Discrete Time Markov Chains

EECS 126 Fall 2019

October 15, 2019

Agenda

Announcements

- Homework 7 due Tomorrow night (10/16)!
- Lab self-grades due on Monday night (10/21).

Recap of Discrete Time Markov Chains

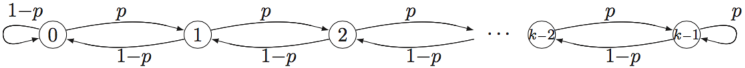


Figure: Example of a Markov chain

- State changes at discrete times
- State X_n belongs to a finite set S (for now)
- Satisfies the Markov property for transitions from state $i \in S$ to state $j \in S$

$$\mathbb{P}(X_{n+1} = j \mid X_n = i, X_{n-1} = x_{n-1} \dots X_1 = x_1)$$

$$= \mathbb{P}(X_{n+1} = j \mid X_n = i) = p_{ij}$$

  Where, $p_{ij} \geq 0, \sum_j p_{ij} = 1$
- Time homogeneous: the evolution of the system or transition probabilities are time independent

Recap of Discrete Time Markov Chains

The probability transition matrix $\boldsymbol{P}$ contains all the information about transitions between different states

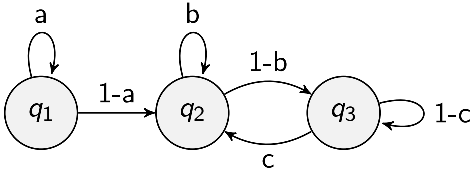
$$\boldsymbol{P} = \begin{bmatrix} p_{11} & p_{12} & \dots & p_{1n} \\ p_{21} & p_{22} & \dots & p_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ p_{n1} & p_{n2} & \dots & p_{nn} \end{bmatrix}$$

Let $\pi^{(n)} = \begin{bmatrix} \mathbb{P}(X_n = 1) & \dots & \mathbb{P}(X_n = m) \end{bmatrix}$ then

$$\pi^{(n+1)} = \pi^{(n)} \boldsymbol{P}$$

$$\Rightarrow \pi^{(n)} = \pi^{(0)} \boldsymbol{P}^n$$

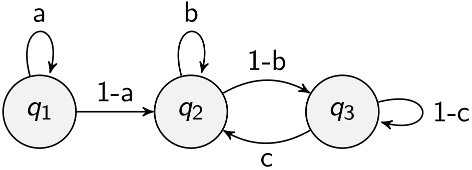
Example



$$\pi_0 = \begin{bmatrix} 0.3 & 0.3 & 0.4 \end{bmatrix}$$

- Write the probability transition matrix $\boldsymbol{P}$
- What is $\mathbb{P}(X_0 = q_1, X_1 = q_2, X_3 = q_1)$?
- What is $\mathbb{P}(X_0 = q_1, X_1 = q_1, X_2 = q_2, X_3 = q_3, X_4 = q_3)$?

Answers



- $\boldsymbol{P} = \begin{bmatrix} a & 1-a & 0 \\ 0 & b & 1-b \\ 0 & c & 1-c \end{bmatrix}$
- $\mathbb{P}(X_0 = q_1, X_1 = q_2, X_3 = q_1) = 0$. You cannot go to q_1 from q_2.
- Use the Markov Property.

$$\begin{aligned} &\mathbb{P}(X_0 = q_1, X_1 = q_1, X_2 = q_2, X_3 = q_3, X_4 = q_3) \\ &\quad = \mathbb{P}(X_0 = q_1) \cdot \mathbb{P}(X_1 = q_1 \mid X_0 = q_1) \ldots \mathbb{P}(X_4 = q_3 \mid X_3 = q_3) \\ &\quad = 0.3 \cdot a \cdot (1-a) \cdot (1-b) \cdot (1-c) \end{aligned}$$

n-step Transition Probabilities

Let $r_{ij}(n) = \mathbb{P}(X_n = j \mid X_0 = i)$ represent the probability that you are in state j exactly n steps after reaching state i. The value of $r_{ij}(n)$ can be calculated recursively as

$$r_{ij}(n) = \sum_{k \in S} r_{ik}(n-1) p_{kj}$$

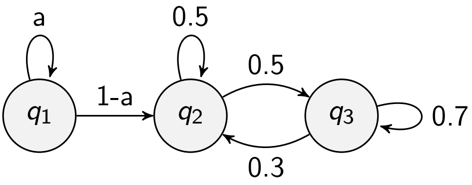
Observe that $r_{ij}(1) = p_{ij}$.

$$\Rightarrow r_{ij}(n) = \boldsymbol{P}^n_{i,j}$$

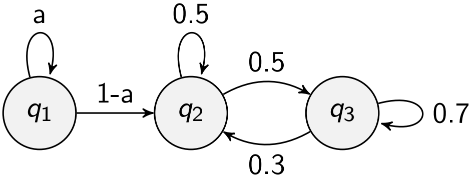
the i, j entry of $\boldsymbol{P}^n$.

Recurrent and Transient States

- **Accessible:** State j is accessible or reachable from state i if $\exists n \in \mathbb{N}$ such that $r_{ij}(n) > 0$.
- **Recurrence:** A state i is recurrent if $\forall j$ reachable from i, i is reachable from j. That is if $A(i)$ is the set of reachable states from i, then i is recurrent if $\forall j \in A(i) \Rightarrow i \in A(j)$.
- **Transient:** A state i is transient if it is not recurrent.
- Classify the states in the below Markov chain as recurrent or transient.



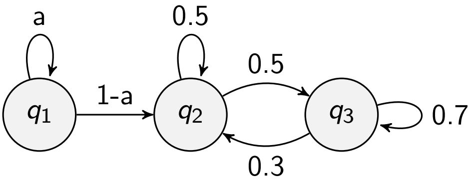
Answer



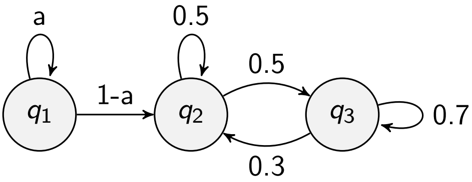
- If $a = 1$, q_1, q_2, q_3 are recurrent.
- If $a < 1$, q_1 is transient and q_2, q_3 are recurrent.

Decomposition of States

- **Recurrent Class:** For any recurrent state i, all states $A(i)$ (the set of states reachable from i) form a recurrent class. Any Markov chain can be decomposed into one ore more recurrent classes.
- A state in a recurrent class is not reachable from states in any other recurrent class (try to prove this).
- Transient states are not reachable from a recurrent state. Moreover, from every transient state atleast one recurrent state is reachable.
- Find the recurrent classes in the following MC:



Answers



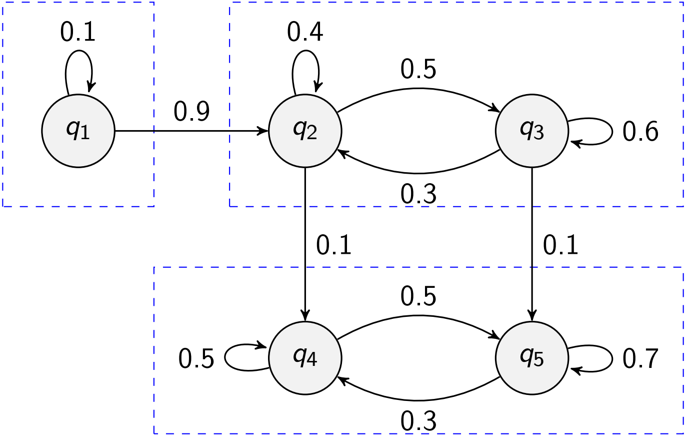
- If $a = 1$, $\{q_1\}, \{q_2, q_3\}$ form two recurrent classes.
- If $a < 1$, q_1 is transient, and $\{q_2, q_3\}$ form a recurrent class.

General Decomposition of States

A Markov chain is called **irreducible** if it only has one recurrent class. For any non-irreducible Markov chain, we can identify the recurrent classes using the following process

- Create directed edges between any two nodes that have a non-zero transition probability between them.
- Find the strongly connected components of the graph.
- Use transitions between different strongly connected components to further topologically sort the graph.
- Each strongly connected component at the bottom of the topologically sorted structure forms a recurrent class. All other nodes in this final structure are transient.

Example



Periodicity

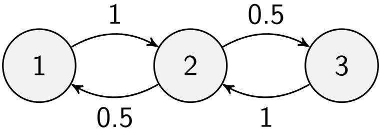
Consider an irreducible Markov chain. Define

$$d(i) := \text{g.c.d.}\{n \geq 1 | r_{ii}(n) > 0\}$$

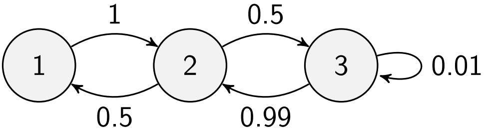
- Remember: $r_{ij}(n) = \mathbb{P}(X_{t+n} = j | X_t = i)$
- "All paths back to i take a multiple of $d(i)$ steps"
- Fact: $\forall i \; d(i)$ is the same.
- Fact: for Markov chains with more than one recurrent class, each class has a separate value for d
- We define a Markov chain as **aperiodic** if $d(i) = 1 \, \forall i$.
- Otherwise, we say it's **periodic** with period d

Periodicity Examples

Are the following Markov chains aperiodic?



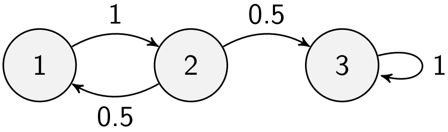
- $d = 2$, so this is **periodic**.



- $d = 1$, so this is **aperiodic**.
- Adding a self loop will make an irreducible Markov chain aperiodic!

Periodicity Examples (continued)

We won't particularly worry about periodicity/aperiodicity for Markov chains with more than 1 recurrent class.



Stationary Distribution

If we choose the initial state of the Markov chain according to the distribution

$$\mathbb{P}(X_0 = j) = \pi_0(j) \quad \forall j$$

and this implies

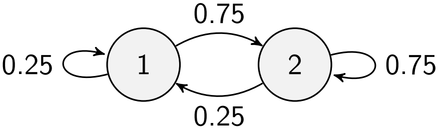
$$\mathbb{P}(X_n = j) = \pi_0(j) \quad \forall j, n$$

then we say that π_0 is **stationary**. The **balance equations** are sufficient for stationarity:

$$\pi_0(j) = \sum_{k=1}^{m} \pi_0(k) p_{kj} \quad \forall j$$

- The balance equations can be written as $\pi_0 = \pi_0 \boldsymbol{P}$. In linear algebra terms, π_0 is a left eigenvector of $\boldsymbol{P}$ that has corresponding eigenvalue $\lambda = 1$
- In general, there can be multiple unique stationary distributions.

Stationary Distribution Example



Let's try $\pi_0 = [1, 0]$.

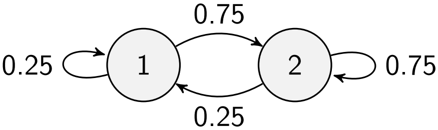
$$\pi_1(1) = \mathbb{P}(X_1 = 1 | X_0 = 1)\pi_0(1) + \mathbb{P}(X_1 = 1 | X_0 = 2)\pi_0(2) \quad (1)$$
$$= (0.25)(1) + (0.25)(0) \quad (2)$$
$$= 0.25 \quad (3)$$

Similarly,

$$\pi_1(2) = \mathbb{P}(X_1 = 2 | X_0 = 1)\pi_0(1) + \mathbb{P}(X_1 = 2 | X_0 = 2)\pi_0(2) \quad (4)$$
$$= (0.75)(1) + (0.75)(0) \quad (5)$$
$$= 0.75 \quad (6)$$

$\pi_1 = [0.25, 0.75] \neq \pi_0$, so $\pi_0 = [1, 0]$ is **not** stationary.

Stationary Distribution Example (continued)

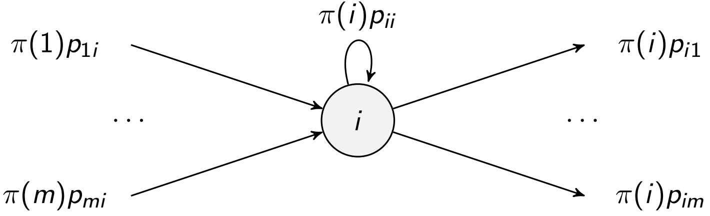


Let's solve for the stationary distribution. Let $\pi_0 = [x, 1 - x]$.

$$
\begin{aligned}
x &= \mathbb{P}(X_1 = 1 | X_0 = 1)x + \mathbb{P}(X_1 = 1 | X_0 = 2)(1 - x) && (7) \\
&= 0.25x + 0.25(1 - x) && (8) \\
1 - x &= \mathbb{P}(X_1 = 2 | X_0 = 1)x + \mathbb{P}(X_1 = 2 | X_0 = 2)(1 - x) && (9) \\
&= 0.75x + 0.75(1 - x) && (10)
\end{aligned}
$$

We see that $1 - x = 3x$, so $x = 0.25$. Our stationary distribution is $\pi_0 = [0.25, 0.75]$

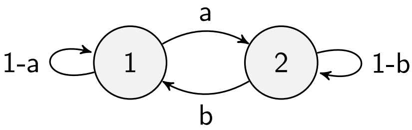
Probability Flow Interpretation



- For any distribution, probability mass flows in and out of every state at each step.
- By subtracting $\pi(i)p_{ii}$ from both sides of the balance equation, we have:

$$\underbrace{\sum_{j \neq i} \pi(j) p_{ji}}_{\text{flow in}} = \underbrace{\pi(i) \sum_{j \neq i} p_{ij}}_{\text{flow out}} \quad \forall i$$

Example Revisited



Let $\pi_0 = [x, 1 - x]$.

$$\sum_{j \neq i} \pi(j) p_{ji} = \pi(i) \sum_{j \neq i} p_{ij} \tag{11}$$

Using this at state 2,

$$xa = (1 - x)b \tag{12}$$

$$x = \frac{b}{a + b} \tag{13}$$

$$1 - x = \frac{a}{a + b} \tag{14}$$

The Big Theorem

- If a Markov chain is finite and irreducible, it has a unique invariant distribution π and $\pi(i)$ is the long term fraction of time that $X(n)$ is equal to i, almost surely.
- If the Markov chain is also aperiodic, then the distribution of $X(n)$ converges to π.

References

Introduction to probability. DP Bertsekas, JN Tsitsiklis - 2002