Note on Discrete Markov Chain

Zepeng CHEN
The HK PolyU

Date: January 12, 2023

1 Markov Chain and Transition probability

Definition 1.1 (Markovian property)

For a Markov chain, the conditional distribution of any future state X_{n+1}, is independent of the past states $X_0, ..., X_{n-1}$ and depends only on the present state X_n. This is called the Markovian property.

Definition 1.2 (Markov Chain)

Consider a stochastic process $\{X_n, n = 0, 1, 2...\}$ that takes on a finite or countable number of possible values, which is denoted by the set of nonnegative integers $\{0, 1, 2...\}$. If $X_n = i$, then the process is said to be in state i at time n. And whenever the process is in state i, there is a fixed probability P_{ij} that it will next be in state j. That is, the follow equation holds for all states and all $n \geq 0$. Such a stochastic process is called a Markov chain.

$$\begin{aligned} & P\{X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, \ldots, X_1 = i_1, X_0 = i_0\} \\ = & P\{X_{n+1} = j \mid X_n = i\} = P_{ij} \end{aligned}$$

Markov chain is a discrete time, discrete state stochastic process with Markovian property.

Definition 1.3 (One-step transition probabilities)

The value P_{ij} represents the probability that the process will, when in state i, next make a transition into state j. And it possesses following properties. There is a matrix form P to present these transition probabilities.

$$P_{ij} \geq 0, \quad i, j \geq 0; \quad \sum_{j=0}^{\infty} P_{ij} = 1, \quad i = 0, 1, \ldots$$

$$P = \begin{bmatrix} P_{00} & P_{01} & \cdots & P_{0j} & \cdots \\ P_{10} & P_{11} & \cdots & P_{1j} & \cdots \\ \vdots & \vdots & \ddots & \vdots & \\ P_{i0} & P_{i1} & \cdots & P_{ij} & \cdots \\ \vdots & \vdots & & \vdots & \ddots \end{bmatrix}$$

Definition 1.4 (n-Step transition probabilities)

The n-step transition probabilities P_{ij}^n is the probability that a process is in state i will be in state j after n additional transitions. That is, as follows. Note that $P_{ij}^1 = P_{ij}$.

$$P_{ij}^n = P\left\{X_{n+m} = j \mid X_m = i\right\}, \quad n \geq 0 \text{ and } i, j \geq 0$$

Theorem 1.1 (Chapman-Kolmogorov equations)

$$P_{ij}^{n+m} = \sum_{k=0}^{\infty} P_{ik}^n P_{kj}^m \quad \text{for all } n, m \geq 0 \text{ and } i, j \geq 0$$

Let $P^{(n)}$ denote the matrix of n-step transition probabilities P_{ij}^n, then we have

$$P^{(n+m)} = P^{(n)} \cdot P^{(m)}$$

$$= \begin{bmatrix} P_{00}^n & P_{01}^n & \cdots & P_{0j}^n & \cdots \\ P_{10}^n & P_{11}^n & \cdots & P_{1j}^n & \cdots \\ \vdots & \vdots & \ddots & \vdots & \\ P_{i0}^n & P_{i1}^n & \cdots & P_{ij}^n & \cdots \\ \vdots & \vdots & & \vdots & \ddots \end{bmatrix} \cdot \begin{bmatrix} P_{00}^m & P_{01}^m & \cdots & P_{0j}^m & \cdots \\ P_{10}^m & P_{11}^m & \cdots & P_{1j}^m & \cdots \\ \vdots & \vdots & \ddots & \vdots & \\ P_{i0}^m & P_{i1}^m & \cdots & P_{ij}^m & \cdots \\ \vdots & \vdots & & \vdots & \ddots \end{bmatrix}$$

And this implies the follow equation. Let $P^{(0))} = I$, then $P^{(n)} = P^{(n)} \cdot P^{(0)}$.

$$P^{(n)} = P \cdot P^{(n-1)} = P \cdot P \cdot P^{(n-2)} = \cdots = P^n$$

Proof

$$\begin{aligned} P_{ij}^{n+m} &= P\left\{X_{n+m} = j \mid X_0 = i\right\} \\ &= \sum_{k=0}^{\infty} P\left\{X_{n+m} = j, X_n = k \mid X_0 = i\right\} \\ &= \sum_{k=0}^{\infty} P\left\{X_{n+m} = j \mid X_n = k, X_0 = i\right\} P\left\{X_n = k \mid X_0 = i\right\} \\ &= \sum_{k=0}^{\infty} P_{kj}^m P_{ik}^n \end{aligned}$$

■

2 States and Class

Definition 2.1 (Communicate: $i \leftrightarrow j$)

Two states i and j accesible to each other are said to communicate, and we write $i \leftrightarrow j$.

Lemma 2.1 (Communicate's property)

- $i \leftrightarrow i$
- *if $i \leftrightarrow j$, then $j \leftrightarrow i$*
- *if $i \leftrightarrow j$ and $j \leftrightarrow k$, then $i \leftrightarrow k$*

Definition 2.2 (Class and class property)

Two states that communicate are said to be in the same class. Note that any two classes are either disjoint or identical. A property is called a class property if the property is shared by states in the same class.

Definition 2.3 (Irreducible Markov chain)

A Markov chain is said to be irreducible if there is only one class.

Definition 2.4 (Periodicity)

State i is said to have period d if $P_{ii}^n = 0$ whenever n is not divisible by d and d is the greatest common divisor of $\{n : P_{ii}^n > 0\}$.

Remark The system cannot go back to the original state if $n \neq kd$, where k is an non-negative integer; otherwise, the system may go back to the original state.

Definition 2.5 (Aperiodic)

A state with period 1 is said to be aperiodic.

Lemma 2.2

Let $d(i)$ denote the period of i, this is a class property. That is, if $i \leftrightarrow j$, then $d(i) = d(j)$.

3 Recurrency

Definition 3.1 (f_{ij}^n and f_{ij})

f_{ij}^n is the probability that, starting in i, the first transition into j occurs at time n.

$$f_{ij}^n = P\{X_n = j, X_k \neq j, k = 1, \ldots, n-1 \mid X_0 = i\}$$

f_{ij} denotes the probability of ever making a transition into state j, given that the process starts i.

$$f_{ij} = \sum_{n=1}^{\infty} f_{ij}^n$$

Remark For $i \neq j$, f_{ij} is positive iff j is accessible from i.

Definition 3.2 (Recurrent and transient)

State j is said to be recurrent if $f_{jj} = 1$, and transient otherwise.

Lemma 3.1 (Recurrency's and Transient's property)

- *State j is recurrent iff $\sum_{n=1}^{\infty} P_{jj}^n = \infty$, that is,*

$$E\,[\,\text{\# of visits to } j \mid X_0 = j] = \infty$$

- *State j is transient, then each time the process returns to j with a fail probability*

$1 - f_{jj}$, *hence the number of visits is geometric with finite mean* $m_{jj} = 1/(1 - f_{jj})$.

- *If* i *is recurrent and* $i \leftrightarrow j$, *then* j *is recurrent. (Class property)*
- *If* $i \leftrightarrow j$ *and* j *is recurrent, then* $f_{ij} = 1$.

Lemma 3.2

A finite-state Markov chain must have at least one of the states being recurrent.

Example 3.1 uplow, 2021 For example, states 1 and 2 are recurrent, and states 3 and 4 are transient.

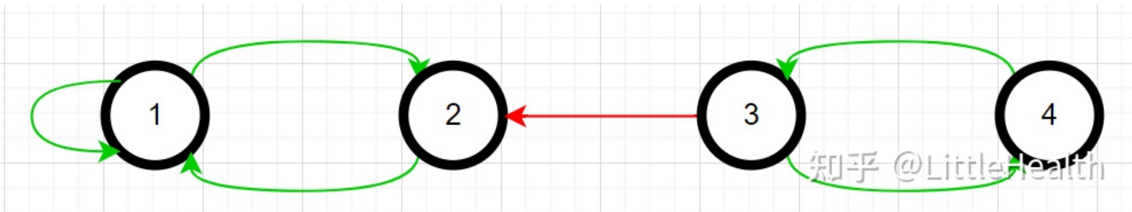

Figure 1: Recurrent vs. Transient

Definition 3.3 (Expected # of transitions for return μ_{jj})

Let μ_{jj} *denote the expected number of transitions needed to return to state* j.

$$\mu_{jj} = \begin{cases} \infty & \text{if } j \text{ is transient} \\ \sum_{n=1}^{\infty} n f_{jj}^{n} & \text{if } j \text{ is recurrent.} \end{cases}$$

Remark Note that the expected times to go back to state j is different from the expected times to visit state j.

Definition 3.4 (Positive and Null Recurrence)

If state j *is recurrent, then we say that it is positive recurrent if* $\mu_{jj} < \infty$ *and null recurrent if* $\mu_{jj} = \infty$.

Lemma 3.3

Positive (Null) recurrence is a class property.

Example 3.2 uplow, 2021 Consider a Markov chain which goes back to state 1 when $n = 2^k$, where k is an non-negative integer.

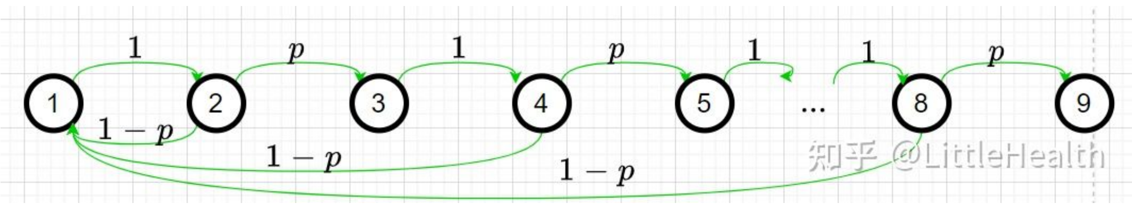

Figure 2: Positive and Null Recurrence

1. When $p = 1/2$, $f_{11}^{(2)} = \frac{1}{2}, f_{11}^{(4)} = \frac{1}{2^2}, f_{11}^{(8)} = \frac{1}{2^3}, \ldots$ and other f_{11}^{i} are 0. So $f_{11} = \sum_{i=}^{\infty} \frac{1}{2^i} = 1$ and $\mu_1 = \sum_{i=}^{\infty} \frac{1}{2^i} 2^i = \infty$.
2. When $p = 1/4$, $f_{11}^{(2)} = \frac{3}{4}, f_{11}^{(4)} = \frac{3}{4^2}, f_{11}^{(8)} = \frac{3}{4^3}, \ldots$ and other f_{11}^{i} are 0. So $f_{11} = \sum_{i=}^{\infty} \frac{3}{4^i} = 1$ and $\mu_1 = 3 \sum_{i=}^{\infty} \frac{1}{4^i} 2^i = 3$.

The key part to distinguish between positive and null recurrence is that if the series of the product of f^i and i, i.e., the probability of going back to the state at the i times and the times needed to go back to the state, converges, then the state is positive recurrent; otherwise if the series diverges, then the state is null recurrent.

Definition 3.5 (Ergodic State)

A positive recurrent, aperiodic state is called ergodic.

Lemma 3.4 (Finite irreducible means all positive recurrent)

For any arbitrary irreducible Markov chain with a finite number of states, all states, denoted by $\{0, 1, \ldots, M\}$ are positive recurrent.

Proof

$$\text{finite states} \rightarrow \text{at least one recurrent state} \xrightarrow{\text{irreducible}} \text{all states recurrent}$$

■

4 Stationary Distribution

Definition 4.1 (Stationary Distribution)

A probability distribution P_j is said to be stationary for the Markov chian if

$$P_j = \sum_{i=0}^{\infty} P_i P_{ij}, \quad j \geq 0$$

Lemma 4.1 (Property of Stationary Distribution)

If the probability distribution of X_0 ($P_j = P\{X_0 = j\}, j \geq 0$) is a stationary distribution, then X_n will have the same distribution (stationary distribution) for all n.

Remark

$$P\{X_1 = j\} = \sum_{i=0}^{\infty} P\{X_1 = j, X_0 = i\} = \sum_{i=0}^{\infty} P\{X_1 = j \mid X_0 = i\} P\{X_0 = i\} = \sum_{i=0}^{\infty} P_i P_{ij} = P_j.$$

By induction,

$$P\{X_n = j\} = \sum_{i=0}^{\infty} P\{X_n = j, X_{n-1} = i\} = \sum_{i=0}^{\infty} P\{X_n = j \mid X_{n-1} = i\} P\{X_{n-1} = i\} = \sum_{i=0}^{\infty} P_i P_{ij} = P_j$$

Theorem 4.1 (Class of Irreducible Aperiodic Markov Chain)

An irreducible aperiodic Markov chain belongs to one of the following two classes:

1. *Either the states are all transient or all null recurrent, in this case,* $P_{ij}^n \to 0$ *as* $n \to \infty$ *for all* i, j *and there exists no stationary distribution.*
2. *Or else, all states are positive recurrent, that is,*

$$\pi_j = \lim_{n\to\infty} P_{ij}^n = 1/\mu_{jj} > 0$$

In this case, $\{\pi_j, j = 0, 1, 2, \ldots\}$ *is a stationary distribution and there exists no other stationary distribution.*

Lemma 4.2 (Property of π_j)

π_j *must be interpreted as the long-run proportion of time that the Markov chain is in state* j*, and*

$$\pi_j = \sum_i \pi_i P_{ij}, \quad \sum_j \pi_j = 1$$

Theorem 4.2 (Interpreting a Markov chain as a renewal process)

Let $N_j(t)$ *denote the number of transitions into* j *by time* t*.*

- *If* j *is recurrent and* $X_0 = j$*, then* $N_j(t)$ *is a renewal process with interarrival distribution* $\{f_{jj}^n, n \geq 1\}$*.*
- *If* $X_0 = i, i \leftrightarrow j$ *and* j *is recurrent, then* $N_j(t)$ *is a delayed renewal process with initial interarrival distribution* $\{f_{ij}^n, n \geq 1\}$*.*

If i *and* j *are communicate, then*

- $P\left\{\lim_{t\to\infty} \frac{N_j(t)}{t} = \frac{1}{\mu_{jj}} \mid X_0 = i\right\} = 1$
- $\lim_{n\to\infty} \frac{\sum_{k=1}^n P_{ij}^k}{n} = \frac{1}{\mu_{jj}}$
- *If* j *is aperiodic, then* $\lim_{n\to\infty} P_{ij}^n = \frac{1}{\mu_{jj}} = \pi_j$
- *If* j *has period* d*, then* $\lim_{n\to\infty} P_{jj}^{nd} = \frac{d}{\mu_{jj}} = d\pi_j$

5 Transitions among classes with transient states

Lemma 5.1

Let R *be a recurrent class of states. If* $i \in R$ *and* $j \notin R$*, then* $P_{ij}^m = 0$ *for all* $m \geq 1$*.*

Theorem 5.1 (f_{ij} among classes)

Let j *be a given recurrent state and let* T *denote the set of all transient states. The set of probabilities* $\{f_{ij}, i \in T\}$ *satisfies*

$$f_{ij} = \sum_{k\in T} P_{ik} f_{kj} + \sum_{k\in R} P_{ik}, \quad i \in T$$

where R *denotes the set of states communicating with* j*.*

Example 5.1Gambler's Ruin Problem Consider a gambler who at each play of the game has probability p of winning 1 unit and probability $q = 1 - p$ of losing 1 unit. Assuming successive plays of the game are independent, we are interested in the probability $f_i (= f_{iN})$ that starting with i units, the gambler's fortune will reach N before reaching 0. Alternatively, we can consider a gambler with wealth i playing against an opponent with wealth $N - i$. In this case, f_i corresponds to the probability that the gambler wins the opponent's wealth. If we let X_n denote the player's fortune at time n, then the process $\{X_n, n = 0, 1, 2, \ldots\}$ is a Markov chain with transition probabilities

$$P_{00} = P_{NN} = 1, \quad P_{i,i+1} = p = 1 - P_{i,i-1}, i = 1, 2, \ldots, N-1$$

Solution *This Markov chain has three classes:* $\{0\}, \{1, 2, \ldots, N-1\}$*, and* $\{N\}$*, the first and third class being recurrent and the second transient.*

$$f_i = pf_{i+1} + qf_{i-1}, \quad i = 1, 2, \ldots, N-1 \quad (\textit{Theorem } 5.1)$$

$$\Updownarrow p + q = 1$$

$$f_{i+1} - f_i = \frac{q}{p}(f_i - f_{i-1}), \quad i = 1, 2, \ldots, N-1$$

$$\begin{aligned}
f_2 - f_1 &= \frac{q}{p}(f_1 - f_0) = \frac{q}{p} f_1 \\
f_3 - f_2 &= \frac{q}{p}(f_2 - f_1) = \left(\frac{q}{p}\right)^2 f_1 \\
&\vdots \\
f_i - f_{i-1} &= \frac{q}{p}(f_{i-1} - f_{i-2}) = \left(\frac{q}{p}\right)^{i-1} f_1 \\
&\vdots \\
f_N - f_{N-1} &= \frac{q}{p}(f_{N-1} - f_{N-2}) = \left(\frac{q}{p}\right)^{N-1} f_1.
\end{aligned}$$

$$\Downarrow \textit{Adding the first } i-1 \textit{ of these equations}$$

$$f_i - f_1 = f_1 \left[\frac{q}{p} + \left(\frac{q}{p}\right)^2 + \cdots + \left(\frac{q}{p}\right)^{i-1}\right], \quad i = 2, 3, \ldots, N$$

or

$$f_i = f_1 \sum_{k=0}^{i-1} \left(\frac{q}{p}\right)^k = \begin{cases} \frac{1-(q/p)^i}{1-q/p} f_1 & \textit{if } q/p \neq 1 \\ i f_1 & \textit{if } q/p = 1 \end{cases} \quad i = 2, 3, \ldots, N$$

Using $f_N = 1$ *yields*

$$f_i = \begin{cases} \frac{1-(q/p)^i}{1-(q/p)^N} & \textit{if } p \neq 1/2 \\ \frac{i}{N} & \textit{if } p = 1/2 \end{cases} \quad i = 0, 1, \ldots, N$$

It is interesting to note that as $N \to \infty$

$$f_i \to \begin{cases} 1-(q/p)^i & \text{if } p > 1/2 \\ 0 & \text{if } p \le 1/2 \end{cases}$$

Theorem 5.2 (Expected # of periods spent in transient state: m_{ij})

Consider a finite state Markov chain and suppose that the states are numbered so that $T = \{1, 2, ..., t\}$ *denotes the set of transient states. For transient states* i *and* j*, let* m_{ij} *denote the expected total number of periods spent in state* j *given that the chain starts in state* i*. Conditioning on the initial transition yields the following equations, where* $\delta(i,j)$ *is equal to 1 when* $i = j$ *and 0 otherwise.*

$$m_{ij} = \delta(i,j) + \sum_k P_{ik} m_{kj} = \delta(i,j) + \sum_{k=1}^{t} P_{ik} m_{kj} \quad m_{kj} = 0 \forall k \notin T$$

Let

$$\mathbf{Q} = \begin{bmatrix} P_{11} & P_{12} & \cdots & P_{1t} \\ \vdots & \vdots & & \vdots \\ P_{i1} & P_{i2} & \cdots & P_{it} \\ \vdots & \vdots & & \vdots \\ P_{t1} & P_{t2} & \cdots & P_{tt} \end{bmatrix} \quad \mathbf{M} = \begin{bmatrix} m_{11} & m_{12} & \cdots & m_{1t} \\ \vdots & \vdots & & \vdots \\ m_{i1} & m_{i2} & \cdots & m_{it} \\ \vdots & \vdots & & \vdots \\ m_{t1} & m_{t2} & \cdots & m_{tt} \end{bmatrix}$$

Then $\mathbf{M} = (\mathbf{I} - \mathbf{Q})^{-1}$.

Proof

$$\mathbf{M} = \mathbf{I} + \mathbf{QM} \to (\mathbf{I} - \mathbf{Q})\mathbf{M} = \mathbf{I} \to \mathbf{M} = (\mathbf{I} - \mathbf{Q})^{-1}$$

■

Lemma 5.2 (Relations of m_{ij} and f_{ij})

$$f_{ij} = m_{ij}/m_{jj} \quad f_{jj} = 1 - 1/m_{jj}$$

Proof

$$\begin{aligned} m_{ij} =& E[\text{ number of transitions into state } j \mid \text{ start in } i] \\ =& E[\text{ number of transitions into state } j \mid \text{ start in } i \text{ and visit } j] f_{ij} \\ &+ E[\text{ number of transitions into state } j \mid \text{ start in } i \text{ and never visit } j]\,(1 - f_{ij}) \\ =& m_{jj} f_{ij} \end{aligned}$$

Then we see $f_{ij} = m_{ij}/m_{jj}$, by Lemma 3.1, we have $f_{jj} = 1 - 1/m_{jj}$. ■

6 Reversed Chain and Time reversible

Definition 6.1 (Stationary Chain or Steady state)

An irreducible positive recurrent Markov chain is stationary if the initial state is chosen according to the stationary probabilities. We say that such a chain is in steady state.

Remark In the case of an ergodic chain, i.e., irreducible, positive recurrent, and aperiodic, this is equivalent to imagining that the process begins at time $t = -\infty$.

Theorem 6.1 (Reversed chain)

Consider an irreducible stationary Markov chain with transition probabilities P_{ij}. If one can find nonnegative numbers $\pi_i, i \geq 0$, summing to unity, and a transition probability matrix $P^* = \left[P_{ij}^*\right]$ such that

$$\pi_i P_{ij} = \pi_j P_{ji}^*$$

then the $\pi_i, i \geq 0$ are the stationary probabilities and P_{ij}^* are the transition probabilities of the reversed chain.

Remark This is useful in solving π_i.

Definition 6.2 (Time reversible)

If $P_{ij}^* = P_{ij}$ for all i, j, then the Markov chain is said to be time reversible. That is,

$$\pi_i P_{ij} = \pi_j P_{ji} \quad \text{for all } i, j$$

For all states i, j, $\pi_i P_{ij}$ means the rate at which the process goes from i to j, $\pi_j P_{ji}$ means the rate at which it goes from j to i.

Theorem 6.2 (Time reversible's condition)

A stationary Markov chain is time reversible iff starting in state i, any path back to i has the same probability as the reversed path, for all i. That is,

$$P_{i,i_1} P_{i_1,i_2} \cdots P_{i_k,i} = P_{i,i_k} P_{i_k,i_{k-1}} \cdots P_{i_1,i} \quad \text{for all states } i, i_1, \ldots, i_k$$

7 Random Walk

Example 7.1General Random Walk Let X_i be i.i.d with $P\{X_i = j\} = a_j, j = 0, \pm 1, \ldots$. And let $S_0 = 0, S_n = \sum_{i=1}^{n} X_i = S_{n-1} + X_n$, then $\{S_n, n \geq 0\}$ is called the general random walk. $\{S_n, n \geq 0\}$ is a Markov chain because S_{n+1} depends on S_n and is independent of S_i for all $i < n$.

$$\begin{aligned} P_{ij} &= P\left\{S_{n+1} = j \mid S_n = i\right\} = P\left\{S_n + X_{n+1} = j \mid S_n = i\right\} \\ &= P\left\{X_{n+1} = j - i\right\} = a_{j-i} \end{aligned}$$

Bibliography

uplow (2021). *How to Understand Null Recurrency*. URL: https://www.zhihu.com/question/46539491/answer/263442039 (visited on 01/12/2023).