Markov Chains

A **Markov chain** is a mathematical system that experiences transitions from one state to another according to certain probabilistic rules. The defining characteristic of a Markov chain is that no matter *how* the process arrived at its present state, the possible future states are fixed. In other words, the probability of transitioning to any particular state is dependent solely on the current state and time elapsed. The **state space**, or set of all possible states, can be anything: letters, numbers, weather conditions, baseball scores, or stock performances.
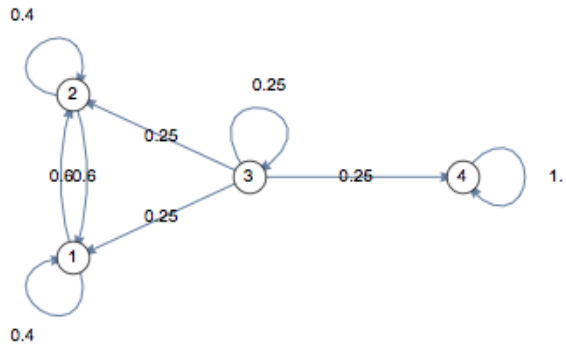
Markov chains may be modeled by finite state machines, and random walks provide a prolific example of their usefulness in mathematics. They arise broadly in statistical and information-theoretical contexts and are widely employed in economics, game theory, queueing (communication) theory, genetics, and finance. While it is possible to discuss Markov chains with any size of state space, the initial theory and most applications are focused on cases with a finite (or countably infinite) number of states.

Many uses of Markov chains require proficiency with common matrix methods.

Contents

Basic Concept

A Markov chain is a stochastic process, but it differs from a general stochastic process in that a Markov chain must be "memory-less". That is, (the probability of) future actions are not dependent upon the steps that led up to the present state. This is called the **Markov property**. While the theory of Markov chains is important precisely because so many "everyday" processes satisfy the Markov property, there are many common examples of stochastic properties that do not satisfy the Markov property.

> **EXAMPLE**
>
> A common probability question asks what is the probability of getting a certain color ball, when selecting uniformly and at random from a bag of multicolored balls. It could also ask what the probability of the next ball is, and so on. In such a way, a stochastic process begins to exist with color for the random variable, and it does not satisfy the Markov property. Depending upon which balls are removed, the probability of getting a certain color ball later may be drastically different.
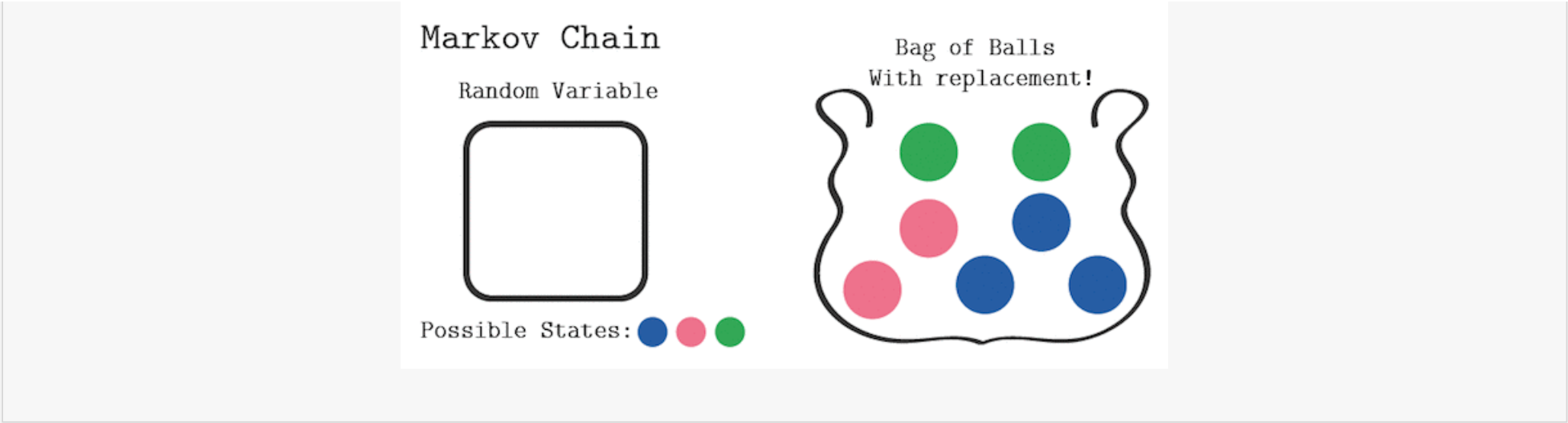>
>
>
> A variant of the same question asks once again for ball color, but it allows replacement each time a ball is drawn. Once again, this creates a stochastic process with color for the random variable. This process, however, **does** satisfy the Markov property. Can you figure out why?

In probability theory, the most immediate example is that of a *time-homogeneous Markov chain*, in which the probability of any state transition is independent of time. Such a process may be visualized with a labeled directed graph, for which the sum of the labels of any vertex's outgoing edges is 1.

EXAMPLE

A (time-homogeneous) Markov chain built on states A and B is depicted in the diagram below. What is the probability that a process beginning on A will be on B after 2 moves?

In order to move from A to B, the process must either stay on A the first move, then move to B the second move; or move to B the first move, then stay on B the second move. According to the diagram, the probability of that is
$$0.3 \cdot 0.7 + 0.7 \cdot 0.2 = \boxed{0.35.}$$

Alternatively, the probability that the process will be on A after 2 moves is $0.3 \cdot 0.3 + 0.7 \cdot 0.8 = 0.65$. Since there are only two states in the chain, the process must be on B if it is not on A, and therefore, the probability that the process will be on B after 2 moves is $1 - 0.65 = \boxed{0.35.}$

In the language of conditional probability and random variables, a Markov chain is a sequence X_0, X_1, X_2, ... of random variables satisfying the rule of conditional independence:

DEFINITION

The Markov Property.

For any positive integer n and possible states i_0, i_1, ..., i_n of the random variables,

$$P(X_n = i_n \mid X_{n-1} = i_{n-1}) = P(X_n = i_n \mid X_0 = i_0, X_1 = i_1, \ldots, X_{n-1} = i_{n-1}).$$

In other words, knowledge of the previous state is all that is necessary to determine the probability distribution of the current state. This definition is broader than the one explored above, as it allows for *non-stationary transition probabilities* and therefore *time-inhomogeneous Markov chains*; that is, as time goes on (steps increase), the probability of moving from one state to another may change.

Transition Matrices

A **transition matrix** P_t for Markov chain $\{X\}$ at time t is a matrix containing information on the probability of transitioning between states. In particular, given an ordering of a matrix's rows and columns by the state space S, the $(i, j)^{\text{th}}$ element of the matrix P_t is given by

$$(P_t)_{i,j} = \mathbb{P}(X_{t+1} = j \mid X_t = i).$$

This means each row of the matrix is a *probability vector*, and the sum of its entries is 1.

Transition matrices have the property that the product of subsequent ones describes a transition along the time interval spanned by the transition matrices. That is to say, $P_0 \cdot P_1$ has in its $(i, j)^{\text{th}}$ position the probability that $X_2 = j$ given that $X_0 = i$. And, in general, the $(i, j)^{\text{th}}$ position of $P_t \cdot P_{t+1} \cdots P_{t+k}$ is the probability $\mathbb{P}(X_{t+k+1} = j \mid X_t = i)$.

EXAMPLE

Prove that, for any natural number t and states $i, j \in S$, the matrix entry $(P_t \cdot P_{t+1})_{i,j} = \mathbb{P}(X_{t+2} = j \mid X_t = i)$.

Denote $M = P_t \cdot P_{t+1}$. By matrix multiplication,

$$M_{i,j} = \sum_{k=1}^{n} (P_t)_{i,k} (P_{t+1})_{k,j}$$

$$= \sum_{k=1}^{n} \mathbb{P}(X_{t+1} = k \mid X_t = i)\mathbb{P}(X_{t+2} = j \mid X_{t+1} = k)$$

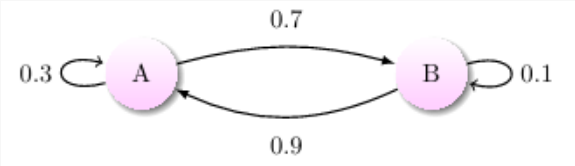$$= \boxed{\mathbb{P}(X_{t+2} = j \mid X_t = i).}$$

The final equality follows from conditional probability.

The k-**step** transition matrix is $P_t^{(k)} = P_t \cdot P_{t+1} \ldots P_{t+k-1}$ and, by the above, satisfies

$$P_t^{(k)} = \begin{pmatrix} \mathbb{P}(X_{t+k} = 1 \mid X_t = 1) & \mathbb{P}(X_{t+k} = 2 \mid X_t = 1) & \ldots & \mathbb{P}(X_{t+k} = n \mid X_t = 1) \\ \mathbb{P}(X_{t+k} = 1 \mid X_t = 2) & \mathbb{P}(X_{t+k} = 2 \mid X_t = 2) & \ldots & \mathbb{P}(X_{t+k} = n \mid X_t = 2) \\ \vdots & \vdots & \ddots & \vdots \\ \mathbb{P}(X_{t+k} = 1 \mid X_t = n) & \mathbb{P}(x_{t+k} = 2 \mid X_t = n) & \ldots & \mathbb{P}(X_{t+k} = n \mid X_t = n) \end{pmatrix}.$$

EXAMPLE

For the time-independent Markov chain described by the picture below, what is its 2-step transition matrix?

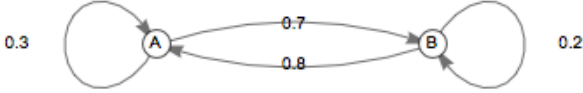

Note the transition matrix is

$$P = \begin{pmatrix} 0.3 & 0.7 \\ 0.9 & 0.1 \end{pmatrix}.$$

It follows that the 2-step transition matrix is

$$P^2 = \begin{pmatrix} 0.3 & 0.7 \\ 0.9 & 0.1 \end{pmatrix} * \begin{pmatrix} 0.3 & 0.7 \\ 0.9 & 0.1 \end{pmatrix} = \begin{pmatrix} 0.3 \cdot 0.3 + 0.7 \cdot 0.9 & 0.3 \cdot 0.7 + 0.7 \cdot 0.1 \\ 0.9 \cdot 0.3 + 0.1 \cdot 0.9 & 0.9 \cdot 0.7 + 0.1 \cdot 0.1 \end{pmatrix} = \begin{pmatrix} 0.72 & 0.28 \\ 0.36 & 0.64 \end{pmatrix}.$$

TRY IT YOURSELF

A Markov chain has first state A and second state B, and its transition probabilities for all time are given by the following graph.

What is its transition matrix?

Note: Note: The transition matrix is oriented such that the kth **row** represents the set of probabilities of transitioning from state k to another state.

$\bigcirc \quad \begin{pmatrix} 1 & 0.8 \\ 0.7 & 1 \end{pmatrix}$

$\bigcirc \quad \begin{pmatrix} 0.3 & 0.8 \\ 0.7 & 0.2 \end{pmatrix}$

$\bigcirc \quad \begin{pmatrix} 0.5 & 1.5 \\ 1.5 & 0.5 \end{pmatrix}$

$\bigcirc \quad \begin{pmatrix} 0.3 & 0.7 \\ 0.8 & 0.2 \end{pmatrix}$

TRY IT YOURSELF

Submit your answer
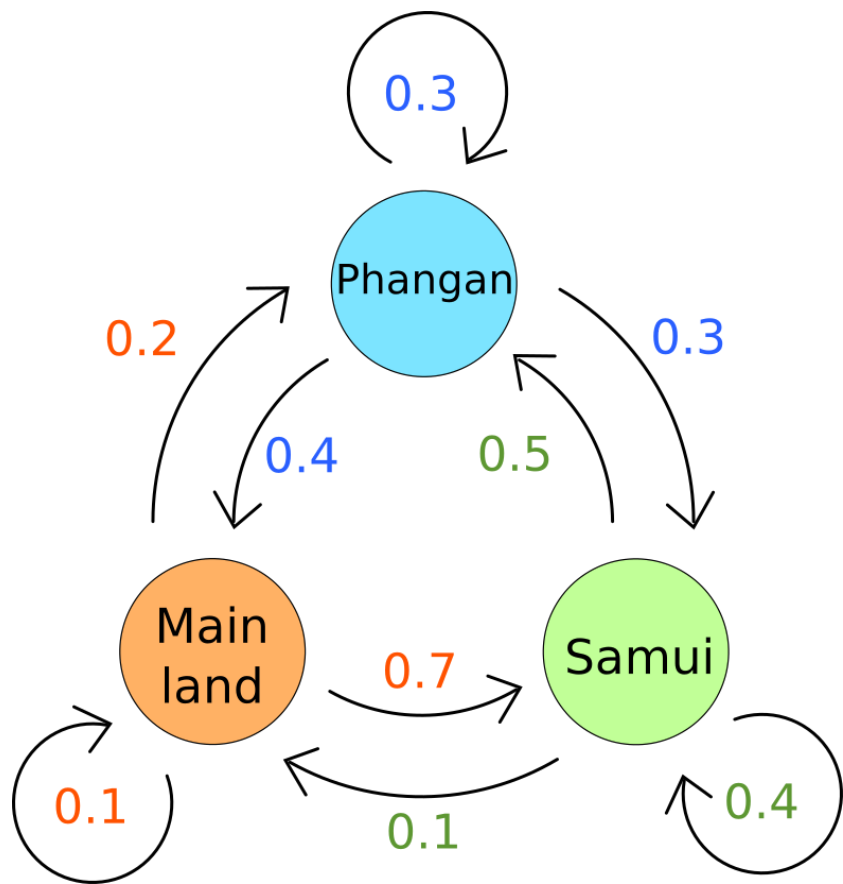
Once people arrive in Thailand, they want to enjoy the sun and beaches on 2 popular islands in the south: Samui Island & Phangan Island.

From survey data, when on the mainland, 70% of tourists plan to go to Samui Island, 20% to Phangan Island, and only 10% remain on shore the next day.

When on Samui Island, 40% continue to stay on Samui, 50% plan to go to Phangan Island, and only 10% return to mainland the next day.

Finally, when on Phangan Island, 30% prolong their stay here, 30% divert to Samui Island, and 40% go back to mainland the next day.

Starting from the mainland, what is the probability (in percentage) that the travelers will be on the mainland at the end of a 3-day trip?

Properties

A variety of descriptions of either a specific state in a Markov chain or the entire Markov chain allow for better understanding of the Markov's chain behavior. Let P be the transition matrix of Markov chain $\{X_0, X_1, \ldots\}$.

- A state i has **period** $k \geq 1$ if any chain starting at and returning to state i with positive probability must take a number of steps divisible by k. If $k = 1$, then the state is known as **aperiodic**, and if $k > 1$, the state is known as **periodic**. If all states are aperiodic, then the Markov chain is known as aperiodic.
- A Markov chain is known as **irreducible** if there exists a chain of steps between any two states that has positive probability.
- An **absorbing** state i is a state for which $P_{i,i} = 1$. Absorbing states are crucial for the discussion of absorbing Markov chains.
- A state is known as recurrent or transient depending upon whether or not the Markov chain will eventually return to it. A recurrent state is known as positive recurrent if it is expected to return within a finite number of steps and null recurrent otherwise.
- A state is known as ergodic if it is positive recurrent and aperiodic. A Markov chain is ergodic if all its states are.

Irreducibility and periodicity both concern the locations a Markov chain could be at some later point in time, given where it started. Stationary distributions deal with the likelihood of a process being in a certain state at an unknown point of time. For Markov chains with a finite number of states, each of which is positive recurrent, an aperiodic Markov chain is the same as an irreducible Markov chain.
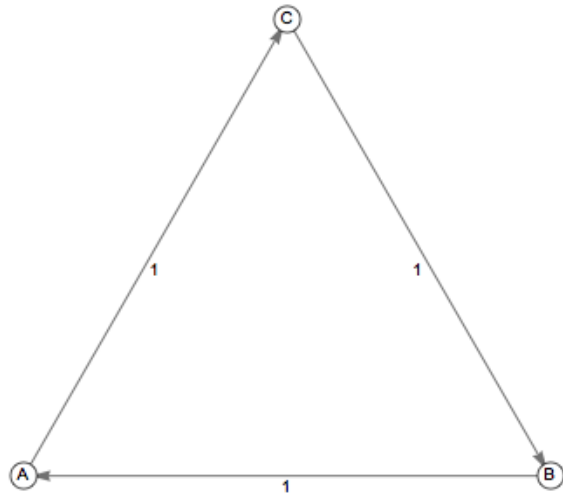

TRY IT YOURSELF

The graphs of two time-homogeneous Markov chains are shown below.

○ Both Markov chains are periodic and irreducible
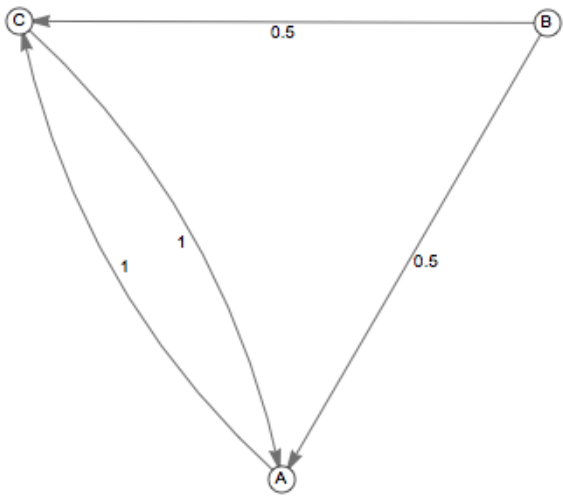
Chain 1

Chain 2

○ Neither Markov chain is irreducible

○ Chain 1 is aperiodic and irreducible, and chain 2 is aperiodic

○ Chain 1 is aperiodic, and chain 2 is irreducible

○ Chain 1 is not irreducible, and chain 2 is not aperiodic

Determine facts about their periodicity and reducibility.

See Also

Markov Chain subpages:

- Stationary Distributions
- Ergodic Markov Chains
- Absorbing Markov Chains
- Transience and Recurrence

Miscellaneous:

- Eigenvectors
- Finite State Machines
- Matrices
- Random Walks

Cite as: Markov Chains. *Brilliant.org*. Retrieved 12:17, May 3, 2018, from https://brilliant.org/wiki/markov-chains/