INTRODUCTION TO BASIC PROPERTIES OF MARKOV CHAIN AND ITS APPLICATION IN SIMPLE RANDOM WALK

GUANYU CHEN

ABSTRACT. A Markov model is a stochastic model which assumes that the transition probabilities between states depend on only the current state rather than the past states (Markov property). Each transition is assigned a probability that defines the chance of the system changing from one state to another. This paper intends to introduce some basic definitions and concepts of Markov chains, such as reducibility, and invariant distribution. After going through the fundamental concepts, we will recurrence and transience properties in Simple Random Walk.

CONTENTS

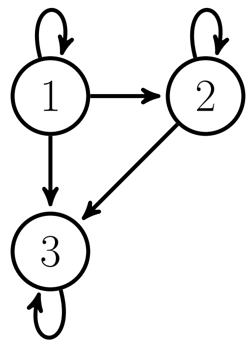
1. TIME HOMOGENEOUS FINITE MARKOV CHAIN

Markov Chains are based on the stochastic model which assumes that the transition probabilities between states depend on only the current state rather than the past state (Markov property). Each transition is assigned a probability that defines the chance of the system changing from one state to another. A stochastic process is any process describing the evolution in time of a random phenomenon, a collection or ensemble of random variables indexed by a variable t.

Definition 1.1. A stochastic process is defined as the Markov process if $P(X_k = x_k|X_0 = x_0, X_1 = x_1, ..., X_{k-1} = x_{k-1})$=$P(X_k = x_k|X_{k-1} = x_{k-1})$, for k greater than 1, where $P(X_k = x_k|X_{k-1} = x_{k-1})$ refers to the conditional probability of being in state x_k at the kth step given that, at the ith step, the process is in state x_i. To make the past conditions independent of the future given the present condition, the Markov property simplifies each of the past conditional transitions.

Definition 1.2. We denote the initial distribution of X_0 by $\mu = (\mu_i)$, for i belongs to S, S refers to the set of possible states, $P(X_0 = i) = \mu_i$

Date: February 2, 2023.



Definition 1.3. The transtition matrix P for the Markov chain is the N*N matrix, N referring the number of states, and we denote the entry as $\boldsymbol{P_{ij}}$, where$\boldsymbol{P_{ij} = P(X_k = j|x_{k-1} = i)}$. Therefore, the matrix P is a stochastic matrix, that is, each entry is between 0 and 1, and $\sum \boldsymbol{P_{ij} = 1}$. The N*N transitional matrix is represented as

$$\begin{bmatrix} \boldsymbol{p_{11}} & \cdots & \boldsymbol{p_{1n}} \\ \vdots & \ddots & \vdots \\ \boldsymbol{p_{n1}} & \cdots & \boldsymbol{p_{nn}} \end{bmatrix}$$

A time-homogeneous finite Markov chain (X_k), for k greater than 0, is entirely determined by: 1. an initial distribution $\mu = (\mu_i)_i \in S$, with the interpretation $\mu_i = P(X_0 = i)$

2. a Stochastic matrix in N*N size, $P = (p_{ij})_{i,j} \in S$, i.e., the entries of P satisfy $p_{ij} \geq 0$, for all $i, j \in S$, and $\Sigma p_{ij} = 1$ for each $i \in S$, with the interpretation $p_{ij} = P(X_k = j|X_{k-1} = i)$ for all $k \geq 1$.

2. Large-Time Behavior and invariant Probability

Theorem 2.1. *The 1*n row vector μP^n is the distribution of X_n given X_0 has distribution μ in the sense that $P_\mu(X_n = i)$ equals the ith entry of (μP^n)*

We can understand the large-time behavior of a Markov chain denote $\lim_{n\to\infty} P^n$ by Π as the limit matrix, that is $\Pi = \lim_{n\to\infty}$ we denote the row vector of the limit matrix π. If $\vec{v}$ is any probability vector, then

$$\lim_{n\to\infty} \vec{v}P^n = \pi$$

Example 2.2. for the matrix P

$$\begin{bmatrix} \mathbf{0.75} & \mathbf{0.25} \\ \mathbf{1/6} & \mathbf{5/6} \end{bmatrix}$$

We calculate the limit matrix of P with a computer algorithm, and we get $\boldsymbol{\lim_{n\to\infty} P^n}$

$$\begin{bmatrix} \mathbf{0.4} & \mathbf{0.6} \\ \mathbf{0.4} & \mathbf{0.6} \end{bmatrix}$$

As $\boldsymbol{\Pi = \lim_{n\to\infty} P^n}$, the row vectors of $\boldsymbol{Pi}$ are identical, which means that the probability of being in a given state after a long time if independent of the initial state

Suppose π is a limiting probability vector,

$$\pi = \lim_{n\to\infty} \vec{v}P^n \tag{2.3}$$

Then

$$\lim_{n\to\infty} P^{n+1} = (\lim_{n\to\infty} P^n)P \tag{2.4}$$

since when n approaches infinity, n+1 also approaches infinity

$$\pi = \lim_{n\to\infty} \bar{v}P^n = \lim_{n\to\infty} P^{n+1} \tag{2.5}$$

then

$$\lim_{n\to\infty} P^{n+1} = \pi = (\lim_{n\to\infty} P^n)P = \pi * P \tag{2.6}$$

Thus, we have

$$\pi = \pi * P \tag{2.7}$$

Definition 2.8. if the equation $\pi = \pi P$ holds, the probability vector π is known as an **invariant distribution** for P Once a Markov process reaches the stationary distribution, the row vector will always stay invariant regardless of the future itinerary.

Example 2.9. Consider an example of a two-state Markov chain, with the matrix P

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

when $p > 0$ and $q < 1$, the matrix has the eigenvalue of 1-p-q and 1. After diagonalization, we have $A = B^{-1}PB$, where

A=$\begin{bmatrix} 1 & 0 \\ 0 & 1-p-q \end{bmatrix}$ B=$\begin{bmatrix} 1 & -p \\ 1 & q \end{bmatrix}$ B^{-1}=$\begin{bmatrix} q/(p+q) & p/p+q \\ -1/(p+q) & 1/(p+q) \end{bmatrix}$

After diagonalization, we can easily calculate the power of P

$$\begin{aligned} P^n &= (BAB^{-1})^n \\ &= BA(BB^{-1})\cdots(BB^{-1})AB^{-1} \\ &= BA^nB^{-1} \\ &= B\begin{bmatrix} 1 & 0 \\ 0 & (1-p-q)^n \end{bmatrix}B^{-1} \\ &= \begin{bmatrix} [q+p(1-p-q)^n]/(p+q)[p-p(1-p-q)^n]/(p+q) \\ [q-q(1-p-q]^n]/(p+q)[p+q(1-p-q)^n]/(p+q) \end{bmatrix} \end{aligned}$$

since $0 < p,q < 1$, we have $|1-p-q| < 1$, thus, as n approaches infinity,

$$\lim_{n\to\infty} P^n = \begin{bmatrix} q/(p+q)p/(p+q) \\ q/(p+q)p/(p+q) \end{bmatrix}$$

Denote the first column vector of π as π_1, and the second column vector of π as π_2. According to the equation

$$\pi = \pi P$$

we have

$$\begin{cases} \pi_1 = (1-p)\pi_1 + q\pi_2 \\ \pi_2 = p\pi_1 + (1-q)\pi_2 \end{cases}$$

Solving the equation, we have $p\pi_1$= $q\pi_2$. Also, because π is a probability distribution, we have that $\pi_1 + \pi_2 = 1$
By solving this system of equations, we have the equation

$$\pi = (\pi_1, \pi_2) = \begin{bmatrix} \frac{q}{p+q} & \frac{p}{p+q} \end{bmatrix}$$

Now we check that π is indeed a stationary distribution by multiplying with P

$$\pi P = \begin{bmatrix} \frac{q}{p+q} & \frac{p}{p+q} \end{bmatrix} \begin{bmatrix} 1-p & p \\ q & 1-q \end{bmatrix}$$
$$= \begin{bmatrix} \frac{q(1-p)}{p+q} + \frac{pq}{p+q} & \frac{qp}{p+q} + \frac{p(1-q)}{p+q} \end{bmatrix}$$
$$= \begin{bmatrix} \frac{q}{p+q} & \frac{p}{p+q} \end{bmatrix} = \pi$$

We can observe that the row vector is the same, which is known as the unique invariant probability distribution. Thus, π is the unique invariant probability distribution for P.

3. Reducibility and Periodicity

Definition 3.1. We use $p_m(i, j)$ to denote the probability of being in state j after m steps if starting in state i. if there exists $m, n \geq 0$ such that $p_m(i, j) > 0$ and $p_n(j, i) > 0$, then we say the two states i and j of the Markov chain **communicate** with each other, which means that it is possible to move back and forth between these two states.

Theorem 3.2. *The relation $\leftrightarrow$ is an equivalence relation on the state space.*

Proof. a.reflexive:$i \leftrightarrow i$, as $p_0(i, i) = 1 > 0$
b.symmetric $i \leftrightarrow j$ implies that $j \leftrightarrow i$, as the definition illustrates the mutual relation
c.transitive: $i \leftrightarrow j$ and $j \leftrightarrow k$ imply $i \leftrightarrow k$. By the definition, there exists m,n such that $p_m(i, j) > 0$ and $p_n(j, k) > 0$, thus:

$$\begin{aligned} P_{m+n}(i, k) &= P(X_{m+n} = k | X_0 = i) \geq P(X_{m+n} = k, X_m = j | X_0 = i) \\ &= P(X_{m+n} = k | X_m = j) P(X_m = j | X_0 = i) \\ &= p_m(i, j) p_n(j, k) \text{ which is greater than } 0 \end{aligned}$$

□

Definition 3.3. If there is only one communication class in the Markov Chain, that is, if for all i,j there exists an n with $p_n(i, j) > 0$, then the Markov Chain is called **irreducible**

Definition 3.4. For an irreducible Markov chain with stochastic matrix denoted as **P**, we define the **period** d_k of a state k, to be the greatest common divisor of $J_k := \{n \geq 0 : p_n(i, i)\} > 0$.
If d=1, then we call the transition matrix **P aperiodic**

Theorem 3.5. *If the transition matrix $\boldsymbol{P}$ is irreducible and aperiodic, then there exists an $N > 0$ such that for all $n \geq N$, all the entries of $\boldsymbol{P}^n$ are strictly larger than 0.*

Proof. Firstly, since **P** is irreducible there exits some k(i,j) such that $p_{k(i,j)}(i,j) > 0$. Secondly, as **P** is aperiodic, there must exist some N(i) such that for any $n \geq N(i)$, $p_n(i,i) > 0$.
Therefore, for all $n \geq N(i)$,

$$p_{k+n(i,j)}(i,j) \geq p_k(i,i)p_{n(i,j)}(i,j) > 0 \tag{3.6}$$

As long as the state matrix is finite, there exists a maximum value of N(i)+k(i,j) for any entries in the transition matrix, and let denote this maximum value as M. If $n \geq M$, then $p_n(i,j) > 0$ for all i, j.

□

4. Recurrence and Transience

Definition 4.1. If a Markov chain has more than one communication class, and for some classes, the probability that the Markov chain will eventually leave the class is 1, we named classes with this property as **transient** classes.

Definition 4.2. a Markov chain will never leave or eventually end with each state is called a **recurrent** class. This is, for state x, $P(X_n = x) = 1$ for infinitely many n
If a certain event has a positive probability of occurring, and we have an infinite number of trials, then the event will certainly happen an infinite number of times.

Theorem 4.3. *The expected number of times an irreducible Markov chain will return to a state is finite, $\sum_{k=0}^{\infty} p_k(x,x) < \infty$, if and only if the state of an irreducible Markov chain is transient.*

Proof. Let T be the time of first return to state x,

$$T = min(n > 0 : X_n = x)$$

If the chain will never return to x, then $T = \infty$
Suppose that $P(T = \infty) = 0$, that is, $P(T < \infty) = 1$, then with probability one, the chain always returns. Furthermore, we deduce that with probability 1, the chain returns infinitely often.
Suppose that $P(T < \infty) = q < 1$. First, the chain never returns if and only if $R = 1$; $P(R = 1) = 1 - q$. if $k > 1$, then $R = k$ if and only if the chain returns k-1 times and then does not return for the kth time.
Thus, $P(R = k) = (1 - q)q^{k-1}$.
Therefore, when $q < 1$, which is in the transient state

$$E(R) = \sum_{k=1}^{\infty} kP(R = k) = \sum_{k=1}^{\infty} kq^{k-1}(1 - q) = \frac{1}{1 - q} < \infty.$$

□

5. Recurrence and Transience in Simple Random Walk

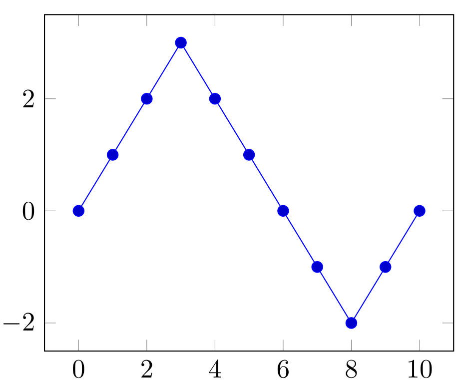
Definition 5.1. A random walk in mathematics is a path comprised of sequential random steps within a mathematical space, described as a random process. A common example would be a random walk along the integer number line starting at 0 and moving +1 or -1 at each step with equal possibility.

Theorem 5.2. *Simple random walks are recurrent in lattices of dimensions 1 and 2 and are transient when the dimension is larger than or equal to 3.*

Proof. First, consider the simple random walk in dimension 1, where the Markov chains have the transition probabilities

$$p(x, x+1) = p(x, x-1) = \frac{1}{2}$$

We use $X_n = k$ to note the location of random walk in state n, and we assume that $X_0 = 0$. Suppose the walker ended up at the origin after certain steps, then the steps the walker takes towards the left must equal the steps taken towards the right. Thus, the total steps must be 2n, n is any integer larger than 0, and the walker takes exactly n steps towards left and n steps towards right
pgfplots



The probability of returning to the origin after 2n steps are $(\frac{1}{2})^{2n}$, as at each step, we have two directions to choose from. Also, there are $\binom{2n}{n}$ ways to choose which of the n steps among 2n are taken towards left, and then the remaining n steps are towards the right:

$$p_{2n}(0,0) = \binom{2n}{n}(\frac{1}{2})^{2n} = \frac{(2n)!}{n!n!}(\frac{1}{2})^{2n}$$

Applying Stirling's formula to estimate the factorials (2n)! and n!:

$$n! \sim \sqrt{2\pi n}n^n e^{-n}$$

After plugging the estimation into our equation above, we have:

$$p_{2n}(0,0) \sim \frac{1}{\sqrt{\pi n}}$$

Therefore, the series of $\frac{1}{\sqrt{\pi n}}$is not convergent, $\sum_{n=0}^{\infty} p_{2n}^{-1/2} = \infty$. According to the theorem, we have proved that a simple random walk in one dimension is recurrent.

Next, consider the situation when the dimension d is greater than 1. The Markov chain has the transition probabilities of

$$p(x,y) = 1/2d, \text{for any two adjacent states x, y}$$

That is, in dimension 2, $p(x,y) = \frac{1}{4}$, while in dimension 3, $p(x,y) = \frac{1}{6}$ Similarly, $p_n(0,0) = 0$, as the walker still needs to take an even number of steps to return.'Now consider each dimension as a component, then we need the walker to take even steps in each component in order to return to the origin. For a sufficiently large n, the probability that we will have even steps in each component is $(\frac{1}{2})^{d-1}$. (If d-1 components have even steps since we have even steps in total, the remaining component must have even steps.)
When n is large enough, by the law of large numbers, we estimate that $\frac{2n}{d}$ steps will be taken in every component. Therefore, for each component, there are $\binom{2n/d}{n/d}$ ways to return to the origin and each has a probability of $(\frac{1}{2})^{2n/d}$. Applying the sterling formula, we have the result:

$$p_{2n}(0,0) \sim (\frac{1}{2})^{d-1}(\frac{d}{n\pi})^{d/2}$$

For n from 0 to ∞, the sum of the series n^{-a} will converge if and only if a is greater than 1, that is $\sum_0^\infty n^{-a} < \infty$ if and only if $a > 1$
Therefore, according to equation (5.7), for d=1 or 2, as 1/2 is less than 1 and 2/2 is equal to 1, we apply Theorem 4.3:

$$p_{2n}(0,0) = \infty$$

while for $d > 2$, $d/2 > 1$

$$p_{2n}(0,0) < \infty$$

Therefore, we have finished the proof

Acknowledgments

It is my pleasure to thank my mentor, Callum Sutton, who guided me through determining my research topic, finding resources, and editing this paper. I would also like to thank Professor Daniil Rudenko and Professor Peter May for organizing this brilliant program. This paper would not have been possible without this guidance and support.

References

[1] Gregory F. Lawler. Introduction to Stochastic Processes. Taylor & Francis Group. 2006.