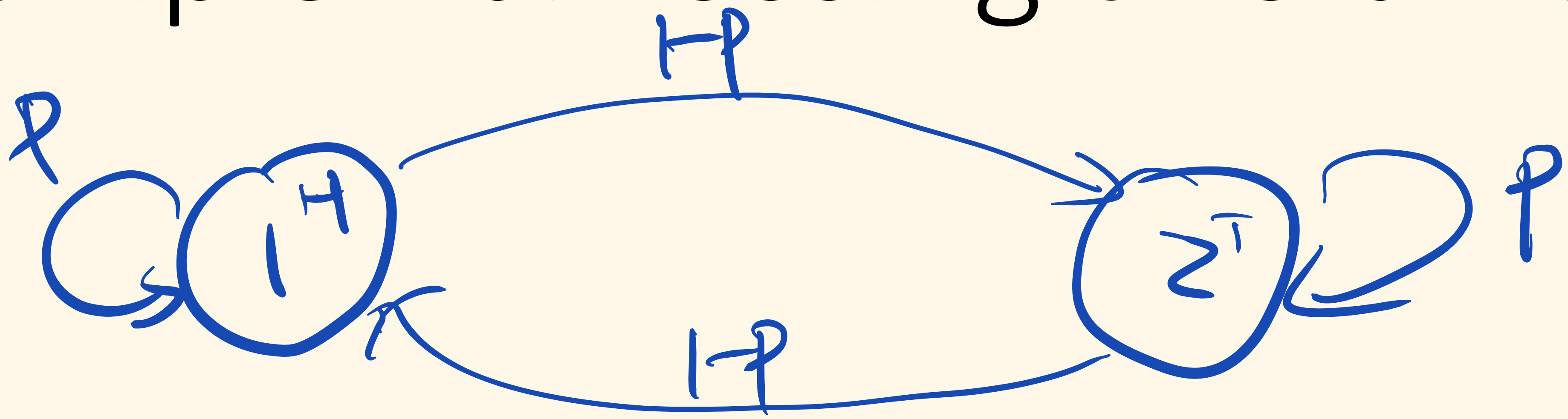
2) If this Markov Chain is also aperiodic:

then $\pi_n \to \pi$, $n \to \infty$

π_n. π_{n+1}, $\pi_{n+2} = \pi_n$, $\pi_{n+3} = \pi_{n+1}$

$d=2$

Example 1b. Tossing two unfair coins

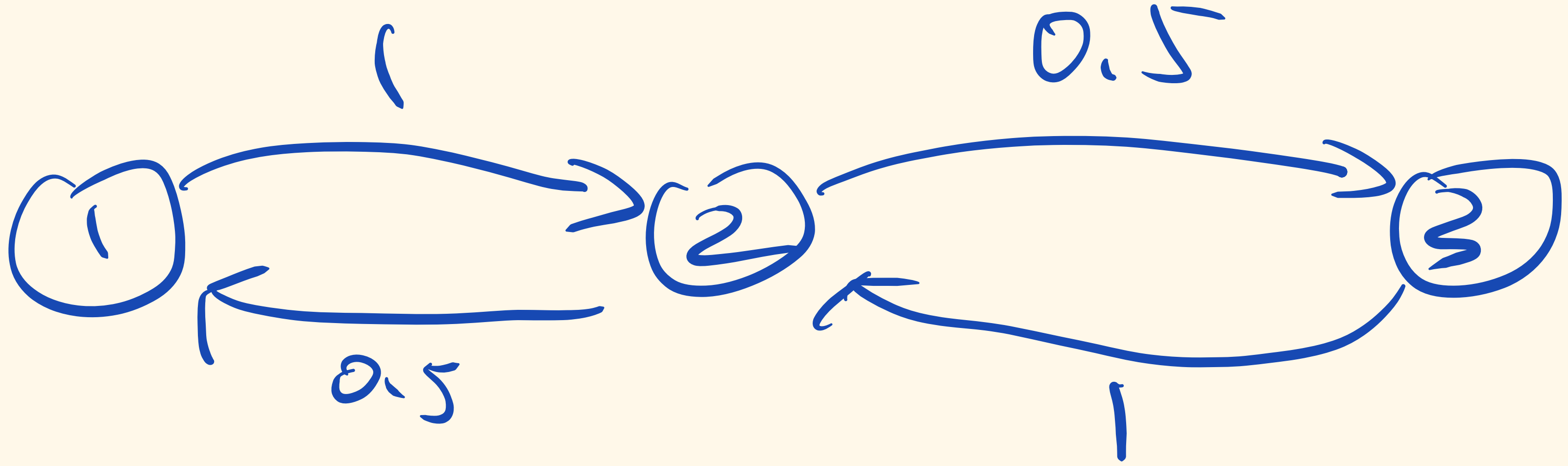


$$\pi = [0.5, 0.5]$$

One more example

$$P = \begin{bmatrix} 0 & 1 & 0 \\ 0.5 & 0 & 0.5 \\ 0 & 1 & 0 \end{bmatrix}$$

$d = 2$



P^2, $P^3 = P$

this is a periodic MC, but we can still define the stationary distribution π

$\pi = \pi P \Rightarrow \pi = [\frac{1}{4}, \frac{1}{2}, \frac{1}{4}]$, long term fraction of time at state 2 is $1/2$.

For people who are interested, this is not required for exams.

lemma: for an irreducible MC.

$d(i) := \gcd\{n \geq 1 \mid P^n(i,i) > 0\}$ has same value

for all state i.

proof: pick $\forall j \neq i$, we show that $d(j) \leq d(i)$

and the flip j & i, $d(i) \leq d(j)$ so $d(i) = d(j)$

MC is irreducible, so $\exists\, m$, $P^m(j,i) > 0$ & $\exists\, n$ $P^n(i,j) > 0$

now by definition of $d(i)$

$\exists\, N$, such that $P^N(i,i) > 0$ and $P^{N+d(i)}(i,i) > 0$

$P^{m+N+n}(j,j) > 0$ and $P^{m+N+d(i)+n}(j,j) > 0$

$\Rightarrow \underbrace{m+N+n}_{k}, \underbrace{m+N+d(i)+n}_{k+d(i)} \in \{n \geq 1 \mid P^n(j,j) > 0\}$

so $d(j) \leq d(i)$ because $\gcd\{k, k+d(i), \cdots\} \leq d(i)$

reverse i, j, do it again we have $d(i) \leq d(j)$

so $d(i) = d(j)$ $\forall\, i \neq j$ $\blacksquare$