20

MARKOV CHAIN AND RANDOM WALK

20.1 Discrete-time Markov chain

20.1.1 The model

Definition 20.1.1 (Markov chain). *Consider a discrete stochastic process $\{X_n\}$ with a countable state space E; that is for $\omega \in \Omega, X_n(\omega) \in E$. The discrete stochastic process $\{X_n\}$ is called a Markov chian if*

$$P(X_{n+1} = j | X_0, X_1, ... X_n) = P(X_{n+1} = j | X_n)$$

Remark 20.1.1. A Markov chain is a sequence of random variables such that for any n, X_{n+1} is conditionally independent of $X_0, X_1, ... X_{n-1}$ given X_n. Note that this is simply conditionally independence, because X_{n+1} is generally not independent of any $X_i, i \leq n$.

Definition 20.1.2 (Markov chain, stochastic matrix). *A stationary finite-state Markov chain can also be denoted as a pair (S, P), where S is the finite state space, P is the transition probability matrix with $P_{i,j} = P(X_{n+1} = j | X_n = i), X_i \in S$. P is also known as the* ***stochastic matrix****, with row sum equals 1.*

Remark 20.1.2 (The stochastic matrix can be quite 'irregular'). The stochastic matrix P associated with a Markov chain can have:

- Eigenvalues can be positive and negative, zero and even complex.
- The matrix might not be diagonalized.
- Absolute values of the eigenvalues must be bounded by 1.

See section 20.5 for more details.

Example 20.1.1. In Figure 20.1.1, we show some example Markov chains and their stochastic matrix representation.

- $$P = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}.$$
- $$P = \begin{bmatrix} 0.5 & 0.5 \\ 0 & 1 \end{bmatrix}.$$

- $$P = \begin{bmatrix} 0.5 & 0 & 0.5 \\ 0.5 & 0.5 & 0 \\ 0 & 0.5 & 0.5 \end{bmatrix}.$$
- $$P = \begin{bmatrix} 0.4 & 0.2 & 0.2 \\ 0.2 & 0.4 & 0.2 \\ 0.2 & 0.2 & 0.4 \end{bmatrix}.$$

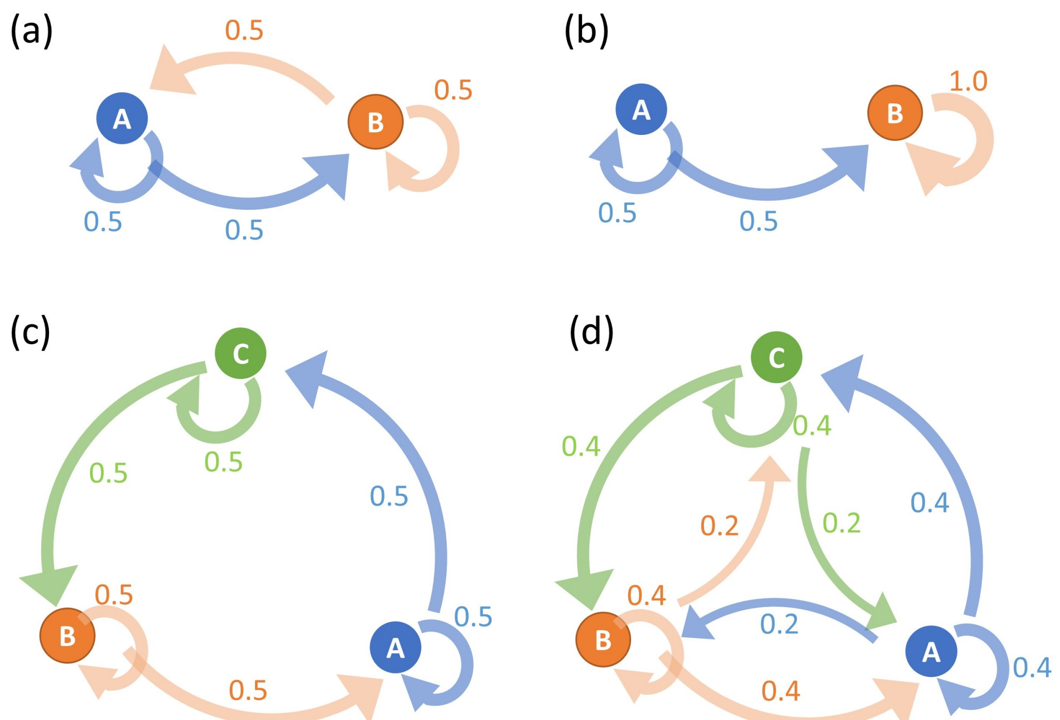


Figure 20.1.1: Example Markov chains. Arrows and numbers are transition directions and probabilities.

Example 20.1.2 (random walk). Figure 20.1.2 Markov chain diagram for a random walk on the state space $\mathbb{Z}$, its transition matrix is given by

$$\begin{bmatrix} \cdots & 1-p & 0 & p & 0 & 0 & \ldots \\ \cdots & 0 & 1-p & 0 & p & 0 & \ldots \\ \cdots & 0 & 1-p & 0 & p & 0 & \ldots \\ \cdots & 0 & 0 & 1-p & 0 & p & \cdots \\ & & & & & \ddots & \end{bmatrix}$$

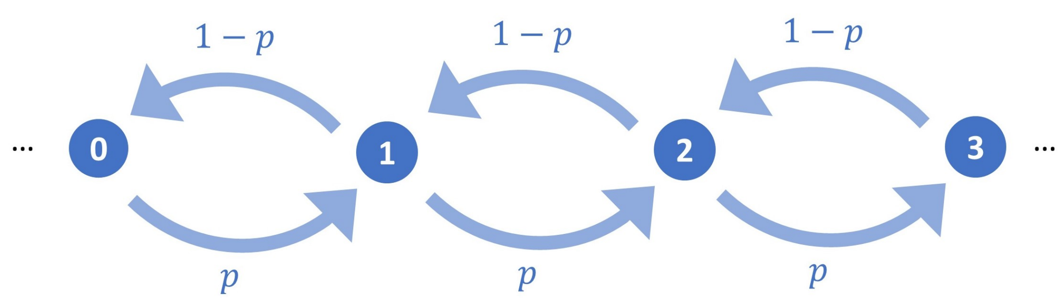


Figure 20.1.2: Markov chain diagram for a random walk on the state space $\mathbb{Z}$.

20.1.2 Evolution of discrete chain

Lemma 20.1.1 (Chapman-Kolmogorov Equation). *Let P be the transition matrix of a Markov chain with state space S.*

$$P(X_n = j|X_0 = i) = \sum_k P(X_n = j|X_m = k)P(X_m = k|X_0 = i)$$

where $1 \leq m \leq n-1$.

Proof. Use the law of total probability [Theorem 11.2.1]. □

Lemma 20.1.2 (Kolmogorov forward equation). *Let P be the transition matrix of a Markov chain and $x(n)$ be the probability row vector at step n, we have*

- $x(n+1) = x(n)P.$

- $x(n) = x(0)P^n$.

Proof. (1) Use Chapman-Kolmogorov Equation for one step. (2) Iteratively use (1). $\square$

20.2 Classification of states

20.2.1 accessibility and communicating classes

Definition 20.2.1 (accessibility and communicate). *[1, p. 235] Given a Markov chain (S, P), we say that state i can* ***access*** *j if there exist a positive integer m such that $P_{ij}^m > 0$. We say two states i and j* ***communicate*** *if there exist two positive integers k_1 and k_2 such that $p_{ij}^{k_1} > 0$ and $p_{ji}^{k_2} > 0$, i.e., i and j can access each other.*

If i can access j, we write $i \to j$.

Remark 20.2.1.

- In other words, states i and j can **communicate** if both can be reached from the other with positive possibility.
- The property of accessibility is not **symmetric**: if state i can access state j, j not necessarily can access i.

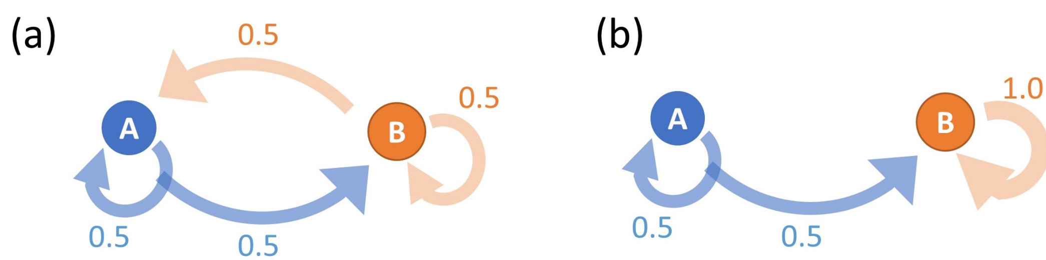


Figure 20.2.1: Demonstration accessibility in a Markov chain. In chain (a), states A and B are accessible to each other or they can communicate. In chain (b), state A can access B but B cannot access A.

Theorem 20.2.1 (partition the state space via communication relationship).

- *Communication is a equivalence relationship;*
- *The state space of a Markov chain can be divided into communicating classes. Each state within a class communicating with every other state in the class, and with no other state.*

Proof. It is easy to show that communication satisfies the transitivity, reflectivity and symmetric conditions.We only show transitivity. If there exists integer m and n for states i, j, k such that

$$P_{ij}^m > 0, P_{jk}^n > 0,$$

then

$$P_{ik}^{m+n} \geq P_{ij}^m P_{jk}^n > 0.$$

That is, i can access k. □

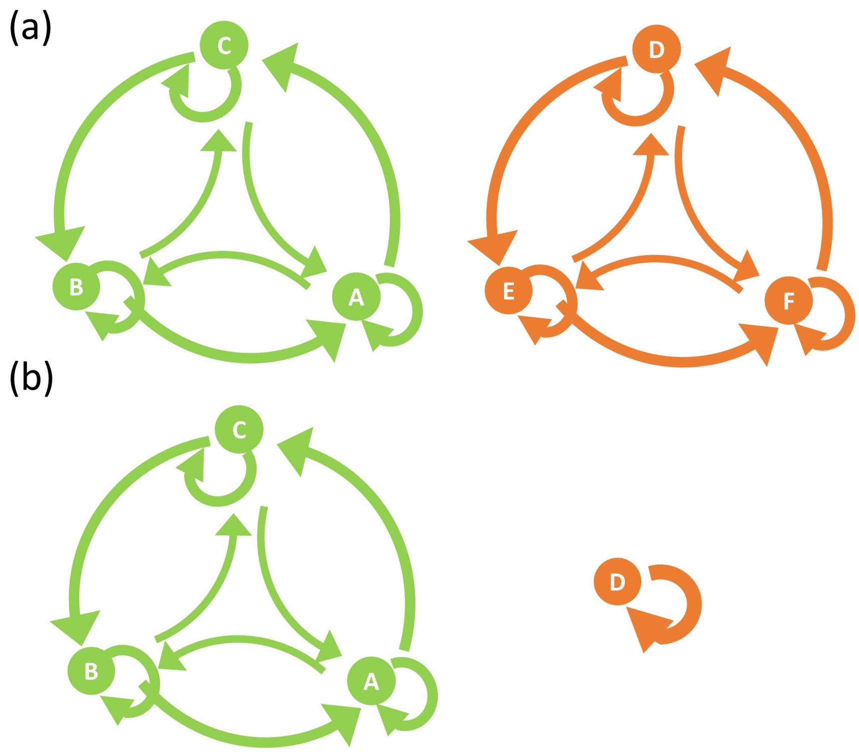


Figure 20.2.2: Demonstration of partitioning state space by communicating classes. Green and orange states belong to different communicating classes. Note that a communicating class can consist of only one state.

20.2.2 Transient and recurrent states and classes

20.2.2.1 *Transient and recurrent states*

We classify each state in the state space by their recurrent or transient properties. Particularly, if a chain starting from a state i, the probability that it will *ever* come back is used to deem the state i is recurrent or transient. More formally,

Definition 20.2.2 (transient state and recurrent state). *Let f_{ii} denote the probability $f_{ii} = P(\text{ever revisit } i | X_0 = i)$. A state i is **transient** is $f_{ii} < 1$ and **recurrent** if $f_{ii} = 1$.*[a]

[a] ever revisit means there exists a finite n such that $P(X_n = i | X_0 = i) = 1$.

Example 20.2.1. Consider a two state Markov chain $S = \{1, 2\}$ with transition matrix

$$P = \begin{bmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{bmatrix}$$

The probability of never visiting state 1 is to stay at 2 forever. We have

$$1 - f_{11} = \lim_{n\to\infty} (0.5)^n = 0 \implies f_{11} = 1.$$

Therefore, these two states (by symmetric reasoning) are classified as recurrent states.

Example 20.2.2. Consider a two state Markov chain $S = \{1, 2\}$ with transition matrix

$$P = \begin{bmatrix} 1 & 0 \\ 0.5 & 0.5 \end{bmatrix}$$

Clearly state 1 is a recurrent state. For state 2, the probability of never visiting state 2 is to visit state 1 at step 1 and stay at 1 forever. We have

$$1 - f_{22} = 0.5 \implies f_{22} = 0.5.$$

Therefore, state 2 is classified as transient state.

Lemma 20.2.1 (expected number of visits).

- *the expected number of visits to state i starting from state i is given by*

$$E[\# \text{ of visits to } i | X_0 = i] = \frac{1}{1 - f_{ii}}$$

- *the expected number of visits to state i starting from state i is given by*

$$E[\# \text{ of visits to } i | X_0 = i] = \sum_{n=1}^{\infty} P_{ii}^n$$

Proof. Given that the system is initially at state i, the probability that the system visit state i exactly n (throughout its infinite life) is $f_{ii}^{n-1}(1 - f_{ii})$, which is the product of the probability visiting state i $n-1$ times and then never visit again.

Therefore, the expected number of visits to state i starting from state i is given by

$$\sum_{n=1}^{\infty} n f_{ii}^{n-1}(1 - f_{ii}) = \frac{1}{1 - f_{ii}}$$

where we use the result from geometric distribution.

$$\begin{aligned}
E[\# \text{ of visits to } i | X_0 = i] &= E[\sum_{n=0}^{\infty} I_n | X_0 = i] \\
&= \sum_{n=0}^{\infty} P(X_n = i | X_0 = i) \\
&= \sum_{n=0}^{\infty} P_{ii}^n
\end{aligned}$$

In fact, we can calculate the expected number of visits to state j starting from state i is given by

$$\begin{aligned}
E[\# \text{ of visits to } j | X_0 = i] &= E[\sum_{n=0}^{\infty} I_n | X_0 = i] \\
&= \sum_{n=1}^{\infty} P_{ii}^n
\end{aligned}$$

□

We have alternative characterization on transient and recurrent states

Theorem 20.2.2 (characterization of transient and recurrent states by expected number of visits).

- *A state i is transient if and only if $\sum_{n=0}^{\infty} P_{ii}^n < \infty$.*
- *A state i is recurrent if and only if $\sum_{n=0}^{\infty} P_{ii}^n = \infty$.*
- *If states i and j are in the same transient class, then $\sum_{n=0}^{\infty} P_{ij}^n < \infty$; that is, the expected number of visits to j starting from $X_0 = i$ is finite.*

Proof. (1)(2) According to Lemma 20.2.1, the expected number of visits to state i starting from state i is given by

$$E[\text{\# of visits to } i | X_0 = i] = \frac{1}{1 - f_{ii}} = \sum_{n=1}^{\infty} P_{ii}^n.$$

Since a transient state is defined by $f_{ii} < 1$ and a recurrent state is defined by $f_{ii} = 1$, the results can be showed. (3) Suppose i can access j, after the visit, the expected number of visits to j from state i is the expected number of visits to j from state i. Suppose i cannot access j, then the expected number of visits to j from state i is zero. Considering both cases,

we have $\sum_{n=0}^{\infty} P_{ij}^n \leq \sum_{n=0}^{\infty} P_{jj}^n - 1 < \infty, \forall i \neq j$

□

If $\sum_{n=0}^{\infty} P_{ii}^n < \infty$ and $i \to j$ then

$$\sum_{n=0}^{\infty} P_{ij}^n < \infty$$

Suppose

$$\sum_{n=0}^{\infty} P_{ij}^n < \infty$$

$$\begin{aligned}
\sum_{n=0}^{\infty} P_{ij}^n &= \sum_{n=0}^{k-1} P_{ij}^n + \sum_{n=k}^{\infty} P_{ij}^n \\
&= \sum_{n=0}^{k-1} P_{ij}^n + \sum_{n=k}^{\infty} P_{ij}^n \\
&= \sum_{n=0}^{k-1} P_{ij}^n + \sum_{n=0}^{\infty} P_{ii}^n P_{ij}^k < \infty
\end{aligned}$$

If i can not access j, and $\sum_{n=0}^{\infty} P_{ij}^n = 0$.

Theorem 20.2.3. *Regarding a transient state, the following are equivalent:*

- *A state i is transient*
- *Let* $T_i = \min\{n \geq 1 | X_n = i\}$. *Then* $P(T_i < \infty | X_0 = i) < 1$
- $\sum_{n=0}^{\infty} P_{ii}^n < \infty$
- $P(X_n = i for some n | X_0 = i) < 1$
- $P(X_n = i for infinitely many n | X_0 = i) = 0$

Regarding a recurrent state, the following are equivalent:

- *A state i is recurrent*
- *Let* $T_i = \min\{n \geq 1 | X_n = i\}$. *Then* $P(T_i < \infty | X_0 = i) = 1$
- $\sum_{n=0}^{\infty} P_{ii}^n = \infty$
- $P(X_n = i for some finite n | X_0 = i) = 1$
- $P(X_n = i for infinitely many n | X_0 = i) = 1$

Example 20.2.3 (One-dimensional random walk). Consider a random walk on $\mathbb{Z}$ with left move probability p and right move probability $1 - p$. The recurrent probability of the system at step $2n - 1$ and $2n$ to position 0 is

$$P_{00}^{2n-1} = 0$$

and

$$P_{00}^{2n} = \binom{2n}{n} p^n (1-p)^n$$

Using Stirling's formula [Lemma A.2.1], we have

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

and

$$\begin{aligned}\binom{2n}{n} &= \frac{(2n)!}{(n!)^2} \\ &\sim \frac{(2n)^{2n} e^{-2n} \sqrt{2\pi 2n}}{n^{2n} e^{-2n} 2\pi n} \\ &= \frac{2^{2n}}{\sqrt{n\pi}}.\end{aligned}$$

We now have the simplified formula

$$\begin{aligned} P_{00}^{2n} &\sim \frac{2^{2n}}{\sqrt{n\pi}} p^n (1-p)^n \\ &\sim \frac{1}{\sqrt{n\pi}} (4p(1-p))^n \end{aligned}$$

If $p = 0.5$ (i.e., symmetric random walk), then

$$P_{00}^{2n} \sim \frac{1}{n\pi}, \sum_{n=0}^{\infty} P_{00}^{2n} = \infty.$$

Therefore, any state is recurrent when $p = 1/2$.

When $p \neq 1/2$ such that $4p(1-p) < 1$, then we have a convergent series, i.e.,

$$\sum_{n=0}^{\infty} P_{00}^{2n} < \infty.$$

Therefore, any state is transient when $p = 1/2$.

20.2.2.2 *From states to classes*

Theorem 20.2.4 (recurrence is class property). *If state i can access j and state i is recurrent, then j can access i and j is recurrent.*

Further more, in a Markov chain with a single communicating class, either all states are recurrent or all states are transient.

Proof. Because $i \to j$, then there exists a positive integer such that $P_{ij}^n > 0$. Because of the positive probability of i visits j, i will eventually visit j. Suppose after i visits j but never come back to i, this will contradict with the assumption that i is recurrent. Because i will return to itself infinitely many times, since the state starting from i will also visit j and since the system will revisit i infinitely many times, the system will revisit j infinitely many times. Therefore j is recurrent.

For other states in the same communicating class, system starts from i can visit them all and all of them will be recurrent.

If i is transient but the any of the rest is recurrent, i will be recurrent instead of transient. Therefore, all the rest of the states must be transient. $\square$

Lemma 20.2.2 (At least one recurrent class in the state space). *[1, p. 237] Every finite state Markov chain must have at least one recurrent communicating class.*

Proof. Suppose there are no recurrent class, then there are only transient classes. By definition, every transient class has at least one path pointing outside. Because every class is transient, which means eventually there is zero probability in it. However, total probability has to sum to 1, therefore contradiction. □

Lemma 20.2.3 (Trajectory must end in a recurrent/closed/adsorbing class). *[1, p. 237] The system state of a finite Markov chain will eventually with probability 1 enter some closed communicating class.*

Proof. As the step $k \to \infty$, the probability of system state lying in transient state will be 0 (by definition). Therefore, the probability of system state can only accumulate at some closed class. □

Definition 20.2.3 (irreducible). *If all states in a Markov chain belong to a communicating recurrent class, then the Markov chain is irreducible.*

Corollary 20.2.4.1. *Every finite state irreducible Markov chain has and only has one recurrent class.*

20.2.2.3 *Qualitative classification of recurrent and transient classes*

Note 20.2.1 (Qualitative classification of communicating classes: recurrent class and transient class). [1, p. 237] Let $\tilde{S} \subset S$ such that

- All states in $\tilde{S}$ communicate.
- If $i \in \tilde{S}, j \notin \tilde{S}$, then $p_{ij}^k = 0, \forall k$; that is, states in $\tilde(S)$ will never visit states outside $\tilde(S)$.

Then we say $\tilde{S}$ form a **recurrent** class of states.

A communicating class C is transient if some state outside of C is accessible from C, or equivalently, there must exist at least one path from one of its members to some state outside the class.

Remark 20.2.2 (interpretation).

- A communicating class can contain only one state.
- There are **only two types of communicating classes: recurrent class and transient class**. When system enters into an recurrent class, it will never get out.When system enters into a transient class, it will eventually get out.
- states in a recurrent class have zero probability to access states outside.
- states outside might access state inside, and once entering into this class, it has no chance of leaving; therefore it is also called adsorbing class.
- Any state of a finite Markov chain is with probability 1 to eventually enter some recurrent communicating class.

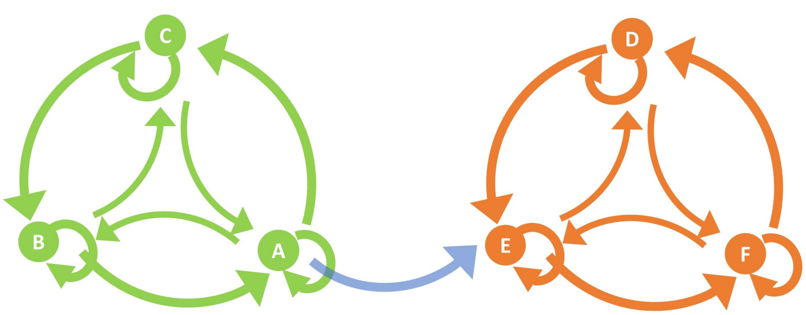


Figure 20.2.3: Classification of communicating classes into recurrent class and state space by communicating classes. Green states form a communicating class belonging to the transient class. Orange states form a communicating class belonging to the recurrent/closed/adsorbing class.

20.2.3 Periodicity

Definition 20.2.4 (periodicity of a Markov chain). *Given a Markov chain, define the set*

$$N(i) = \{n \geq 1 | P_{ii}^n > 0\}.$$

The period of a state i, denoted by $d(i)$, is the greatest common divisor of values in $N(i)$. More precisely, we have

$$d(i) \triangleq \begin{cases} 0 & \text{if } N(i) = \emptyset \\ \gcd(N(i)) & \text{otherwise} \end{cases}$$

*. If the $d(i) = 1$, the state is **aperiodic**; otherwise, the state is periodic.*

Remark 20.2.3 (why aperiodicity is necessary). If a Markov has periodicity of 2, then $P_{ii}^{2k+1} = 0$ for some states, therefore, no matter how large is k, P will still have zero entries.

Methodology 20.2.1 (criteria for an aperiodic chain).

- *Note that if $\{n \geq 1 : P_{ii}^n > 0\}$ contains two distinct numbers that are relatively prime to each other, then the state i is aperiodic.*
- *If $P_{ii} > 0$, then the state is aperiodic. If the chain is irreducible, then the chain is irreducible.*
- *If there exists a positive integer n such that all elements in matrix P^n are strictly positive, then the chain is aperiodic.*[a]

[a] note that for all $k > n$, P^k will be strictly positive. Then we can find two coprime numbers $k_1, k_2 > n$ such that P^{k_1}, P^{k_2} that are positive.

Example 20.2.4 (periodic and aperiodic chains).

- The chain [Figure 20.2.4(a)] with
$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, P^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, P^3 = P$$
is periodic with periodicity of 2.
- The chain [Figure 20.2.4(a)] with
$$P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}, P^2 = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, P^3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
is periodic with periodicity of 2.
- The chain [Figure 20.2.4(a)] with
$$P = \begin{bmatrix} 0 & 0.5 & 0.5 \\ 0.5 & 0.5 & 0 \\ 0.5 & 0 & 0.5 \end{bmatrix}, P^2 = \begin{bmatrix} 0.25 & 0.5 & 0.5 \\ 0.25 & 0.5 & 0.25 \\ 0.25 & 0.5 & 0.25 \end{bmatrix}, P^3 = \begin{bmatrix} 0.25 & 0.375 & 0.375 \\ 0.375 & 0.375 & 0.25 \\ 0.375 & 0.25 & 0.375 \end{bmatrix}$$
is aperiodic since P^3 is all positive.

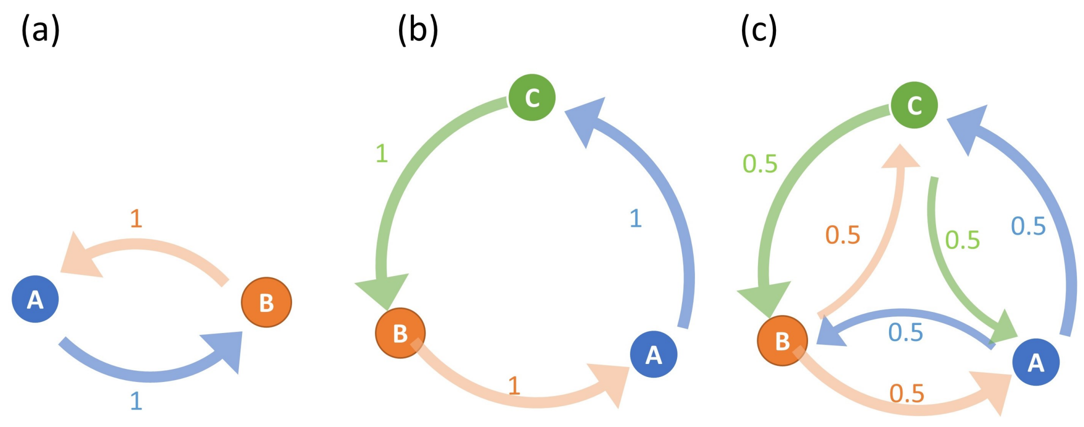


Figure 20.2.4: Example periodic and aperiodic Markov chains.

Theorem 20.2.5. *[2, p. 166] For any Markov chain(with either finite or countable infinite states), all states in the same communicating class have the same period. That is, if states i and j communicate, then $d(i) = d(j)$.*

Proof. If states i and j can communicate, then $d(i) = d(j)$. That is, every communicating class have the same periodicity

We first introduce a inequality

$$P_{ij}^{m+n} \geq P_{ik}^{m} P_{kj}^{n}, \forall m, n \in \mathbb{Z}_{+}.$$

Suppose first that $d(i) = 0$. Since $i \to j, j \to i$, $P_{ij}^n > 0$ and $P_{ji}^m > 0$ for some $n, m \geq 0$. This implies that $P_{ii}^{m+n} > 0$, which forces $m + n = 0$ or $m = n = 0$, and hence $j = i$.

Next assume $d(i) > 0$ and $N(i) \neq \emptyset$. Since $i \to j, j \to i$, there are $r, s \geq 0$ such that $P_{ji}^r > 0, P_{ji}^s > 0 \implies P_{ji}^{r+s} > 0, r + s \in N(j)$. If we choose any $n \in N(i)$, we have $P_{jj}^{r+n+s} \geq P_{ji}^r P_{ii}^n P_{ij}^s > 0$ and thus $r + s + n \in N(j)$.

Since $d(j)|r + s, d(j)|r + s + n$, we have $d(j)|n$. Because n is an arbitrary element in $N(i)$, then $d(j)$ is the common divisor for all values in $N(j)$ and $d(j) \leq d(i)$.

Conversely, we can show $d(i) \leq d(j)$. Therefore $d(i) = d(j)$. □

20.2.4 Positive and null recurrent

In Markov chain with countable infinitely many state, we further distinguish recurrent states with positive and null recurrence. Both types of states can revisit itself infinitely many times but takes either finite or infinite steps on average.

Definition 20.2.5 (first visit time, recurrent time). *Let T_{ij} denote the stochastic first arrival time starting from state i to state j, which can be more formally defined by*

$$T_{ij} = \min_n \{n \geq 1 | X_n = j, X_0 = i\}.$$

Then T_{ii} is the recurrent time for state i.

Definition 20.2.6 (positive recurrent, null recurrent). *[3, p. 125] A **recurrent** state $i \in S$ is said to be **positive recurrent** if the mean recurrent time is finite*

$$\mu_i = E[T_{ii}] < \infty,$$

*and **null recurrent** if*

$$\mu_i = E[T_{ii}] = \infty.$$

Theorem 20.2.6 (positive/null recurrence are class properties).

- *If recurrent state i is positive recurrent and states i and j communicate to each other, then state j is also positive recurrent.*
- *If recurrent state i is null recurrent and states i and j communicate to each other, then state j is also null recurrent.*

Proof. Because the states are recurrent states, systems starts from i will almost surely revisit i. We can partition all trajectories starting from i and then return to i into two sets. Let A denote the event that system starting from j visit i before revisit j and $\overline{A}$ denote the event that system tarting from j will not visit i before revisit j (i.e., directly revisit)

Then $E[T_{jj}] = P(A)(E[T_{ji}] + E[T_{ij}]) + P(\overline{A})E_{T_{jj}}$

Let B denote the event the system starting from i visit j before revisit i.

Then

$$E[T_{ii}] = P(B)(E[T_{ji}] + E[T_{ij}]) + P(\overline{B})E_{T_{ii}} \implies E[T_{ii}] = \frac{P(B)}{1 - P(\overline{B})}(E[T_{ji}] + E[T_{ij}]).$$

Clearly, If $E[T_{ii}]$ is finite, then $E[T_{jj}], E[T_{ij}], E[T_{ji}]$ must all be finite.

□

Lemma 20.2.4 (finite state space chain are positive recurrent). *[3, p. 125]*

- *An irreducible Markov chain with finite state space have all its states being positive recurrent.*
- *If a Markov chain has a finite state space, then all of its recurrent states are positive recurrent.*

Proof. (1) See Lemma 20.4.3. (2) Note that all these recurrent states will not classified into different recurrent classes. Since there is no path out inside a recurrent class, we can isolate each recurrent class and apply (1) to it. □

20.2.5 Summary

For historical reasons, there are also other naming conventions besides what we have discussed above. For example, a recurrent state $i \in S$ is said to be **ergodic** if it is both positive recurrent and aperiodic. Note the concepts of positive recurrence and aperiodicity also apply to class. A Markov chain consisting entirely of one ergodic class is called an**ergodic chain.**.

We have following summary.

Table 20.2.1: Summary of Markov chain state property[3, p. 140]

Property	**Definition**
absorbing (state)	$P_{i,i} = 1$
recurrent (state)	$P(T_i^r < \infty \vert X_0 = i) = 1$
transient (state)	$P(T_i^r < \infty \vert X_0 = i) < 1$
positive recurrent (state)	recurrent and $E[T_i^r \vert X_0 = i] < \infty$
null recurrent (state)	recurrent and $E[T_i^r \vert X_0 = i] = \infty$
aperiodic (state or chain)	period = 1
ergodic (state or chain)	positive recurrent and aperiodic
irreducible (chain)	all states communicate
regular	all coefficients of $P^n > 0$ for some $n \geq 1$

20.3 Absorption analysis

20.3.1 Matrix structure for adsorption analysis

One of most important applications of Markov chain is adsorption analysis, which is about the system's transition dynamics from transient states to recurrent states (as if being adsorbed into recurrent states). We will first reorder transition matrix P according to different class classifications. It turns out that matrix power of this re-ordered matrix reveal interesting and in-depth connection to adsorption analysis.

By ordering r recurrent class states and followed by $n-r$ transient class states, we can write the transition matrix in the partitioned form as[1, p. 239]

$$P = \begin{pmatrix} P_1 & 0 \\ R & Q \end{pmatrix}$$

where $P_1 \in \mathbb{R}^{r\times r}$ is the stochastic matrix between r states within recurrent class, $R \in \mathbb{R}^{(n-r)\times r}$ is the stochastic matrix for $n-r$ transient states to r states in recurrent class, $Q \in \mathbb{R}^{(n-r)\times r}$ is the stochastic matrix among $n-r$ states within recurrent class. This matrix form is known as **canonical form of Markov chain**.

Note that

$$P = \begin{pmatrix} P_1 & 0 \\ R & Q \end{pmatrix}, P^2 = \begin{pmatrix} P_1^2 & 0 \\ RP_1 + QR & Q^2 \end{pmatrix}, P^2 = \begin{pmatrix} P_1^3 & 0 \\ RP_1^2 + QRP_1 + Q^2R & Q^3 \end{pmatrix}.$$

Remark 20.3.1 (relationship to the power of P)**.**

- Note that P^n is still a lower triangle block matrix.
- In P^n the lower right block is given by Q^n

We can investigate a number of interesting quantities, such as absorbing time using the canonical form.

Theorem 20.3.1 (absorption analysis via matrix structure). *[1, p. 240]*

- *Define $M = (I-Q)^{-1}$. $M \in \mathbb{R}^{r\times r}$, which is known as the* ***fundamental matrix****. Then the matrix $M = I + Q + Q^2 + Q^3 + \cdots$ exists and is positive. Note that the element m_{ij} in M is equal to the expected number of visits to j if the system is initiated in transient state i.*

- *The summation of the ith row of M is equal to the mean number of steps before entering a recurrent class when the system is initiated in transient state i.*
- *Let $B = MR, B \in \mathbb{R}^{(n-r)\times r}$, the B_{ij} is the probability that if a Markov chain starts with transient state i, it will first enter a recurrent class via j in the recurrent class.*

Proof. (1) From the definition, we know that

$$M_{ij} = I_{ij} + Q_{ij} + Q_{ij}^2 + ... + Q_{ij}^k + ...$$

where we interpret Q_{ij}^k as the probability that the system is at state j at step k when initially is at $X_0 = i$. Therefore, the summation will equal to the expected visits to state j when the state is initially at state i [Lemma 20.2.1]. For transient states, the expected number of visits are finite [Theorem 20.2.2], therefore $(I + Q + \cdots)$ exists.

Since $(I - Q)(I + Q + Q^2 + \cdots) = I - Q^n, n \to \infty$, we have

$$(I - Q)(I + Q + Q^2 + \cdots) = I.$$

(2) The summation of the i row in M is the total expected number of visits to transient states.

(3)$B = R + QB$ A state i in transient class has two possibilities to enter a recurrent class via j: (1) it directly enter in one step via R_{ij}; (2) it first enter other transient states before its final enter into j, where $QR + Q^2R + Q^3R + ...$ gives the probability of indirect enter using 2 step, 3 steps, 4 steps.... □

Example 20.3.1. Consider a Markov chain with transition matrix given by

$$P = \begin{bmatrix} 1 & 0 & 0 \\ 0.5 & 0.25 & 0.25 \\ 0.5 & 0.25 & 0.25 \end{bmatrix},$$

where it is clear that $Q = \begin{bmatrix} 0.25 & 0.25 \\ 0.25 & 0.25 \end{bmatrix}$, therefore we have $(I - Q)^{-1} = \begin{bmatrix} 1.5 & 0.5 \\ 0.5 & 1.5 \end{bmatrix}$.

To verify, we have $Q = \begin{bmatrix} 0.25 & 0.25 \\ 0.25 & 0.25 \end{bmatrix}, Q^2 = \begin{bmatrix} 0.5^3 & 0.5^3 \\ 0.5^3 & 0.5^3 \end{bmatrix}, Q^n = \begin{bmatrix} 0.5^{n-1} & 0.5^{n-1} \\ 0.5^{n-1} & 0.5^{n-1} \end{bmatrix}$ It can be verified that

$$(I - Q)^{-1} = I + Q + Q^2 + \cdots$$

20.3.2 Absorbing Markov chains

Definition 20.3.1 (absorbing Markov chain). *A Markov chain is absorbing if it have at least one **absorbing state** (i.e., $P_{ii} = 1$ if state i is adsorbing state) and if from every state it is possible to go to an absorbing state. An absorbing state i is characterized by $p_{ii} = 1, p_{ij} = 0, j \neq i$.*

Lemma 20.3.1 (matrix structure). *By ordering absorbing states and followed by transient states, we can write the transition matrix in the partitioned form as*

$$P = \begin{pmatrix} I_r & 0 \\ R & Q \end{pmatrix}$$

where $I_r \in \mathbb{R}^{r\times r}$ is the stochastic matrix between r absorbing states, $R \in \mathbb{R}^{(n-r)\times r}$ is the stochastic matrix for $n-r$ transient states to r states in closed class, $Q \in \mathbb{R}^{(n-r)\times r}$ is the stochastic matrix between $n-r$ transient states and absorbing states.

Lemma 20.3.2 (adsorption probability analysis). *[3, p. 96] Consider the M state Markov chain. The probability to reach a specific adsorbing state i is given by $a_1, ..., a_M$ such that*

- *$a_i = 1$ for i being an absorbing state.*
- *$a_i = \sum_{j=1}^{m} a_j p_{ij}$ for i being a transient state.*

- *If we partition the a vector as* $a = [a_{ab}\ a_{tran}]$, *then we have*

$$\begin{bmatrix} a_{ab} \\ a_{tran} \end{bmatrix} = \begin{bmatrix} I_r & 0 \\ R & Q \end{bmatrix} \begin{bmatrix} a_{ab} \\ a_{tran} \end{bmatrix}$$

and

$$a_{tran} = (I-Q)^{-1} R a_{ab} = ([(I-Q)^{-1}R]_i)^T$$

Proof. (1)(2) We can use Chapman-Kolmogov Equation [Lemma 20.1.1] such that

$$P(X_n = i | X_0 = j) = \sum_{k=1}^{m} P(X_n = i | X_1 = j) P(X_1 = j | X_0 = i).$$

Let A be the event of starting from j and arriving $i, i \neq j$, then

$$P(A|X_0 = j) = \sum_{n=1}^{\infty} P(X_n = i | X_0 = j) = \sum_{n=1}^{\infty} P(X_n = i | X_1 = j)$$

since $P(X_1 = i | X_1 = j) = \delta_{ij}$. Then we have

$$P(A|X_0 = j) = a_j = \sum_j a_j p_{ij} = \sum_k P(A|X_0 = a_j) p_{ij}$$

(3)**This is a special case of Theorem 20.3.1.** □

Corollary 20.3.1.1 (expected time to adsorption). *[3, p. 99] Consider the* M *state Markov chain. The expected time to reach a specific adsorbing state* s *is given by* $u_1, ..., u_M$ *such that*

- $u_i = 0$ *for* i *being an absorbing state.*
- $u_i = 1 + \sum_j^m u_j p_{ij}$ *for* i *being a transient state.*
- *If we partition the* u *vector as* $u = [u_{ab}\ u_{tran}]$ *with* $u_{ab} = 0$, *then we have*

$$\begin{bmatrix} u_{ab} \\ u_{tran} \end{bmatrix} = \begin{bmatrix} I_r & 0 \\ R & Q \end{bmatrix} \begin{bmatrix} 0 \\ u_{tran} \end{bmatrix} + \begin{bmatrix} 0 \\ \mathbf{1} \end{bmatrix}$$

and

$$u_{tran} = (I-Q)^{-1}\mathbf{1}$$

i.e. the row sum of $(I-Q)$

Remark 20.3.2. This is a special case of Theorem 20.3.1.

20.3.3 Hitting and return analysis

Now we address more general hitting and return analysis.

Definition 20.3.2 (first hitting time and first return time).

- *First hitting time T_{ij} from state i to j is defined as*
$$T_{ij} = \min\{n \geq 0 : X_n = j, X_0 = i\},$$
and the mean first hitting time starting from i to j is denoted as
$$t_{ij} = E[T_{ij}].$$
- *First return time of state i is defined as*
$$R_i = \min\{n \geq 1 : X_n = i, X_0 = i\},$$
and the mean first return time of state i is defined as
$$r_i = E[R_i].$$
- *Hitting probability $h_{i,j}$ from state i to j is the probability of* ***ever*** *reaching state j, starting from initial state i. $h_{i,j}$ is defined as*
$$h_{i,j} = P(X_n = j \text{ for some } n \geq 0 | X_0 = i).$$
- *Returning probability h_i^R of state i is the probability of* ***ever*** *returning to state i, starting from initial state i. h_i^R is defined as*
$$h_i^R = P(X_n = j \text{ for some }, n \geq 1 | X_0 = i).$$

Lemma 20.3.3 (governing equation for mean first hitting time and first return time). *Let i be the starting state and the set A be the finishing states ($i \notin A$). Then*

- $T_{jj} = t_{jj} = 0, \forall j \in A.$
- $t_{ij} = 1 + \sum_{k \in S} p_{ik} t_{kj}, \forall j \in A.$
- $r_i = 1 + \sum_{k \in S} p_{ik} t_{ki}$

Proof. Using law of total expectation. □

Lemma 20.3.4 (governing equation for hitting probability). *[3, p. 104] Let i be the starting state and the set A be the finishing states ($i \notin A$). Then*

- $h_{jj} = 1, \forall j \in A$.
- $h_{ij} = \sum_{k \in S} p_{ik} h_{kj}, \forall j \in A$.
- $h_i^R = \sum_{k \in S} p_{ik} h_{ki}$

Proof. Using law of total probability. □

Lemma 20.3.5 (hitting and return probabilities are one in an irredicible finite chain). *The hitting and returnning probabilities in an irreducible finite chain is 1*

Proof. the governing equations hold if we plug 1 into it. □

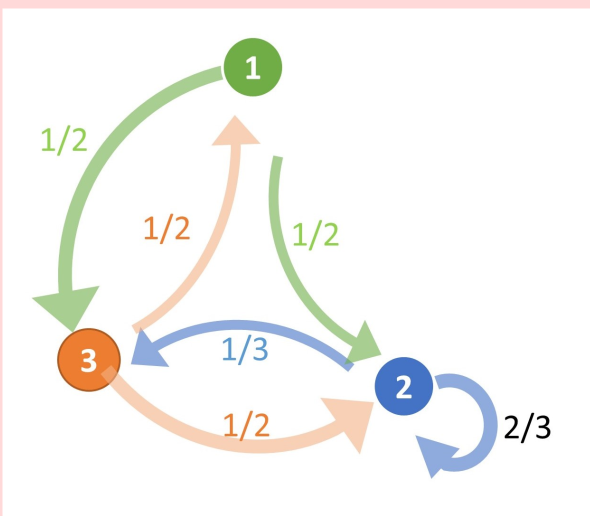
Remark 20.3.3 (mean first hitting time vs. mean first return time). First hitting time and first return time has quite subtle difference. Indeed, their definitions are given by

$$T_{ii} = \min_n \{n \geq 0, X_n = i | X_0 = i\}$$

$$R_i = \min_n \{n \geq 1, X_n = i | X_0 = i\}$$

More illustrations can be found in the following example.

Example 20.3.2. Consider following Markov chain.



Let the target state be 1, we have

$$\begin{aligned} t_{11} &= 0 \\ t_{21} &= 1 + \frac{2}{3}t_{21} + \frac{1}{3}t_{31} \\ t_{31} &= 1 + \frac{1}{2}t_{11} + \frac{1}{2}t_{21} \end{aligned}$$

which has results of $t_{11} = 0, t_{21} = 8, t_{31} = 5$.

On the other hand, the mean return time to state 1 is

$$r_1 = 1 + \frac{1}{2}t_{21} + \frac{1}{2}t_{31}.$$

20.3.4 Examples

20.3.4.1 *Consecutive coin toss game*

Example 20.3.3 (consecutive coin toss). A fair coin is tossed repeatedly until 5 consecutive heads occurs. What is the expected number of coin tosses?
Solution: Let e be the expected number of tosses. We have the following situations:

- If we get a tail immediately, the expected number is $e + 1$.
- If we get a head and then a tail, the expected number is $e + 2$.
- If we get two heads and then a tail, the expected number is $e + 3$
- If we get 3 heads and then a tail, the expected number is $e + 4$
- If we get 4 heads and then a tail, the expected number is $e + 4$
- If we get 5, the expected number is 5

Then, we have

$$e = \frac{1}{2}(e+1) + \frac{1}{2^2}(e+2) + \cdots + \frac{1}{2^5}(e+5) + \frac{1}{2^5}5,$$

which gives $e = 62$.

Remark 20.3.4.

- The different scenarios cannot contain same 'substring'
- This can generalize the biased coin straight forward.

Lemma 20.3.6. *Let p be the probability of flipping a head. Let x be the number of flips needed to to achieve h consecutive heads. Then*

$$E[x] = \frac{1-p^h}{p^h(1-p)}.$$

Remark 20.3.5. For excellent examples, see [4, p. 110].

20.4 Limiting behavior & distributions

20.4.1 Preliminary: eigenvalue proproties of stochastic matrices

20.4.1.1 *Preliminary: Frobenius-Perron matrix theory*

Definition 20.4.1. *If $A = [a_{ij}]$ is matrix, we write:*

- *$A > 0$ if all $a_{ij} > 0$, A is called strictly positive.*
- *$A \geq 0$ if all $a_{ij} \geq 0$ and for at least one $a_{ij} > 0$, A is called positive or strictly non-negative.*
- *$A \geqq 0$ if all $a_{ij} \geq 0$, A is called non-negative.*

Theorem 20.4.1 (Frobenius-Perron). *[1, p. 191]If $A > 0$, then there exists $\lambda_0 > 0$ and $x_0 > 0$ such that*

- *$Ax_0 = \lambda_0 x_0$, λ_0 is known as the Frobenius-Perron eigenvalue.*
- *if $\lambda_i \neq \lambda_0$ is any eigenvalue of A, then $|\lambda| < \lambda_0$;*
- *λ_0 is an eigenvalue with geometry and algebra multiplicity 1.*

Caution!
Non-negative matrices are not necessarily positive semidefinite; that is, they might have negative eigenvalues. For example,

$$A = \begin{pmatrix} 1 & 100 \\ 100 & 1 \end{pmatrix}$$

then $x^T A x < 0, x = [1, -100]^T$.

Theorem 20.4.2 (extension to strictly non-negative). *[1, p. 193]Let $A \geq 0$, and suppose there is an integer m such that $A^m > 0$. Then there exists $\lambda_0 \geq 0$ and $x_0 \geq 0$ such that*

- *$Ax_0 = \lambda_0 x_0$;*
- *if $\lambda_i \neq \lambda_0$ is any eigenvalue of A, then $|\lambda| < \lambda_0$;*
- *λ_0 is an eigenvalue with geometry and algebra multiplicity 1.*

Proof. Note that the eigenvalues of A^m is the m power of the eigenvalues of A. Therefore the comparison order among eigenvalues preserved after taking powers. The Frobenius-

Perron eigenvalue and eigenvectors, whose geometry and algebra multiplicity, will also be 1 after taking powers. □

Lemma 20.4.1 (bounds on Frobenius-Perron eigenvalue). *Let $A \geq 0$, and suppose there is an integer m such that $A^m > 0$. Let λ_0 denote the Frobenius-Perron eigenvalue. We have*

$$\min_i \Delta_i \leq \lambda_0 \leq \max_i \Delta_i$$

and

$$\min_i \delta_i \leq \lambda_0 \leq \max_i \delta_i$$

where Δ_i is the sum for column i, δ_i is the sum of the row i.

Proof. Let x be the normalized eigenvector corresponds to λ_0, then $Ax = \lambda_0 x$ Sum up rows, we have

$$\sum_{i=1}^{n} \Delta_1 x_1 = \lambda_0$$

i.e., λ_0 is the convex combination of Δ_is. Similarly, we have $A^T x' = \lambda_0 x'$, where x' is the eigenvector for A^T. □

Theorem 20.4.3 (Frobenius-Perron eigenvalue and eigenvectors in a finite-state, irreducible, and aperiodic Markov chain).

- *The stochastic matrix P of an irreducible and aperiodic Markov chain has Frobenius-Perron eigenvalue 1 and constant right eigenvector.*
- *The doubly stochastic stochastic matrix P[a] has Frobenius-Perron eigenvalue 1, and constant left and right eigenvectors.*

[a] Both rows and columns sum to 1

Proof. First, for a finite-state, irreducible, and aperiodic Markov chain, there will exist a positive integer k such that P^k is strictly positive (see the following sections).

Using the bounds of Frobienius-Perron eigenvalue lemma [Lemma 20.4.1], the Frobienius-Perron eigenvalue can only take value 1.

For a stochastic matrix whose rows sum to 1, the right eigenvector associated with eigenvalue 1 is a constant vector, i.e., $P\mathbf{1} = \mathbf{1}$.

For a doubly stochastic matrix whose rows and columns sum to 1, the left and right eigenvectors associated with eigenvalue 1 is a constant vector, i.e., $P\mathbf{1} = \mathbf{1}$ and $\mathbf{1}^T P = \mathbf{1}^T$.

□

20.4.1.2 *More general situations*

Lemma 20.4.2 (eigenvalues of stochastic matrix). *The stochastic matrix P of a Markov chain always has eigenvalue 1. And all other eigenvalues $|\lambda| \leq 1$*

Proof. Since row sum is 1, then we always have right eigenvalue of 1 with eigenvector $\mathbf{1}$. Since left and right eigenvalues are the same. We always have eigenvalue of 1. Let η be a left eigenvector with eigenvalue λ. Then

$$\lambda \eta_i = \sum_j \eta_j p_{ij}.$$

So

$$|\lambda| \sum_i |\eta_i| \leq \sum_i \left| \sum_j \eta_j p_{ij} \right| \leq \sum_{ij} \left| \eta_j \right| p_{ij} = \sum_j \left| \eta_j \right|.$$

therefore,

$$|\lambda| \leq 1.$$

□

Remark 20.4.1 (possible eigenvalues of a stochastic matrix). In the following situation, we shows that we could have following situation for a general stochastic matrix.

- Eigenvalues of a Markov chain can be positive and negative, zero and even complex.
- The matrix might not be diagonalized.
- Absolute values of the eigenvalues must be bounded by 1.

Example 20.4.1 (eigenvalue can be 0). The stochastic matrix

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 1 \end{pmatrix}$$

have eigenvalue 1 and 0, with associated eigenvector $[0, 1]$ and $[\sqrt{2}/2, -\sqrt{2}/2]$. **Note that the eigenvector can contain positive and negative entries at the same time**. Usually, these types of eigenvector does not have meaningful interpretation in terms of probability flow, but only have mathematical interpretation.

Example 20.4.2 (eigenvalue can be negative). The stochastic matrix

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

have eigenvalue 1 and -1, with associated eigenvector $[\sqrt{2}/2, \sqrt{2}/2]$ and $[\sqrt{2}/2, -\sqrt{2}/2]$. **Note that the eigenvalues can contain negative values and it is difficult to interpret its meaning in terms of probability flow**.

Example 20.4.3 (eigenvalues can have multiple 1). The stochastic matrix

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

have eigenvalue 1 and 1.

Example 20.4.4 (eigenvalue might be complex number). The stochastic matrix

$$A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$$

have eigenvalue 1 and $-\frac{1}{2} \pm i\sqrt{3}/2$.

Example 20.4.5 (stochastic matrix might not be diagonalized). The stochastic matrix

$$A = \begin{pmatrix} 5/12 & 5/12 & 1/6 \\ 1/4 & 1/4 & 1/2 \\ 1/3 & 1/3 & 1/3 \end{pmatrix}$$

cannot be diagonalized.

20.4.2 Limiting theorem

20.4.2.1 *Limiting distribution*

In this section, we focus on the long-term behavior of Markov chains. In particular, we would like to know the fraction of times that the Markov chain spends in each state in the long run and the properties of $P^n, n \to \infty$.

For a finite-state chain with including both transient classes and recurrent classes, system starting from a transient class will enter one recurrent class and stay there forever. We will first limit our discussion to consider irreducible finite-state Markov chains, which only contain a recurrent class. Then we briefly discuss the case with multiple recurrent classes.

In the section, we also limit our discussion to aperiodic chains, more general chains with countable states and periodicity will be discussed in the next section.

Definition 20.4.2 (stationary distribution, invariant distribution). *Consider a discrete Markov chain characterized by transition matrix P. A stochastic vector π is called* ***stationary distribution****, or* ***invariant distribution*** *if*

$$\pi P = \pi$$

Remark 20.4.2 (existence and uniqueness of stationary distribution)**.** The stationary distribution might not exists(such as periodic chains); even if it exists, it might not be unique(for example, a chain that is not irreducible might have multiple stationary distribution.)

Example 20.4.6. Consider a Markov chain with $P = \begin{bmatrix} \frac{3}{5} & \frac{2}{5} \\ \frac{1}{5} & \frac{4}{5} \end{bmatrix}$. The distribution $\pi = (\frac{1}{3}, \frac{2}{3})$ is a stationary/invariant distribution since

$$(\frac{1}{3}, \frac{2}{3}) = (\frac{1}{3}, \frac{2}{3}) \begin{bmatrix} \frac{3}{5} & \frac{2}{5} \\ \frac{1}{5} & \frac{4}{5} \end{bmatrix}$$

Example 20.4.7. Consider a Markov chain with $P = \begin{bmatrix} \frac{3}{5} & \frac{2}{5} \\ \frac{1}{5} & \frac{4}{5} \end{bmatrix}$. The distribution $\pi = (\frac{1}{3}, \frac{2}{3})$ is a stationary/invariant distribution since

$$(\frac{1}{3}, \frac{2}{3}) = (\frac{1}{3}, \frac{2}{3}) \begin{bmatrix} \frac{3}{5} & \frac{2}{5} \\ \frac{1}{5} & \frac{4}{5} \end{bmatrix}$$

Example 20.4.8 (A Markov chain with multple adsorbing states). Consider a one-dimensional random walk defined on state space $S = \{1, 2, 3, 4, 5\}$ with two adsorbing boundaries, whose transition matrix is given by

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

And

$$P^{\infty} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ \frac{2}{3} & 0 & 0 & \frac{1}{3} \\ \frac{1}{3} & 0 & 0 & \frac{2}{3} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Theorem 20.4.4 (fundamental limit theorem for an irreducible, aperiodic, and finite-state Markov chain). *Let P be the transition matrix of an irreducible and aperiodic finite state Markov chain. Then:*

1. *There is an unique probability vector (which is the stationary distribution) $\pi > 0, \sum \pi = 1$ such that*
$$\pi^T P = \pi^T$$

2. $\pi_j = \lim_{n\to\infty} P(X_n = j|X_0 = i) = \lim_{n\to\infty} P^n_{ij}, \forall i$. *That is,*

$$\lim_{n\to\infty} P^n = \begin{bmatrix} \pi \\ \pi \\ \cdots \\ \pi \end{bmatrix} = \begin{bmatrix} \pi_1 & \pi_2 & \cdots & \pi_N \\ \pi_1 & \pi_2 & \cdots & \pi_N \\ \vdots & \vdots & \ddots & \vdots \\ \pi_1 & \pi_2 & \cdots & \pi_N \end{bmatrix}.$$

3.
$$\pi_i = \frac{1}{E[R_i]}$$

, where $E[R_i]$ is the mean return time to state i from state i [subsection 20.3.3].

Proof. (1) From Frobenius-Perron theorem that there exists an eigenvalue of 1 [Theorem 20.4.3], whose algebra and geometry multiplicity are both 1. Therefore, the eigenvector (after normalization) associated with eigenvalue 1, denoted by π, satisfies $\pi > 0, \pi P = \pi$.
(2)

Using Jordan decomposition [Theorem 4.10.4], we have

$$P = U\Lambda V = \begin{bmatrix} u_1 & U_J \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & J \end{bmatrix} \begin{bmatrix} v_1^T \\ V_J \end{bmatrix},$$

where $UV = I$, and J is a block matrix consisting other Jordan bloacks.

Note that $u_1 = \pi^T$ is the Frobenius-Perron left eigenvector since $u_1 = u_1 P$, and $v_1 = \mathbf{1}$ is the Frobenius-Perron right eigenvector since $Pv_1 = v_1$.

Because

$$P^n = U\Lambda^n V = U \begin{bmatrix} 1 & 0 \\ 0 & J^n \end{bmatrix} V.$$

Note that as $n \to \infty$, $J^n \to 0$, we have $P^n \to u_1 v_1^T = \pi^T \mathbf{1}^T$. (3) See proof in next section Theorem 20.4.5. □

Example 20.4.9. Consider a two-state chain with transition matrix

$$P = \begin{bmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{bmatrix}, \alpha, \beta \in (0,1)$$

We can verify that that

$$\begin{aligned} \mathbb{P}^n &= \begin{bmatrix} \frac{\beta}{\alpha+\beta} + \frac{\alpha}{\alpha+\beta}(1-\alpha-\beta)^n & \frac{\alpha}{\alpha+\beta} - \frac{\alpha}{\alpha+\beta}(1-\alpha-\beta)^n \\ \frac{\beta}{\alpha+\beta} + \frac{\beta}{\alpha+\beta}(1-\alpha-\beta)^n & \frac{\alpha}{\alpha+\beta} - \frac{\beta}{\alpha+\beta}(1-\alpha-\beta)^n \end{bmatrix} \\ &\to \begin{bmatrix} \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \\ \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \end{bmatrix} \quad \text{as } n \to \infty \end{aligned}$$

Corollary 20.4.4.1 (limit distribution for an irreducible, aperiodic, and finite-state Markov chain with doubly stochastic transition matrix). *Let P be the transition matrix of an irreducible and aperiodic finite state Markov chain. Further assume P is doubly stochastic. Let N be the number states in the chain. Then: $\pi_j = \lim_{n\to\infty} P(X_n = j|X_0 = i) = \lim_{n\to\infty} P_{ij}^n, \forall i$. That is,*

$$\lim_{n\to\infty} P^n = \begin{bmatrix} \pi \\ \pi \\ \cdots \\ \pi \end{bmatrix} = \begin{bmatrix} 1/N & 1/N & \cdots & 1/N \\ 1/N & 1/N & \cdots & 1/N \\ \vdots & \vdots & \ddots & \vdots \\ 1/N & 1/N & \cdots & 1/N \end{bmatrix}.$$

Proof. Using additional fact from Frobenius-Perron theorem on doubly stochastic matrix [Theorem 20.4.3]. □

Note 20.4.1 (Limiting distribution when there are multiple classes). We can represent P the canonical form

$$P = \begin{pmatrix} P_1 & 0 \\ R & Q \end{pmatrix}$$

where $P_1 \in \mathbb{R}^{r\times r}$ is the stochastic matrix between r states within recurrent class, $R \in \mathbb{R}^{(n-r)\times r}$ is the stochastic matrix for $n-r$ transient states to r states in recurrent class, $Q \in \mathbb{R}^{(n-r)\times r}$ is the stochastic matrix among $n-r$ states within closed class.

We have

Remark 20.4.3.

- system starting from a transient class will enter one recurrent class and stay there forever. Which recurrent class the system will first enter is addressed in subsection 20.3.1.
- The (left) eigenvectors corresponding to the eigenvalue 1 must have the form $p^T = [p_1^T, 0]$, where p_1^T is a r dimensional probability vector, representing that **only states in recurrent class can occur with positive probability in equilibrium**.
- The number of eigenvectors corresponding to the eigenvalue 1 equal the number of recurrent classes. The recurrent classes act like separate Markov chain and have equilibrium distributions. **Transient classes cannot sustain an equilibrium.**
- If there are c recurrent classes, then we will have c linearly independent equilibrium probability vector, each non-zero only over the elements of the corresponding recurrent class.

20.4.2.2 *Extensions via Long run return analysis*

In the previous section, we consider finite-state aperiodic chains. If we consider chains with periodicity, then the limit distribution will not exists. Consider following example.

Example 20.4.10 (a two-state periodic chain). Consider a two-state Markov chain with

$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

We have $P^{2n} = P, P^{2n+1} = I$; that is, $\lim_{n\to\infty} P^n$ does not exist.

Further, we would like to extend limiting distribution analysis to chains with countably many states. Our tool is average long run return time analysis. Let's start with definitions.

The π_j denotes the expected long-run proportion of time that the chain spends in state j.

$\pi_j = \lim_{n\to\infty} \frac{1}{n}\sum_{m=1}^{n} E[I(X_m = j)|X_0 = i] = \lim_{n\to\infty} \frac{1}{n}\sum_{m=1}^{n} P(X_m = j|X_0 = i) = \lim_{n\to\infty} \frac{1}{n}\sum_{m=1}^{n} P_{ij}^m$

Theorem 20.4.5 (fundamental limit theorem for an irreducible Markov chain). *Let P be the transition matrix of an irreducible Markov chain with all states being positive recurrent. Then:*

- *If all states are positive recurrent, then π exists and is given by*

$$\pi_j \triangleq \lim_{n\to\infty} \frac{1}{n}\sum_{m=1}^{n} P_{ij}^m = \frac{1}{E[R_j]} > 0.$$

  where $E[R_i]$ is the mean return time to state i from state i, and π satisfies $\pi = \pi P$; that is, π is the unique stationary distribution of the train.
- *If all states are null recurrent or transient, then $\pi_j = 0$. No stationary distribution exists.*

Proof. Assume $X_0 = j$. Let $t_0, t_1, ..., t_n$ be the time j is visited. Define $Y_i = t_i - t_{i-1}$, then $Y_1, .., Y_n$ are iid random variables with mean $E[Y_i] = E[R_j]$. Then

$$\begin{aligned} \frac{1}{\pi_j} &= \lim_{m\to\infty} \frac{m}{\sum_{k=1}^{m} I(X_k = j)} \\ &= \lim_{n\to\infty} \frac{\sum_{i=1}^{n} Y_i}{n} \\ &= E[R_j] \end{aligned}$$

□

Remark 20.4.4 (interpretation and extensions). [2, p. 170]

- If there are c recurrent classes, then we will have c linearly independent equilibrium probability vector, each non-zero only over the elements of the corresponding recurrent class.
- If the Markov chain has one or more periodic recurrent classes, then P^n does not converge.

Example 20.4.11 (one-dimensional random walk). Consider a symmetric random walk defined on $\mathbb{N}$. This Markov chain is aperiodic, irreducible and recurrent. Its limiting distribution is 0, and it does not have stationary distribution.

Lemma 20.4.3 (positive recurrence in finite-state irreducible chain). *All states in a finite-state irreducible Markov chain are positive recurrent.*

Proof. Let's assume all states are null recurrent.

Then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} P_{ij}^m = 0, \forall i, j \in S$$

Sum over j, we have

$$\sum_{j\in S} \lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} P_{ij}^m = 0, \forall i \in S$$

Because the state space is finite-sized, we can exchange the summation and limit and get

$$\lim_{n\to\infty} \frac{1}{n} \sum_{j\in S} \sum_{m=1}^{n} P_{ij}^m = 0, \forall i \in S$$

However, on the other hand if we sum j first

$$\lim_{n\to\infty} \frac{1}{n} \sum_{m=1}^{n} \sum_{j\in S} P_{ij}^m = 1, \forall i \in S$$

Therefore we have a contradiction and all states must be positive recurrent. □

20.4.3 Application: PageRank algorithm

In the PageRank algorithm, the user's browsing behavior is modeled as a Markov chain model with a finite state space defined by all the websites and their links to each other. The goal is find out the probability that a random user will browse each website. The website that has a higher probability will receive a high score and rank before those with lower score.

More formally, we consider a universe with N websites. The transition matrix is given by

$$P_{ij} = \begin{cases} 1/n_i, \text{ there is a outgoing link from website i to website j} \\ 0, otherwise \end{cases}$$

where n_i is the total number of outgoing links from website i.

If all websites are connected properly such that the resulting Markov chain is irreducible, finite state, and aperiodic, then there exists an unique stationary distribution. The value of the stationary distribution will simply be the rank(importance) of the web/node [Theorem 20.4.4].

In practice, P define above does not give an irreducible, finite state, and aperiodic chain. We can consider a modified transition matrix given by $M = (1-\lambda)P + \lambda J$ where $J_{ij} = \frac{1}{N}$. In Google's primitive algorithm, $\lambda = 0.15$. Clearly, M will give an irreducible, finite state, and aperiodic Markov chain. The interpretation of the modification in terms of user behavior is that a random surfer would click on a link on the current page with probability $(1-\lambda)$ and open up a random page with probability λ.

Remark 20.4.5 (computation of PageRank score and a basic web search design). We already knew the PageRank score is the top eigenvector associated with transition matrix M. Since all other eigenvalues are less than 1, we can use power iteration method [Theorem 4.17.2] to compute the page rank score v_0.

With the page rank score, we can design a web search in the following way: In the top K pages, find out the top L pages that are most similar to the query.

20.5 Detailed balance and spectral properties

In Remark 20.4.1, we see that a general stochastic matrix might not be diagonalizable. In this section, we consider a special class of Markov chains that satisfy the detailed balance condition. Chains satisfying this condition is diagonalizable and are commonly in chemical physics research. Spectral analysis on such chains can give in-depth view on the transition matrix.

Definition 20.5.1 (detailed balance). *If an irreducible, positive recurrent, and aperiodic Markov chain X with transition matrix P satisfies the detailed balance condition with respect to a stochastic vector π if*

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

Remark 20.5.1 (verification of stationary distribution). From detailed balance, we have

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

Sum up each side, we have

$$\begin{aligned}\sum_i \pi_i P_{ij} &= \sum_i \pi_j P_{ji} \\ \Rightarrow (\pi P)_j &= \pi_j\end{aligned}$$

Remark 20.5.2 (How to check detailed balance).

- For an irreducible, positive recurrent, and aperiodic chain, we know that there exists an unique stationary distribution π. Then we can verify whether π satisfy $\pi_i P_{ij} = \pi_j P_{ji}$.
- Usually it is difficult to check if there exists a detailed balanced distribution by observing the matrix structure.

Theorem 20.5.1 (spectral decomposition of Markov chain satisfying detailed balance). *Let P be the stochastic matrix associated with a Markov chain. Further assumes that P satisfies detailed balance. Then*

- *P can be symmetrized via $V = \Pi P \Pi^{-1}$ where Π is the diagonal matrix with entries $\sqrt{\pi_1}, ..., \sqrt{\pi_n}$ $(\pi = (\pi_1, ..., \pi_n)$ is the equilibrium probability vector).*
- *Let $\lambda_1, ..., \lambda_n$ and $w_1, ..., w_n$ be the* ***real eigenvalues*** *and unit eigenvectors of V. Then P has the* ***same real eigenvalues****. The left eigenvectors of P are given as*

$$\psi_j = \Pi w_j$$

The right eigenvectors of P are given as

$$\phi_j = \Pi^{-1} w_j$$

- *P has the following spectral decomposition:*

$$P = \sum_{k=1}^{n} \lambda_k \phi_k \psi_k^T = \sum_{k=1}^{n} \lambda_k \Pi^2 \phi_k \phi_k^T$$

Proof. (1)(2)(3) Since V is symmetric, it will have real eigenvalues. Moreover, we have $V = \Pi P \Pi^{-1}, V = W \Lambda W^{-1}$, then it can be showed that

$$(\Pi W)^T P = W^T \Pi \Pi^{-1} W \lambda W^T P i = \Lambda (\Pi W)^T$$

that is, ΠW are the left eigenvector matrix. Similarly, we can prove the rest. □

Remark 20.5.3 (How to check detailed balance).

- For an irreducible, positive recurrent, and aperiodic chain, we know that there exists an unique stationary distribution π. Then we can verify whether π satisfy $\pi_i P_{ij} = \pi_j P_{ji}$.
- Usually it is difficult to check if there exists a detailed balanced distribution from observation of matrix structure.

Corollary 20.5.1.1 (kinetics on a detailed balance Markov chain). *On a detailed balance Markov chain of M states,denoted by stochastic matrix P, we can decompose a given probability vector v as*

$$v = \sum_{i=1}^{M} a_i^{(0)} w_i$$

where w_i are the left eigenvalues of P. The coefficients a_i evolves as

$$a_i^{(n+1)} = \lambda_i a_i^{(n)}, i = 1, 2, ..., M$$

where λ_i are the eigenvalues of P.

Remark 20.5.4 (interpretation). The first k slowest non-trivial kinetics is characterized by the first k largest non-trivial eigenvalues.That is, the smaller the eigenvalue, the faster the kinetics.

20.6 Random walk

20.6.1 Basic concepts and properties

Definition 20.6.1 (random walk). *The stochastic process $\{B_n, n \in \mathbb{Z}_+\}$ is called a random walk $S_n = X_1 + X_2 + ... + X_n$ and X_is are iid discrete random variables taking σ and $-\sigma, \sigma > 0$ with probability p and $1-p$ respectively. If $p = 1/2$, then B_n is called symmetric random walk.*

Lemma 20.6.1 (basic properties). *Let B_n be a random walk with step size σ, then*

- $E[B_n] = 0$
- $Var[B_n] = n\sigma^2$.
- $cov(B_t, B_s) = \min(s,t)\sigma^2$

Proof. Note that $B_n = \sum_{i=1}^n X_n$ with $E[X_n] = 0$ and $Var[X_n] = \frac{1}{2}\sigma^2 + \frac{1}{2}\sigma^2$. Then use linearity of expectation, we can get (1)(2). For (3), let $s < t$, then

$$cov(B_s, B_t) = cov(B_s, B_s + \sum_{i=s+1}^{t} X_i) = Var[B_s] = s\sigma^2.$$

□

Lemma 20.6.2 (martingale property). *Let $\mathcal{F}_n$ be the filtrations associated with the random walk(i.e., $\mathcal{F}_n = \sigma(X_1, ..., X_n)$). Then we have*

- *B_n is a martingale.*
- *$B_n^2 - n$ is a martingale*

Proof. (1) $E[B_{n+1}|\mathcal{F}_n] = E[B_n + X_{n+1}|\mathcal{F}_n] = B_n$ (2) similar to (1). □

20.6.2 Persistent random walk

Definition 20.6.2 (persistent random walk). *The stochastic process $\{B_n, n \in \mathbb{Z}_+\}$ is called a random walk $B_n = X_1 + X_2 + ... + X_n$ and X_is are discrete random variables with properties*

$$E[X_i] = 0, Var[X_i] = \sigma^2, Cov(X_i, X_j) = \sigma^2\rho^{|i-j|}, |\rho| < 1.$$

X_i will take σ and $-\sigma, \sigma > 0$ with equal probability. ρ is called step size correlation coefficient and σ is called step size.

Lemma 20.6.3 (basic properties of persistent random walk). *Let B_n be a persistent random walk with correlation coefficient $\rho, |\rho| < 1$ and step size σ, then*

- $E[B_n] = 0$
- $$Var[B_n] = \sigma^2 (n + \frac{2\rho(n-1)}{1-\rho} - \frac{2\rho^2(1-\rho^{n-1})}{(1-\rho)^2})$$
  .
- $$cov(B_n, B_{n+h}) = Var[B_n] + \sum_{j=1}^{h} \rho^j (\frac{\rho^{n+1} - \rho}{\rho - 1}), h > 0$$

Proof. Note that $B_n = \sum_{i=1}^{n} X_n$ with $E[X_n] = 0$. Then use linearity of expectation, we can get (1).(2)

$$\begin{aligned} Var[B_n] &= \sum_{i=1}^{n} Var[X_i] + 2 \sum_{i=1}^{n-1} \sum_{j>i}^{n} Cov(X_i, X_j) \\ &= n\sigma^2 + 2 \sum_{i=1}^{n-1} \sum_{j>i}^{n} \sigma^2 \rho^{j-i} \\ &= n\sigma^2 + 2\sigma^2 \sum_{i=1}^{n-1} (n-i)\rho^i \\ &= \sigma^2 (n + \frac{2\rho(n-1)}{1-\rho} - \frac{2\rho^2(1-\rho^{n-1})}{(1-\rho)^2}). \end{aligned}$$

(3)

$$\begin{aligned} Cov(B_n, B_{n+h}) &= Cov(B_n, B_n + \sum_{i=n+1}^{n+h} X_i) \\ &= Cov(B_n, B_n) + Cov(B_n, \sum_{i=n+1}^{n+h} X_i) \\ &= Var[B_n] + \sigma^2 \sum_{j=1}^{h} \rho^j (\rho + \rho^2 + ... + \rho^n) \\ &= Var[B_n] + \sum_{j=1}^{h} \rho^j (\frac{\rho^{n+1} - \rho}{\rho - 1}) \end{aligned}$$

□

20.6.3 Asymptotic properties

Lemma 20.6.4 (unboundedness of random walk). *With probability 1 (i.e. almost surely)*

$$\limsup_n |B_n| = \infty$$

Proof. One intuitive way to prove: given any number M, we can find a N, such that when $n > N$, any trajectories that contains $3/4n$ of up steps will be greater than M. And these trajectories have $\binom{3/4n}{n}/2^n > 0$ probability to happen. □

Corollary 20.6.0.1 (finiteness of first passage time). *Define* $T_a = \inf\{t : B_n = a, a \in \mathbb{Z}\}$. $T_a < \infty$ *almost surely.(Note that the expected first passage time will be infinite.)*

20.6.4 Gambler's ruin problems

Definition 20.6.3. *A player with some initial money plays a game. He get 1 if he wins and loses 1 otherwise. He will continue to play the games until he reach a total fortune of N, or he gets ruined(running out of money).*

Lemma 20.6.5 (winning probability). *Let P_i being probability he will reach fortune N without being ruined with initial money i. The probability he wins a single game is p; the probability he loses a single game is $q = 1 - p$. Then*

$$P_i = \begin{cases} 0, if\ i = 0 \\ 1, if\ i = N \\ pP_{i+1} + (1-p)P_{i-1}, if\ 0 < i < N \end{cases}$$

Moreover, solving this recurrence equation, we have

$$P_i = \begin{cases} \frac{1-r^i}{1-r^N}, if\ p \neq q \\ i/N, if\ p = q = 0.5 \end{cases}$$

where $r = q/p$.

Proof. We can rewrite the recurrence relationship as

$$P_{i+1} - P_i = q/p(P_i - P_{i-1})$$

and we have

$$P_{i+1} - P_i = r^i P_1$$

Add equations of different i, we have

$$P_{i+1} - P_1 = \sum_{k=1}^{i} r^k P_1$$

Use the boundary condition of $P_N = 1$, and we can solve it. □

Lemma 20.6.6 (winning probability for symmetric case, martingale method). *[3, p. 220] Let P_i being probability he will reach fortune N without being ruined with initial money i. The probability he wins a single game is $p = 0.5$; the probability he loses a single game is $q = 1 - p = 0.5$.Further let M_n be the money he has after n steps, then*

- *M_n is a martingale.*
- *M_τ is a martingale, where τ is the stopping time.*
- *The winning probability is i/N.*

Proof. (1) Straight forward. (2) using optional stopping theorem. [Theorem 18.7.2] (3)

$$E[M_\tau] = NP_i + 0(1 - P_i) = M_0 = i \implies P_i = i/N.$$

□

Lemma 20.6.7 (winning probability for asymmetric case, martingale method). *[3, p. 220] Let P_i being probability he will reach fortune N without being ruined with initial money i. The probability he wins a single game is p; the probability he loses a single game is $q = 1 - p$.Further let X_n be the money he has after n steps, and let*

$$M_n = (\frac{q}{p})^{X_n}.$$

- *M_n is a martingale.*
- *M_τ is a martingale, where τ is the stopping time.*
- *The winning probability is $\frac{r^i-1}{r^N-1}, r = q/p$.*

Proof. (1)

$$E[M_{n+1}|\mathcal{F}_n] = E[(\frac{q}{p})^{X_n}(\frac{q}{p})^{X_{n+1}-X_n}|\mathcal{F}_n] = (\frac{q}{p})^{X_n}E[(\frac{q}{p})^{X_{n+1}-X_n}|\mathcal{F}_n] = (\frac{q}{p})^{X_n}((q/p)p + (q/p)^{-1}q) = M_n.$$

(2) using optional stopping theorem. [Theorem 18.7.2] (3)

$$E[M_\tau] = r^N P_i + (1 - P_i) = M_0 = r^i \implies P_i = \frac{r^i - 1}{r^N - 1}.$$

□

Lemma 20.6.8 (mean game duration, symmetric case). *[3, p. 220] Consider the case $p = q = 0.5$. Let X_n be the money he has after n steps. Let M_i be the mean game duration with initial wealth i. Let P_i being probability he will reach fortune N without being ruined with initial money i. We have:*

- $X_n^2 - n$ *is a martingale.*
- $X_\tau^2 - -\tau$ *is a martingale, where τ is the stopping time.*
- $M_i = i(N - i)$.

Proof. (1) from Lemma 20.6.2. (2) using optional stopping theorem. [Theorem 18.7.2] (3)

$$E[X_\tau^2 - \tau] = (N^2)P_i + (0)(1 - P_i) - M_i = i^2 \implies M_i = i(N - i),$$

where $P_i = i/N$ is used. □

Lemma 20.6.9 (mean game duration, asymmetric case). *[3, p. 220] Consider the case $p \neq q$. Let X_n be the money he has after n steps. Let M_i be the mean game duration with initial wealth i. Let P_i being probability he will reach fortune N without being ruined with initial money i. We have:*

- $X_n - n(p - q)$ *is a martingale.*
- $X_\tau - \tau(p - q)$ *is a martingale, where τ is the stopping time.*
- $$M_i = \frac{1}{p - q}(N\frac{r^i - 1}{r^N - 1} - i)$$
  .

Proof. (1) from Lemma 20.6.2. (2) using optional stopping theorem. [Theorem 18.7.2] (3)

$$E[X_\tau - \tau(p - q)] = (N)P_i + 0(1 - P_i) - M_i(p - q) = i \implies M_i = \frac{1}{p - q}(N\frac{r^i - 1}{r^N - 1} - i),$$

where $P_i = i/N$ is used. □

Corollary 20.6.0.2 (asymptotic properties).

- *If $p > 0.5, r < 1$, then*
  $$\lim_{N\to\infty} P_i = 1 - r^i > 0$$
  That is, the player has non-zero probability to get infinitely rich.
- *If $p \leq 0.5, r \geq 1$, then*
  $$\lim_{N\to\infty} P_i = 0$$

That is, the player will go broke with probability 1.

Remark 20.6.1 (interpretation).

- When $N \to \infty$, we are considering the case that the player does not set a winning criterion and he will continue to play until he is ruined.
- If $p \leq 0.5$, and if **the player does not set a winning criterion**, then he will definitely go broke, with no chances of being infinitely rich.

Corollary 20.6.0.3 (lose probability). *Let P_i being probability he will go bankrupt without reaching fortune N with initial money i. Then*

$$P_i = \begin{cases} 1, if\ i = 0 \\ 0, if\ i = N \\ pP_{i+1} + (1-p)P_{i-1}, if\ 0 < i < N \end{cases}$$

Moreover, solving this recurrence equation, we have

$$P_i = \begin{cases} \frac{1-r^{N-i}}{1-r^N}, if\ p \neq q \\ 1 - i/N, if\ p = q \end{cases}$$

where $r = q/p$.

Proof. Use symmetry. □

Remark 20.6.2. Gambler's ruin problem is just special cases of absorbing Markov chain problem.

Lemma 20.6.10 (asymptotic behavior). *Let T denote the first time the player's fortune reaches finite N or gets ruined. Then $T < \infty$ almost surely, that is, the player will stop playing (because either he wins N or gets ruined)after finite number of games.*

Proof. (informal idea) in the finite state absorbing chain, the expected number of step to adsorbing state cannot be infinite. □

Remark 20.6.3 (finite money is impossible in a finite state game in the long run). If N is finite, then the player must hit N or 0 in the long run, the player cannot have finite money if he play infinitely.

Lemma 20.6.11. *Let S denotes the number of games the player played when his fortune first reach N or gets ruined. Then the expected time to win or lose given we start with 0 dollars is*

$$E_i = \begin{cases} 0, if\ i = 0 \\ 0, if\ i = N \\ 1 + pE_{n+1} + (1-p)E_{n-1}, if\ 0 < i < T \end{cases}$$

Proof. Use conditional expectation. □

Remark 20.6.4. Even though the player will definitely stop after finite number of games. The expected number of games played might be infinite.

Lemma 20.6.12. *If you start with n dollars and $p = 1/2$, and you play until you go broke, then for all $n > 0$, $P(gobroke) = 1$(that is, eventually, you will go broke no matter how rich you are initially).However, the* ***expected number of games played*** *is infinity.*

20.7 Notes on bibliography

For detailed description of non-negative matrix theory,see [5][6].

For elementary treatment on Markov chains, see [3][7][1]. In particular, for hitting time analysis in Markov chain, see [3].

For in-depth treatment on Markov chains, see [2][8][9].

For treatment of Markov chains from the perspective of non-negative matrix theory, see [10].

For Markov process, see[11].

For random walks on graph, see [12].

BIBLIOGRAPHY

1. Luenberger, D. *Introduction to dynamic systems: theory, models, and applications* (Wiley, 1979).
2. Gallager, R. G. *Stochastic processes: theory for applications* (Cambridge University Press, 2013).
3. Privault, N. *Understanding Markov Chains: Examples and Applications* ISBN: 9789814451505 (Springer Singapore, 2013).
4. Zhou, X. *A Practical Guide to Quantitative Finance Interviews* (Lulu. com, 2008).
5. Meyer, C. D. *Matrix analysis and applied linear algebra* (Siam, 2000).
6. Cinlar, E. *Introduction to Stochastic Processes* ISBN: 9780486497976 (Dover Publications, Incorporated, 2013).
7. Miranda, H.-C. *Applied Stochastic Analysis lecture notes* (NewYork University, 2015).
8. Sericola, B. *Markov Chains: Theory and Applications* ISBN: 9781118731536 (Wiley, 2013).
9. He, S. *Stochastic Process (chinese)* (Peking University, 2007).
10. Seneta, E. *Non-negative matrices and Markov chains* (Springer Science & Business Media, 2006).
11. Pavliotis, G. A. *Stochastic Processes and Applications: Diffusion Processes, the Fokker-Planck and Langevin Equations* (Springer, 2014).
12. Lovász, L. Random walks on graphs.