Markov Chains

- Consider a set of states $S = \{1, 2, \dots, k\}$
- Assume states evolve over time in discrete time steps
- Let X_n = state at time $n = 0, 1, 2, \dots$

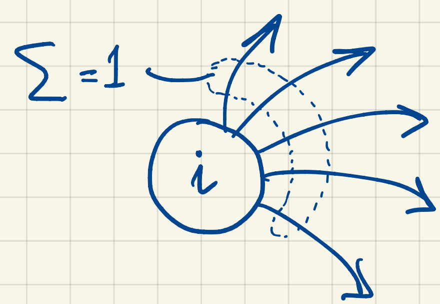
We say we have a Markov Chain iff $(m < n)$:

$$P(X_n = j \mid X_0, X_1, X_2, \dots, X_m = i) = P(X_n = j \mid X_m = i)$$

(Conditioning on X_m, makes X_n independent of $X_0, \dots, X_{m-1}$)

Furthermore, when $m = n-1$, above probability is given by P_{ij} and is called the transition prob. from state i to state j

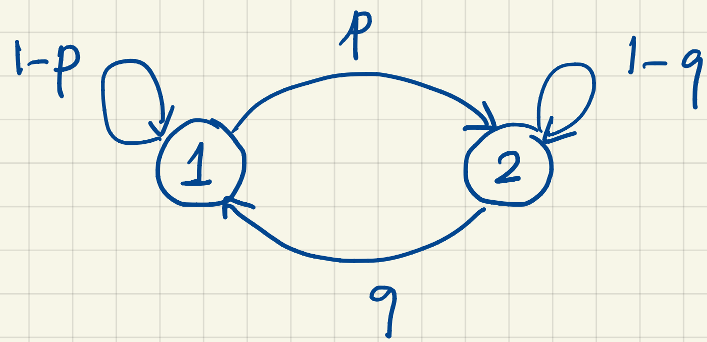
$$\sum_j P_{ij} = 1$$



Matrix Representation

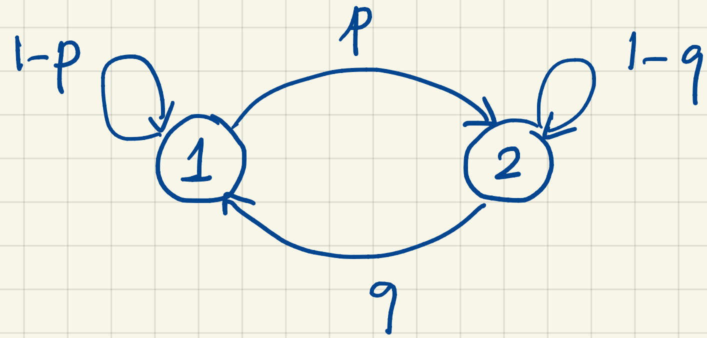
Here's a simple Markov chain given as a matrix and illustrated by a state diagram

$$P = \underbrace{\begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix}}_{\text{All the transition probabilities}}$$



$P_{11} = 1-p$ $P_{12} = p$ $P_{21} = q$ $P_{22} = 1-q$

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



Let's do some calculations: What's $P(X_2 = 2 \mid X_0 = 1)$?

"Probability of being in state 2 at time 2 given we are in state 1 at time 0"

Visual interpretation:

$$1 \xrightarrow{1-p} 1 \xrightarrow{p} 2$$

$$1 \xrightarrow{p} 2 \xrightarrow{1-q} 2$$

$$(1-p)p + p(1-q)$$

We have "visually" used the Markov property. (see next)

$$P(X_2=2 \mid X_0=1) = \sum_i P(X_2=2 \mid X_1=i, X_0=1)\, P(X_1=i \mid X_0=1)$$

$$= P(X_2=2 \mid X_1=1, X_0=1)\, P(X_1=1 \mid X_0=1) + P(X_2=2 \mid X_1=2, X_0=1)\, P(X_1=2 \mid X_0=1)$$

$$= P(X_2=2 \mid X_1=1)\, P(X_1=1 \mid X_0=1) + P(X_2=2 \mid X_1=2)\, P(X_1=2 \mid X_0=1)$$

$$= p(1-p) + (1-q)p$$

Observation: this is the entry in first row second colum of P^2

$$P^2 = \begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix} \times \begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix} = \begin{bmatrix} - & (1-p)p + p(1-q) \\ - & - \end{bmatrix}$$

In general $P(X_n = j \mid X_m = i) = P_{ij}^{\,n-m}$ for $n > m$.

Two Properties

Irreducible and aperiodic

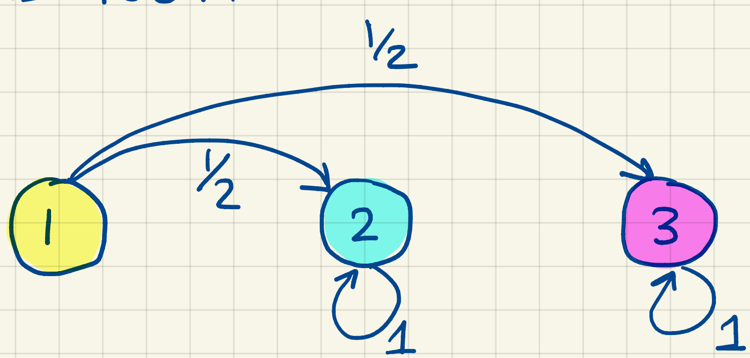
- The matrix P is irreducible iff:

$$\forall\, i,j .\ \exists\, n .\ P(X_n = j \mid X_0 = i) > 0$$

- In terms of the state diagram, this means there is a path from i to j (not necessarily an immediate transition)

- The previous Markov Chain is irreducible.

- The one below is NOT.



$$\begin{bmatrix} 0 & 1/2 & 1/2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

What's important about irreducibility?

Irreducible $\Rightarrow$ $\exists$ unique stationary distribution of states

Unique Vector $\pi \neq 0$ $\left(\sum \pi_i = 1\right)$

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

Why called stationary?

$$\begin{bmatrix} \pi_1 & \pi_2 & \cdots & \pi_i & \cdots & \pi_k \end{bmatrix} \begin{bmatrix} & P_{1i} & \\ & P_{2i} & \\ & \vdots & \\ & P_{ki} & \end{bmatrix} = \begin{bmatrix} \cdots & \pi_i & \cdots \end{bmatrix}$$

$$\sum_j \pi_j P_{ji} = \pi_i$$

$$\pi_1 P_{1i} + \pi_2 P_{2i} + \cdots + \pi_i P_{ii} + \cdots + \pi_k P_{ki} = \pi_i$$

- If π is the prob. distribution over states at time n, then the left hand side is the prob. of being in state i at time $n+1$.
- If this is π_i, then the prob. of state i has not changed.
- This is true for all i.

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

This is irreducible if $p, q > 0$

$$\pi = \left[\frac{q}{p+q} , \frac{p}{p+q} \right]$$

$$P = \begin{bmatrix} 0 & 1/2 & 1/2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

This is not irreducible

any $\pi = [0 \quad p \quad 1-p]$ is stationary

(not unique)

Is is possible to have no stationary distribution? Not if we have a finite set of states. There has to be at least one.

How do we find the stationary dist. ?

Solve a system of linear equations.

Example:
$$[\pi_1 \;\; \pi_2]\begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix} = [\pi_1 \;\; \pi_2]$$

① $\pi_1(1-p) + \pi_2 q = \pi_1$

$\pi_1 p + \pi_2(1-q) = \pi_2$ (same equation!)

add: $\pi_1 + \pi_2 = 1 \Rightarrow \pi_2 = 1 - \pi_1$ ②

① $\Rightarrow$ $\pi_1(1-p) + (1-\pi_1)q = \pi_1$

$\pi_1(1-p) - \pi_1 q - \pi_1 = -q$

$\pi_1(\cancel{1} - p - q - \cancel{1}) = -q \Rightarrow \pi_1(p+q) = q \Rightarrow \pi_1 = \dfrac{q}{p+q}$

K equations, 1 for each $i \in \{1, 2, \dots, k\}$

$$i: \sum_{j=1}^{k} \pi_j P_{ji} = \pi_i$$

If we sum the first $k-1$ equations, we obtain the last

$$\sum_{i=1}^{k-1} \sum_{j=1}^{k} \pi_j P_{ji} = \sum_{i=1}^{k-1} \pi_i$$

$$\sum_{j=1}^{k} \sum_{i=1}^{k-1} \pi_j P_{ji} = \sum_{i=1}^{k-1} \pi_i$$

$$\sum_{j=1}^{k} \pi_j \sum_{i=1}^{k-1} P_{ji} = \sum_{i=1}^{k-1} \pi_i$$

$$\sum_{j=1}^{k} \pi_j \left(1 - P_{jk}\right) = \sum_{i=1}^{k-1} \pi_i$$

$$\sum_{j=1}^{k} \pi_j - \sum_{j=1}^{k} \pi_j P_{jk} = \sum_{i=1}^{k-1} \pi_i$$

$$\pi_k - \sum_{j=1}^{k} \pi_j P_{jk} = 0$$

$$\sum_{j=1}^{k} \pi_j P_{jk} = \pi_k \qquad (k^{th} \text{ equation})$$

In general, solve:

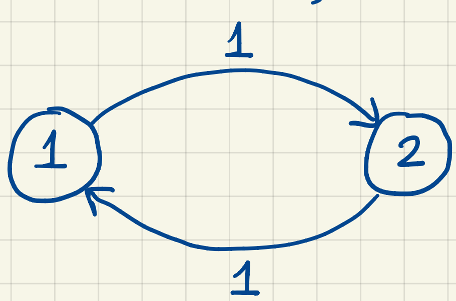
$$\begin{bmatrix} \pi_1 & \pi_2 & \cdots & \pi_{k-1} & \pi_k \end{bmatrix} \underbrace{\begin{bmatrix} & & 1 \\ & & 1 \\ & & \vdots \\ & & 1 \end{bmatrix}}_{P} = \begin{bmatrix} \pi_1 & \pi_2 & \cdots & \pi_{k-1} & 1 \end{bmatrix}$$

The matrix P is aperiodic iff:

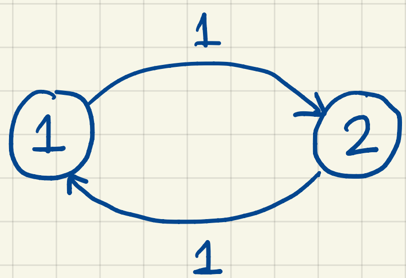
$$\forall i.\ \gcd\{n : P(X_n = i \mid X_0 = i) > 0\} = 1$$

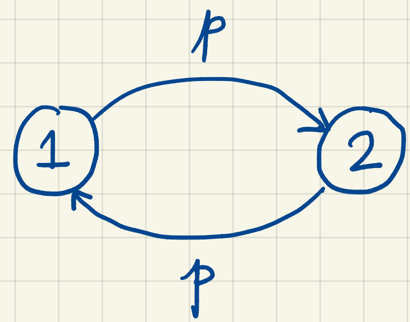
What does that mean? If we consider the lengths of all cycles that brings us back to state i, those lengths have no common divisor (except 1).

This chain is irreducible, but NOT aperiodic



Let's compare :





$p<1$

Both are irreducible with $\pi = [\frac{1}{2} \ \ \frac{1}{2}]$ as the stationary dist.

What if we start in the states with prob. $[\frac{1}{3} \ \ \frac{2}{3}]$?

What will happen next ?

$$[\tfrac{1}{3} \ \ \tfrac{2}{3}] \underbrace{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}}_{\text{periodic}} = [\tfrac{2}{3} \ \ \tfrac{1}{3}] \qquad [\tfrac{2}{3} \ \ \tfrac{1}{3}] \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = [\tfrac{1}{3} \ \ \tfrac{2}{3}]$$

No convergence to $[\frac{1}{2} \ \ \frac{1}{2}]$

$$\begin{bmatrix} q & 1-q \end{bmatrix} \underbrace{\begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix}}_{\text{aperiodic}} = \begin{bmatrix} \underbrace{q(1-p)+(1-q)p}_{q'} & \underbrace{qp+(1-q)(1-p)}_{1-q'} \end{bmatrix}$$

$$\begin{bmatrix} q' & 1-q' \end{bmatrix} \begin{bmatrix} 1-p & p \\ p & 1-p \end{bmatrix} = \begin{bmatrix} q'' & 1-q'' \end{bmatrix} \quad \ldots$$

At the limit, we are going to converge to $[1/2 \;\; 1/2]$

<u>Lesson:</u>

- If you start the system in the stationary dist., you will stay in the stationary dist.

- If you start the system arbitrarily, you may not reach the stationary dist. But you will if aperiodic!

For most practical purposes, we only need irreducibility.

- If a Markov chain is irreducible, then

$$\frac{1}{n+1}\sum_{t=0}^{n} f(x_t) \xrightarrow[n\to\infty]{} \sum_i \pi_i f(x_i)$$

- In addition, if chain is aperiodic, then

$$\lim_{n\to\infty} P^n = \begin{bmatrix} \pi_1 & \cdots & \pi_k \\ \pi_1 & \cdots & \pi_k \\ & \vdots & \\ \pi_1 & \cdots & \pi_k \end{bmatrix} \qquad \lim_{n\to\infty} P(X_n = i \mid \underset{\substack{\uparrow \\ \text{No effect}}}{x_0}) = \pi_i$$