Lecture notes 4
Markov Chain: Introduction

Whatever happened in the past, be it glory or misery, be Markov!

1. Introduction

(1a) General Definitions/Descriptions

- *Stochastic Process*: a family of random variables $\{X_t\}$ index by t.
- *State space*: the set of values of the stochastic processes, which in general does not have to be real numbers.
- *Markov Process*: A stochastic process $\{X_t\}$ indexed by time t such that, at each time t, the future of the process $\{X_s : s > t\}$ is conditionally independent of the past of the process $\{X_s : s < t\}$ given the present of the process X_t taking any fixed value. Another interpretation is: at each time t, the future of the process $\{X_s : s > t\}$ depends on the the past of the process $\{X_s : s < t\}$ only through the present X_t.
  Caution: $X_s : s > t$ is in general not independent of $X_s : s < t$.
- *Markov chain (MC)*: Markov process with discrete state space. Discrete state space is usually denoted by numbers $0, 1, 2, ...$
- *Discrete/continuous time Markov chain*: Discrete time MC: the time domain is $\{0, 1, 2, ...\}$. Continuous time MC: the time domain is $[0, \infty)$.

(1b) *Markov chain*

In this chapter, we focus on *Markov chain*, a stochastic process $\{X_n, n = 0, 1, 2, \ldots\}$ that takes on a finite or countable number of possible values. Unless otherwise mentioned, this set of possible values of the process will be denoted by the set of nonnegative integers $\{0, 1, 2, \ldots\}$.

If $X_n = i$, the process is said to be in state i at time n. We suppose that whenever the process is in state i, there is a fixed probability P_{ij} that it will next be in state j. That is

$$\begin{aligned} & P(X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) \\ = \quad & P(X_{n+1} = j | X_n = i) \end{aligned}$$

for all states $i_0, i_1, \ldots, i_{n-1}, i, j$ and all $n \geq 0$. Let $P_{ij} = P(X_{n+1} = j | X_n = i)$. In other words, the *one-step transition probability* is irrelevant with n, i.e., it's the same for all n, we call the MC $\{X_t\}$ a MC with *stationary* transition probabilities. Throughout the course, we only consider MC with stationary transition probabilities.

Note that

$$P_{ij} \geq 0,\ \ i, j \geq 0;\ \ \sum_{j=0}^{\infty} P_{ij} = 1,\ \ i = 0, 1, 2, \ldots$$

Let $\mathbf{P}$ denote the one-step transition probabilities matrix, or *transition matrix* in brief, which is

$$\mathbf{P} = \begin{array}{c} 0 \\ 1 \\ 2 \\ \vdots \end{array} \begin{pmatrix} P_{00} & P_{01} & P_{02} & \dots \\ P_{10} & P_{11} & P_{12} & \dots \\ P_{20} & P_{21} & P_{22} & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Example 1:
(Forecasting Weather) Suppose that the chance of rain tomorrow depends on previous weather conditions only through whether or not it is raining today and not on past weather conditions. Suppose also that if it rains today, then it will rain tomorrow with probability 0.2; and if it does not rain today, then it will rain tomorrow with probability 0.4.

If we say that the process X_n (X_n denotes the weather on the n-th day) is in

state 0 when it rains and
state 1 when it does not rain,

then X_n is a two-state Markov chain whose transition probabilities are given by

$$\mathbf{P} = \left\| \begin{array}{cc} 0.2 & 0.8 \\ 0.4 & 0.6 \end{array} \right\|$$

Example 2:
(A communications system) Consider a communications system which transmits the digits 0 and 1. Each digit transmitted must pass through several stages, at each of which there is a probability p that the digit entered will be unchanged when it leaves.

Letting X_n denote the digit entering the n-th stage, then $\{X_n, n = 0, 1, \ldots\}$ is a two-state Markov chain having a transition probability matrix

$$\mathbf{P} = \left\| \begin{array}{cc} p & 1-p \\ 1-p & p \end{array} \right\|$$

Example 3:
On any given day Gary is either cheerful(C), so-so(S), or glum(G). If he is cheerful today, then he will be C, S or G tomorrow with respective probabilities 0.5, 0.4, 0.1. If he is feeling so-so today, then he will be C, S or G tomorrow with probabilities 0.3, 0.4, 0.3. If he is glum today, then he will be C, S or G tomorrow with probabilities 0.2, 0.3, 0.5.

Letting X_n denote Gary's mood on the n-th day, then $\{X_n, n \geq 0\}$ is a three-state Markov chain (state $0 = C$, state $1 = S$, state $2 = G$) with transition probability matrix

$$\mathbf{P} = \begin{Vmatrix} 0.5 & 0.4 & 0.1 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.3 & 0.5 \end{Vmatrix}$$

Example 4:
(Transforming a Process into a Markov chain) Suppose that whether or not it rains today depends on previous weather conditions through the last two days. Specifically, suppose that if it has rained for the past two days, then it will rain tomorrow with probability 0.7; if it rained today but not yesterday, then it will rain tomorrow with the probability 0.5; if it rained yesterday but not today, then it will rain tomorrow with probability 0.4; if it has not rained in the past two days, then it will rain tomorrow with probability 0.2.

If we let the state at time n depend only on whether or not it is raining at time n, then the preceding model is not a Markov chain (Why not?). However, we can transform this model into a Markov chain in the following way, by saying that the state at any time is determined by the weather conditions during both that day and the previous day.

In other words, we can say that the process is in:

state 0	if it rained both today and yesterday
state 1	if it rained today but not yesterday
state 2	if it rained yesterday but not today
state 3	if it did not rain either yesterday or today

The preceding would then represent a four-state Markov chain having a transition probability matrix

$$\mathbf{P} = \begin{Vmatrix} 0.7 & 0 & 0.3 & 0 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0.4 & 0 & 0.6 \\ 0 & 0.2 & 0 & 0.8 \end{Vmatrix}$$

We should carefully checked the matrix $\mathbf{P}$, and make sure you understand how it was obtained.

Example 5:
(A random walk model) A Markov chain whose state space is given by the integers $i = 0, \pm 1, \pm 2, \ldots$ is said to be a random walk if, for some number $0 < p < 1$,

$$P_{i,i+1} = p = 1 - P_{i,i-1}, \quad i = 0, \pm 1, \ldots$$

The preceding Markov chain is called a random walk as we may think of it as being a model for an individual walking on a straight line who at each point of time either takes one step to the right with probability p or one step to the left with the probability $1 - p$.

Example 6:
(A gambling model) Consider a gambler who, at each play of the game, either wins \$1 with probability p or losses \$1 with probability $1 - p$. If we suppose that our gambler quits playing either when he goes broke or he attains a fortune of \$N, then the gambler's fortune is a Markov chain having transition probabilities

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

$$P_{00} = P_{NN} = 1$$

States 0 and N are called *absorbing* states since once entered they are never left. Note that the preceding is a finite state random walk with absorbing barriers (states 0 and N).

2. Chapman-Kolmogorov Equations

Define the n-step transition probabilities P_{ij}^n to be the probability that a process is in state i will be in state j after n additional transitions. That is

$$P_{ij}^n = P(X_{n+m} = j \mid X_m = i), \quad n \geq 0, \; i, j \geq 0.$$

Of course $P_{ij}^1 = P_{ij}$. Note that

$$P_{ij}^{n+m} = \sum_{k=0}^{\infty} P_{ik}^n P_{kj}^m$$

for all $n, m \geq 0$, all i, j.

Formally we have

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

If we let $\mathbf{P}^{(n)}$ denote the matrix of n-step transition probabilities P_{ij}^n, then

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

In particular,

$$\mathbf{P}^{(2)} = \mathbf{P}^{(1+1)} = \mathbf{P} \cdot \mathbf{P} = \mathbf{P}^2$$

and by induction

$$\mathbf{P}^{(n)} = \mathbf{P}^{(n-1+1)} = \mathbf{P}^{(n-1)} \cdot \mathbf{P} = \mathbf{P}^n$$

Example 7:
Consider Example 1 in which the weather is considered as a two-state Markov chain. If $\alpha = 0.7$ and $\beta = 0.4$, then calculate the probability that it will rain four days from today given that it is raining today.

Solution: The one-step transition probability matrix is given by

$$P = \begin{Vmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{Vmatrix}$$

Hence,

$$P^{(2)} = P^2 = \left\| \begin{array}{cc} 0.7 & 0.3 \\ 0.4 & 0.6 \end{array} \right\| \cdot \left\| \begin{array}{cc} 0.7 & 0.3 \\ 0.4 & 0.6 \end{array} \right\| = \left\| \begin{array}{cc} 0.61 & 0.39 \\ 0.52 & 0.48 \end{array} \right\|$$

$$P^{(4)} = (P^2)^2 = \left\| \begin{array}{cc} 0.61 & 0.39 \\ 0.52 & 0.48 \end{array} \right\| \cdot \left\| \begin{array}{cc} 0.61 & 0.39 \\ 0.52 & 0.48 \end{array} \right\| = \left\| \begin{array}{cc} 0.5749 & 0.4251 \\ 0.5668 & 0.4332 \end{array} \right\|$$

and the desired probability P_{00}^4 equals 0.5749.

Example 8:
Consider Example 4. Given that it rained on Monday and Tuesday, what is the probability that it will rain on Thursday?

Solution: The two-step transition matrix is given by

$$P^{(2)} = P^2 = \left\| \begin{array}{cccc} 0.7 & 0 & 0.3 & 0 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0.4 & 0 & 0.6 \\ 0 & 0.2 & 0 & 0.8 \end{array} \right\| \cdot \left\| \begin{array}{cccc} 0.7 & 0 & 0.3 & 0 \\ 0.5 & 0 & 0.5 & 0 \\ 0 & 0.4 & 0 & 0.6 \\ 0 & 0.2 & 0 & 0.8 \end{array} \right\| = \left\| \begin{array}{cccc} 0.49 & 0.12 & 0.21 & 0.18 \\ 0.35 & 0.20 & 0.15 & 0.30 \\ 0.10 & 0.12 & 0.20 & 0.48 \\ 0.20 & 0.16 & 0.10 & 0.64 \end{array} \right\|$$

Since rain on Thursday is equivalent to the process being in either state 0 or state 1 on Thursday, the desired probability is given by: $P_{00}^2 + P_{01}^2 = 0.49 + 0.12 = 0.61$.

If the unconditional distribution of the state at time n is desired, it is necessary to specify the probability distribution of the initial state. Let us denote this by

$$\alpha_i \equiv P(X_0 = i),\ i \geq 0 (\sum_{i=0}^{\infty} \alpha_i = 1)$$

All unconditional probabilities may be computed by conditioning on the initial state. That is

$$\begin{aligned} P(X_n = j) &= \sum_{i=0}^{\infty} P(X_n = j \mid X_0 = i) P(X_0 = i) \\ &= \sum_{i=0}^{\infty} P_{ij}^n \alpha_i \end{aligned}$$

For example, if $\alpha_0 = 0.4$ and $\alpha_1 = 0.6$ in Example 7, then the (unconditional) probability that it will rain four days after we begin keeping weather records is

$$\begin{aligned} P(X_4 = 0) &= 0.4 P_{00}^4 + 0.6 P_{10}^4 \\ &= 0.4(0.5749) + 0.6(0.566) = 0.57 \end{aligned}$$

3. Classification of States

Definition 1: State j is said to be accessible from state i if $P_{ij}^n > 0$ for some $n \geq 0$.

Note that this implies state j is accessible from state i if and only if, starting in i, it is possible that the process will ever enter state j. This is true since if j is not accessible from i, then

$$\begin{aligned} P(\text{ever enter } j \mid \text{start in } i) &= P[\cup_{n=0}^{\infty}\{X_n = j\} \mid X_0 = i] \\ &\leq \sum_{n=0}^{\infty} P(X_n = j \mid X_0 = i) \\ &= \sum_{n=0}^{\infty} P_{ij}^n = 0. \end{aligned}$$

Definition 2: Two states i and j that are accessible to each other are said to communicate, and we write $i \leftrightarrow j$.

Note that $P_{ii}^0 = 1$.

The relation of communication satisfies the following three properties:

(a) State i communicates with state i, all $i \geq 0$.
(b) If state i communicates with state j, then state j communicates with state i.
(c) If state i communicates with state j, and state j communicates with state k, then state i communicates with state k.

Properties (a) and (b) follow immediately from the definition of communication. To prove (c), suppose that i communicates with j, and j communicates with k. Then there exists integers n and m such that $P_{ij}^n > 0$ and $P_{jk}^m > 0$. By the Chapman-Kolmogorv equations,

$$P_{ik}^{n+m} = \sum_{r=0}^{\infty} P_{ir}^n P_{rk}^m \geq P_{ij}^n P_{jk}^m > 0.$$

Hence, state k is accessible from state i. By the same argument, we can show that state i is accessible from state k. Hence, state i and k communicate.

Definition 3: Two states that communicate are said to be in the same class.

It is an easy consequence of the (a), (b) and (c) that any two classes of states are either identical or disjoint. In other words, the concept of communication divides the state space up into a number of separate classes.

Definition 4: The Markov chain is said to be irreducible if there is only one class. i.e. if all states communicate with each other.

Example 9:
Consider the Markov chain consisting of the three states 0,1,2 and having transition probability matrix

$$P = \begin{Vmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 1/4 & 1/4 \\ 0 & 1/3 & 2/3 \end{Vmatrix}$$

It is easy to verify that this Markov chains irreducible. For example, it is possible to go from state 0 to state 2 since

$$0 \ \rightarrow 1 \ \rightarrow 2$$

That is, one way of getting from state 0 to state 2 is to go from state 0 to state 1 (with probability 1/2) and then go from state 1 to state 2 (with probability 1/4).

Example 10:
Consider the Markov chain consisting of the four states 0,1,2,3 and having transition probability matrix

$$P = \begin{Vmatrix} 1/2 & 1/2 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 0 & 0 & 1 \end{Vmatrix}$$

The classes of this Markov chain are $\{0{,}1\}$,$\{2\}$ and $\{3\}$. Note that while state 0 (or 1) is accessible from state 2, the reverse is not true. Since state 3 is an absorbing state, that is, $P_{33} = 1$, no other state is accessible from it.

Definition 5: For any state i we let f_i denote the probability that, starting in state i, the process will ever reenter state i. State i is said to be recurrent if $f_i = 1$ and transient if $f_i < 1$.

Suppose that the process starts in state i and i is recurrent. Then, with probability 1, the process will even re-enter state i. However, by the definition of Markov Chain, it follows that the process will be starting over again when it re-enters state i and, therefore, state i will eventually be visited again. Thus we have

Proposition 1:
If state i is recurrent, then starting in start i, the process will re-enter state i again, again

and again - in fact, infinitely often.

On the other hand, suppose that state i is transient. Hence, each time the process enters state i, there will be a positive probability, namely $1 - f_i$ that it will never enter state i again. Therefore, starting in state i, the probability that the process will be in state i for exactly n time periods equals to $f_i^{n-1}(1 - f_i),\ n \geq 1$. Thus we have

Proposition 2:
if state i is transient then, starting in state i, the number of time periods that the process will be in state i has a geometric distribution with finite mean $\frac{1}{1-f_i}$.

Moreover, based on the last two Propositions, we have

Proposition 3: state i is recurrent if and only if, starting in state i, the expected number of time periods that the process is in state i is infinite.

Theorem 1:
State i is recurrent if

$$\sum_{n=1}^{\infty} P_{ii}^n = \infty$$

and state i is transient if

$$\sum_{n=1}^{\infty} P_{ii}^n < \infty$$

This leads to the conclusion that in a finite-state Markov chain not all states can be transient.

Proof of Theorem: Let

$$A_n = \begin{cases} 1, & if\ X_n = i \\ 0, & if\ X_n \neq i \end{cases}$$

Then $\sum_{n=0}^{\infty} A_n$ is the number of periods that the process is in state i.

$$\begin{aligned} E(\sum_{n=0}^{\infty} A_n \mid X_0 = i) &= \sum_{n=0}^{\infty} E(A_n \mid X_0 = i) \\ &= \sum_{n=0}^{\infty} P(X_n = i \mid X_0 = i) \\ &= \sum_{n=0}^{\infty} P_{ii}^n \end{aligned}$$

Thus the result follows.

Corollary:
If state i is recurrent, and state i communicates with state j, then state j is recurrent.

Proof: To prove this we first note that, since state i communicates with state j, there exists integers k and m such that $P_{ij}^k > 0$ and $P_{ji}^m > 0$. Now, for any integer n,

$$P_{jj}^{m+n+k} \geq P_{ji}^m P_{ii}^n P_{ij}^k$$

This follows since the left side of the above is the probability of going from j to j in $m+n+k$ steps, while the right side is the probability of going from j to j in $m+n+k$ steps via a path that goes from j to i in m steps, then from i to i in an additional n steps, then from i to j in an additional k steps.
By the above Proposition, by summing over n, we have

$$\sum_{n=1}^{\infty} P_{jj}^{m+n+k} \geq \sum_{n=1}^{\infty} P_{ji}^m P_{ii}^n P_{ij}^k = P_{ji}^m P_{ij}^k \sum_{n=1}^{\infty} P_{ii}^n = \infty.$$

since $P_{ij}^k > 0$, $P_{ji}^m > 0$ and $\sum_{n=1}^{\infty} P_{ii}^n = \infty$. Therefore, state j is recurrent.

Remarks:

(a) Corollary also implies that transience is a class property. For if state i is transient and communicates with state j, then state j must also be transient. For if j were recurrent then, by corollary, i would also be recurrent and hence could not be transient.

(b) Corollary along with our previous result that not all states in a finite Markov chain can be transient leads to the conclusion that all state of a finite irreducible Markov chain are recurrent.

Example 11:
Let the Markov chain consisting of the four states 0,1,2,3 and having transition probability matrix

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

Determine which states are transient and which are recurrent.

Solution: It is a simple matter to check that all states communicate and, hence, since it is a finite chain, all states must be recurrent.

Example 12:
Consider the Markov chain having states 0,1,2,3,4 and

$$P = \begin{Vmatrix} 1/2 & 1/2 & 0 & 0 & 0 \\ 1/2 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 1/2 & 0 \\ 0 & 0 & 1/2 & 1/2 & 0 \\ 1/4 & 1/4 & 0 & 0 & 1/2 \end{Vmatrix}$$

Determine the recurrent state.

Solution: This claim consists of three classes {0,1},{2,3}, and {4}. The first two classes are recurrent and the third transient.

4. First step analysis

A journey of a thousand miles begins with the first step. –Lao Tse

Example 13. (Coin Tossing) Repeatedly toss a fair coin a number of times. What's the expected number of tosses till the first two consecutive heads occur?

Let $\xi_0 = 0$ and, for $i \geq 1$,

$$\xi_i = \begin{cases} 1 & \text{if the } i\text{-th toss is a Head (with probability } 1/2) \\ 0 & \text{if the } i\text{-th toss is a Tail (with probability } 1/2) \end{cases}$$

Set, for $n \geq 1$,

$$X_n = \begin{cases} 0 & \text{if the } n\text{-th toss is tail} \\ 1 & \text{if } \xi_{n-1} = 0, \xi_n = 1 \\ 2 & \text{if } \xi_{n-1} = 1, \xi_n = 1. \end{cases}$$

Then, X_n is a MC with state space $\{0, 1, 2\}$ and transition matrix

$$\mathbf{P} \quad = \quad \begin{array}{c c} & \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{array}{c} 0 \\ 1 \\ 2 \end{array} & \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 \end{pmatrix} \end{array}.$$

For example,

$$P_{00} = P(X_{n+1} = 0 | X_n = 0) = P(\xi_{n+1} = 0 | \xi_n = 0) = P(\xi_{n+1} = 0) = 1/2.$$

For $k \geq 1$, let $T_k = \min\{n \geq 0 : X_{n+k} = 2\}$, which is the minimum number of additional tosses till the first two consecutive heads occur, starting from (excluding) the k-th toss. Then

$$T_k = \begin{cases} 0 & \text{if } X_k = 2 \\ T_{k+1} + 1 & \text{if } X_k = 0 \text{ or } 1 \end{cases}.$$

Let

$$w_0 = E(T_k | X_k = 0) \qquad w_1 = E(T_k | X_k = 1).$$

Then,

$$\begin{aligned} w_0 &= E(T_1 | X_1 = 0) = E(T_1 1_{\{X_{1+1} = 0 \text{ or } 1 \text{or } 2\}} | X_1 = 0) \\ &= E(T_1 1_{\{X_2 = 0\}} | X_1 = 0) + E(T_1 1_{\{X_2 = 1\}} | X_1 = 0) + E(T_1 1_{\{X_2 = 2\}} | X_1 = 0) \\ &= E(T_1 | X_2 = 0, X_1 = 0) P(X_2 = 0 | X_1 = 0) \\ &\quad + E(T_1 | X_2 = 1, X_1 = 0) P(X_2 = 1 | X_1 = 0) + E(1_{\{X_2 = 2\}} | X_1 = 0) \\ &= E(T_2 + 1 | X_2 = 0, X_1 = 0) P(X_2 = 0 | X_1 = 0) \\ &\quad + E(T_2 + 1 | X_2 = 1, X_1 = 0) P(X_2 = 1 | X_1 = 0) + P_{02} \\ &= (1 + w_0) P_{00} + (1 + w_1) P_{01} + P_{02} \\ &= 1 + w_1 P_{01} + w_0 P_{00} \end{aligned}$$

Likewise,

$$w_1 = 1 + P_{10}w_0 + P_{11}w_1.$$

Together, we have

$$\begin{aligned} w_0 &= 1 + P_{00}w_0 + P_{01}w_1 = 1 + (1/2)(w_0 + w_1) \\ w_1 &= 1 + P_{10}w_0 + P_{11}w_1 = 1 + (1/2)w_0. \end{aligned}$$

Solving the equation, we have $w_0 = 6$ and $w_1 = 4$. Since $X_1 = 0$ or 1 with half chance. The mean number of tosses till the first HH occur is

$$1/2(w_0 + w_1) + 1 = 6.$$

The above example in fact, for the purpose of illustrating the method of first-step analysis, demonstrates a hard way of solving the problem. For this particular problem, there is actually an easier method without invoking the *MC* $\{X_n, n \geq 1\}$, but, rather, directly based on $\{\xi_i, i \geq 1\}$ (Please DIY). The idea is contained in the example called a dice game called craps presented in review and in the following example as well.

Example 14. (Mickey in Maze) Mickey mouse travels in a maze with nine 3×3 cells. The cells are numbered as 0, 1, ..., 8 from left to right and top down. Each step Mickey travels from where it is to one of the surrounding connected cells with equal chance. (please refer to the graph shown in lecture)

(i) Compute the probability Mickey goes to cell 2 before it reaches cell 6, beginning from cell 1.

(ii) Compute the mean number of steps to reach cells 2 or 6, beginning from cell 4.

(iii) Compute the mean number of times Mickey visits cell 2 before reaching cell 6, starting from cell 4.

Let X_n denote the cell number of Mickey at step n. ($X_0 = 4$). Then,

$$\begin{aligned} & P(X_{n+1} = j | X_0 = 0, X_1 = i_1, ..., X_{n-1} = i_{n-1}, X_n = i) \\ = \; & P(X_{n+1} = j | X_n = i). \end{aligned}$$

Suppose currently Mickey is in cell 5, for example, the future movement or path of Mickey is irrelevant with the past movement or path of Mickey. In other words, how Mickey has got to cell 5 in the past has nothing to do with how Mickey would move around in the future. The process $\{X_n : n = 0, 1, 2, ...\}$ is a Markov chain.

For this example:

$$
\mathbf{P} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \end{array}
\begin{array}{c}
\begin{array}{ccccccccc} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \end{array} \\
\left(\begin{array}{ccccccccc}
 & 1/2 & & 1/2 & & & & & \\
1/3 & & 1/3 & & 1/3 & & & & \\
 & 1/2 & & & & 1/2 & & & \\
1/3 & & & & 1/3 & & 1/3 & & \\
 & 1/4 & & 1/4 & & 1/4 & & 1/4 & \\
 & & 1/3 & & 1/3 & & & & 1/3 \\
 & & & 1/2 & & & & 1/2 & \\
 & & & & 1/3 & & 1/3 & & 1/3 \\
 & & & & & 1/2 & & 1/2 &
\end{array}\right)
\end{array}
$$

Solution. (i) Set

$$p_i = P(\text{Mickey reaches cell 2 before reaching 6} \mid \text{starting from i}).$$

Then, for example,

$$
\begin{aligned}
p_0 &= P_{01}p_1 + P_{03}p_3 = 1/2p_1 + 1/2p_3 \\
p_1 &= 1/3p_0 + 1/3p_4 + 1/3p_2 = 1/3(p_0 + p_4) + 1/3 \\
p_2 &= 1 \\
p_3 &= 1/3p_0 + 1/3p_4 + 1/3p_6 = 1/3(p_0 + p_4) \\
p_4 &= 1/4p_1 + 1/4p_3 + 1/4p_5 + 1/3p_7 \\
p_5 &= 1/3p_2 + 1/3p_4 + 1/3p_8 = 1/3(p_0 + p_8) + 1/3 \\
p_6 &= 0 \\
p_7 &= 1/3p_4 + 1/3p_6 + 1/3p_8 = 1/3(p_4 + p_8) \\
p_8 &= 1/2p_5 + 1/2p_7
\end{aligned}
$$

The above linear equations can be solved for the answer

$$p_0 = p_8 = 1/2, \quad p_1 = p_5 = 2/3, \quad p_3 = p_7 = 1/3 \quad p_4 = 1/2.$$

In fact, a short-cut to solving the eight equations is, by the symmetry of the maze, to realize that $p_0 = p_8$, $p_1 = p_5$, $p_3 = p_7$. Together with the fact that $p_2 = 1$ and $p_6 = 0$, the number of equations can be immediately reduced to four about p_0, p_1, p_3 and p_4.

(ii). Set w_i the mean number of steps to reach 2 or 6, starting from cell i. Then, obviously, $w_2 = w_6 = 0$. By symmetry, $w_0 = w_8$, $w_1 = w_3 = w_5 = w_7$. Moreover,

$$
\begin{aligned}
w_0 &= 1 + 1/2(w_1 + w_3) = 1 + w_1 \\
w_1 &= 1 + 1/3(w_0 + w_2 + w_4) = 1 + 1/3(w_0 + w_4) \\
w_4 &= 1 + 1/4(w_1 + w_3 + w_5 + w_7) = 1 + w_1
\end{aligned}
$$

Solving the three equations,

$$w_0 = w_4 = w_8 = 6 \quad \text{and} \quad w_1 = w_3 = w_5 = w_7 = 5.$$

(iii). Set w_i the mean number of visits of cell 2 before reaching cell 6, starting from cell i (including the starting state). Then, obviously, $w_6 = 0$. By symmetry, $w_0 = w_8$, $w_1 = w_5$ and $w_3 = w_7$. Moreover,

$$\begin{aligned}
w_0 &= 1/2(w_1 + w_3) \\
w_1 &= 1/3(w_0 + w_2 + w_4) \\
w_2 &= 1 + 1/2(w_1 + w_5) = 1 + w_1 \\
w_3 &= 1/3(w_0 + w_4 + w_6) = 1/3(w_0 + w_4) \\
w_4 &= 1/4(w_1 + w_3 + w_5 + w_7) = 1/2(w_1 + w_3)
\end{aligned}$$

Solving the three equations,

$$w_0 = w_4 = w_8 = 3/2 \quad w_1 = w_5 = 2 \quad w_2 = 3 \quad \text{and} \quad w_3 = w_7 = 1.$$

Starting from cell 4, Mickey's mean number of visits of cell 2 before reaching cell 6 is $w_4 = 3/2$.