Exploring Markov Chains

Anirudh Vetukuri

12/10/2023

1 Fundamental Concepts

- **Event:** An event is a subset of the set of all possible outcomes in an experiment (a procedure with defined outcomes that is repeatable). For example, in a dice roll, an event could be rolling a number greater than 4, denoted as $\{5, 6\}$.
- **Probability:** Probability can be considered as a measure of the state of knowledge or certainty about an event. I.e. a measure quantifying the "chance" that an event will occur. This is defined as a number between 0 and 1.
- **Probability Function:** A probability function is a mathematical function that assigns a probability value to each outcome in the set of all possible outcomes. The sum of probabilities for all possible outcomes equals 1.
- **Probability Distribution:** A probability distribution describes the probabilities associated with all of the possible outcomes of a random variable.
- **Random Variable:** A random variable is a variable whose values are outcomes of a random process, assigning numerical values to each outcome. In a coin toss, a Random Variable X could be defined as 1 for heads and 0 for tails, mapping the outcomes to numbers.
- **System:** A system can be thought of as a collection of elements that interact with one another; or, in other words, a group of components that function together.
- **State:** A state is the status of a system at a specific time. It is a description of the relationships between the interacting things. For example, in a system of particles, a state can be the relative positions of those particles at a specific time.
- **Discrete:** Discrete refers to variables or systems that take distinct values. For example, the number of people in a classroom is discrete since it can only take distinct, integer values.

- **Process:** A process is simply a series of actions or changes within a system.
- **Stochastic Process:** A stochastic process is a collection of random variables representing the change of a system over time. It is defined as a sequence of random variables: $X_1, X_2, \ldots$, where each variable represents the state of the system at a different point in time.

2 Introduction to Markov Chains

2.1 Definition

A **Markov Chain** is a type of stochastic process that can model a sequence of events occurring in a system. Here, events refer to the transition from one state to another and model refers to the method of representing and analyzing a system's behavior through a set of states and transitions. Specifically, Markov Chains are used to model systems where the future state (the state the system will be in next) depends only on the current state (the state the system is in now) and not on previous states (the sequence of states that the system was in before the current state), this is known as the **Markov Property**.

2.1.1 Markov Property

The Markov Property is a property of some stochastic processes that states that the probability of transitioning to a future state depends only on the current state attained during the previous event, not on the sequence of states that preceded it. For a sequence of states represented by random variables $X_1, X_2, \ldots, X_n$, the Markov Property is satisfied if, for every n:

$$P(X_{n+1} = x | X_n = x_n) = P(X_{n+1} = x | X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n)$$

Here, X_n denotes the current state, and X_{n+1} represents the future/next state. The probability function P gives us the probability of moving from the current state X_n to the next state X_{n+1}.

2.2 Mathematical Formulation

2.2.1 State Space

The **state space** of a Markov Chain is the set of all distinct states that the system may be in. This can be represented as:

$$S = \{s_1, s_2, \ldots, s_n\}$$

where each s_i is a distinct state. The above is also an example of a finite Markov Chain, which means that the state space is finite (the total number of possible states are bounded by some real number).

For example, in a simple weather model, the state space might be $S = \{\text{Sunny}, \text{Cloudy}, \text{Rainy}\}$.

2.2.2 Transition Matrix

The **transition matrix** is a square matrix that describes the probabilities of moving from one state to another in one time step. It is often denoted as P and defined as:

$$P = [p_{ij}]$$

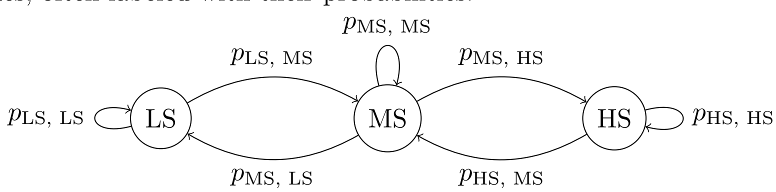
where p_{ij} represents the probability of transitioning from state i to state j in one step. Each row of the transition matrix represents a probability distribution over the states, and therefore, the sum of the probabilities in a row is 1.

For instance, in a retail business model with states Low Sales (LS), Moderate Sales (MS), and High Sales (HS), the transition matrix might look like:

$$P = \begin{bmatrix} p_{\text{LS, LS}} & p_{\text{LS, MS}} & p_{\text{LS, HS}} \\ p_{\text{MS, LS}} & p_{\text{MS, MS}} & p_{\text{MS, HS}} \\ p_{\text{HS, LS}} & p_{\text{HS, MS}} & p_{\text{HS, HS}} \end{bmatrix}$$

2.2.3 Graphical Representation

A Markov Chain can be represented graphically using a state diagram. Each state is a node in the graph, and transitions are directed edges between these nodes, often labeled with their probabilities.



Above is the state diagram visualizing the transitions between states.

2.3 Applications of Markov Chains

Markov Chains are extremely popular and can be found in many scientific areas. In Computer Science, A well-known application is Google's PageRank algorithm, which uses Markov Chains to rank web pages in search engine results. It models the internet as a massive network of web pages (states) and links between them (transitions), predicting the likelihood of a random surfer visiting a particular page (you might recognize this if you took MATH 1554), which ranks different web pages across the internet. In biology and social sciences, Markov Chains can accurately model population dynamics and disease transmission. Other areas include, physics, chemistry, finance, and sports.

3 Simple Example of a Markov Chain

This section presents a practical application of a Markov Chain. Consider a car rental service with a small fleet of cars. The number of cars available each day can be in one of three states: Low (L), Medium (M), or High (H). The state changes (transitions) based on customer rentals and returns.

3.0.1 Defining the Transition Matrix

The transition matrix for this scenario can be defined as follows, where again each element p_{ij} represents the probability of moving from state i to state j:

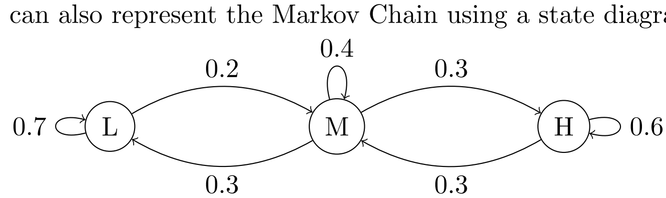
$$P = \begin{bmatrix} p_{\text{L, L}} & p_{\text{L, M}} & p_{\text{L, H}} \\ p_{\text{M, L}} & p_{\text{M, M}} & p_{\text{M, H}} \\ p_{\text{H, L}} & p_{\text{H, M}} & p_{\text{H, H}} \end{bmatrix}$$

Assume that through data analysis of customer behavior on a weekly basis, we are provided the matrix:

$$P = \begin{bmatrix} 0.7 & 0.2 & 0.1 \\ 0.3 & 0.4 & 0.3 \\ 0.1 & 0.3 & 0.6 \end{bmatrix}$$

3.0.2 Graphical Representation

We can also represent the Markov Chain using a state diagram as follows:



3.0.3 Problem Solving

A **steady-state distribution** is a probability vector π (vector with non-negative entries that sums to 1) that remains the same after applying the Markov Chain transition matrix P. It satisfies the equation $\pi P = \pi$, in this state, the probabilities of being in any particular state do not change from one step to the next.

The question is to determine the steady-state distribution for the availability of cars in the rental service (Low (L), Medium (M), and High (H)).

We can write out our system of equations using the matrix we defined above:

$$\begin{aligned} \pi_L \cdot 0.7 + \pi_M \cdot 0.3 + \pi_H \cdot 0.1 &= \pi_L \\ \pi_L \cdot 0.2 + \pi_M \cdot 0.4 + \pi_H \cdot 0.3 &= \pi_M \\ \pi_L \cdot 0.1 + \pi_M \cdot 0.3 + \pi_H \cdot 0.6 &= \pi_H \\ \pi_L + \pi_M + \pi_H &= 1 \end{aligned}$$

Upon solving these equations, we find the steady-state distribution:

$$\pi_L = 0.3947$$
$$\pi_M = 0.2895$$
$$\pi_H = 0.3158$$

The calculated values indicate that the car rental service has the following long-term distribution: a low number of cars available about 39.47% of the time, a medium number of cars available about 28.95% of the time, and a high car availability about 31.58% of the time.

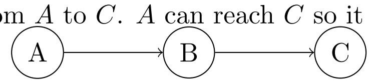
This steady-state distribution is valuable for the car rental service, providing insights into the typical availability of cars. This information can be useful for decisions about car management, marketing strategies, and supply and demand.

4 State Classification in Markov Chains

4.1 Accessibility

A state j is **accessible** from a state i if there is a non-zero probability of transition from i to j in some number of steps. This means that j is accessible from state i if there exists some $n \geq 0$ such that $p_{ij}^{(n)} > 0$, where $p_{ij}^{(n)}$ is the n-step transition probability from i to j. Visually in a state diagram, if the node of state j is reachable from state i then j is accessible from i.

Example: Below is a Markov Chain with states A, B, and C. There is a transition from A to B, and from B to C, although there is no direct transition from A to C. A can reach C so it is accesible.

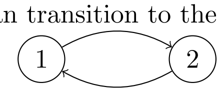


Here, state C is accessible from state A through B, but there is no direct path from A to C.

4.2 Communication

Two states i and j are said to **communicate** with each other if i is accessible from j and j is accessible from i.

Example: Consider a Markov Chain with states 1 and 2 where each state can transition to the other.



States 1 and 2 communicate with each other since there is mutual accessibility.

4.3 Recurrent and Transient States

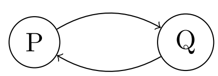
A state i is called **recurrent** if we begin at i, and the probability of going back to i is 1 over some number of steps. In other words, this means that once we reach recurrent state, it will continue to be returned to infinitely. The opposite type of state would be **transient**, which means there is a non-zero probability that the process will not return to this state after leaving it.

5 Types of Markov Chains

5.1 Regular Markov Chains

Regular Markov Chains are a class of Markov Chains where every state is accessible from every other state. This means after some number of steps, the probabilities of transitioning from one state to any other state are positive. Why is this important? This means we can reach a steady-state distribution.

Example: Consider a Markov Chain with states P and Q, where there are transitions between P and Q and vice versa.

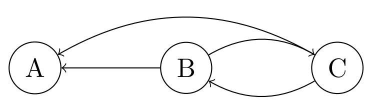


In this chain, it is possible to reach either state from the other, making it a regular Markov Chain.

5.2 Absorbing Markov Chains

Absorbing Markov Chains contain at least one absorbing state, which is a state that once entered cannot be left (think about how recurrent and absorbing states are related). These chains are useful in modeling processes where certain states can never be left (like winning/losing a game).

Example: Consider a Markov Chain with states A, B, and C, where A is an absorbing state, and there are transitions from B to C and C to B, but none out of A. Thus when A is reached, there are no transitions out of it.



Here, state A is an absorbing state. Once the process enters state A, it remains there.

6 The Gambler's Ruin Problem

Problem: Player A has \$1 and Player B has \$2. The winner of each game takes \$1 from the other. Player A is a better player than B, with a probability of winning 2/3 of the games. The game continues until one of the players is bankrupt. We are interested in finding the probability that Player A wins the entire game.

6.1 Absorbing Markov Chain Formulation

This problem can be modeled as an absorbing Markov Chain (due to the termination of the game). The states of the chain represent the amount of money Player A has, which can be 0, 1, 2, or 3 dollars. The states 0 and 3 are absorbing states, where the game ends either with Player A losing everything (state 0) or winning everything (state 3). States 1 and 2, are for Player A holding \$1 and \$2 respectively.

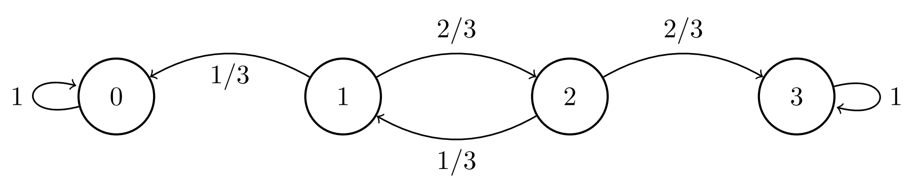
6.2 State Space and Transition Probabilities

The state space is defined as $\{0, 1, 2, 3\}$. The transition probabilities are as follows:

- From state 1 (when Player A has \$1), the probability of going to state 0 (Player A loses a game) is 1/3, and to state 2 (Player A wins a game) is 2/3.
- From state 2 (when Player A has \$2), the probability of going to state 1 (Player A loses a game) is 1/3, and to state 3 (Player A wins a game) is 2/3.

The transition matrix P and diagram are given by:

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



6.3 Solution

To find the probability that Player A wins (reaches state 3 starting from state 1) let a_i be the probability of reaching state 3 starting from a state i. We have:

$$a_0 = 0 \quad \text{(Player A cannot win from state 0)}$$
$$a_3 = 1 \quad \text{(Player A has already won in state 3)}$$
$$a_1 = \frac{1}{3}a_0 + \frac{2}{3}a_2$$
$$a_2 = \frac{1}{3}a_1 + \frac{2}{3}a_3$$

Solving these equations, we find that:

$$a_1 = 0.5714 \quad \text{(Approximately 57.14\% chance of Player A winning from state 1)}$$
$$a_2 = 0.8571 \quad \text{(Approximately 85.71\% chance of Player A winning from state 2)}$$

This shows us player A has a 57.14% chance of winning the game.

7 Further Exercises

Exercise 1: Two-State Markov Chain

Consider we have a Markov Chain with two states, X and Y. The transition probabilities are as follows: A to B is 0.4, and from B to A is 0.6. Create the transition matrix and find the steady-state.

Exercise 2: Absorbing States in a Board Game

Let's play a game! There are 3 stairs. 2 players start at the first stair and either take a step up to the second stair or go back to the floor (stair 0). From stair 2, they can move to 3 or back to 1. This is a short staircase, where reaching stair 3 means you made it to the top and win. If the probability of moving up is twice the probability of moving back, again create the transition matrix and identify the absorbing states.