

Stochastic Processes

Derek Li

Contents

1	Review	2
1.1	Basic Probability Theory	2
1.2	Standard Probability Distribution	2
1.3	Infinite Series and Limit	2
1.4	Bounded, Finite, and Infinite	2
1.5	Conditioning	2
1.6	Convergence of Random Variable	2
1.7	Continuity of Probability	2
1.8	Exchanging Sum and Expectation	2
1.9	Exchanging Expectation and Limit	3
1.10	Exchanging Limit and Sum	3
1.11	Basic Linear Algebra	3
1.12	Mathematical Fact	3
2	Markov Chain Probabilities	4
2.1	Markov Chain	4
2.2	Multi-Step Transitions	5
2.3	Recurrence and Transience	6
2.4	Communicating States and Irreducibility	7
2.5	Application: Gambler's Ruin	10
3	Markov Chain Convergence	12
3.1	Stationary Distributions	12
3.2	Searching for Stationarity	12
3.3	Obstacles to Convergence	14
3.4	Convergence Theorem	14

1 Review

1.1 Basic Probability Theory

Property 1.1. If Z is non negative integer valued, then

$$\mathbb{E}[Z] = \sum_{k=1}^{\infty} P(Z \geq k).$$

1.2 Standard Probability Distribution

1.3 Infinite Series and Limit

Property 1.2. If $x_n \geq 0$ and $\sum_{n=1}^{\infty} x_n < \infty$, then $\lim_{n \rightarrow \infty} x_n = 0$.

Property 1.3. It is possible that $\sum_{n=1}^{\infty} x_n = \infty$ even if $\lim_{n \rightarrow \infty} x_n = 0$.

Property 1.4 (Squeeze's Theorem). If $\{a_n\}, \{b_n\}$ and $\{c_n\}$ are three sequences with $a_n \leq b_n \leq c_n$ and $\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} c_n = L$, then $\lim_{n \rightarrow \infty} b_n = L$.

1.4 Bounded, Finite, and Infinite

1.5 Conditioning

Property 1.5 (Law of Total Expectation). If X and Y are discrete random variables, then

$$\mathbb{E}[X] = \sum_y P(Y = y) \mathbb{E}[X|Y = y].$$

Property 1.6. If $X = \mathbf{1}_A$, then

$$P(A) = \sum_y P(Y = y) P(A|Y = y).$$

1.6 Convergence of Random Variable

1.7 Continuity of Probability

Property 1.7.

$$\lim_{n \rightarrow \infty} P(X \geq n) = P(X = \infty).$$

1.8 Exchanging Sum and Expectation

Property 1.8 (Countable Linearity). If $\{Y_n\}$ is a sequence of non negative random variables, then

$$\sum_{n=1}^{\infty} \mathbb{E}[Y_n] = \mathbb{E} \left[\sum_{n=1}^{\infty} Y_n \right].$$

Property 1.9. If x_{nk} are non negative real numbers, then

$$\sum_n \sum_k x_{nk} = \sum_k \sum_n x_{nk}.$$

1.9 Exchanging Expectation and Limit

1.10 Exchanging Limit and Sum

Property 1.10 (M -Test). Let $\{x_{nk}\}_{n,k \in \mathbb{N}}$ be real numbers. Suppose that $\lim_{n \rightarrow \infty} x_{nk}$ exists for each fixed $k \in \mathbb{N}$, and that $\sum_{k=1}^{\infty} \sup_n |x_{nk}| < \infty$. Then

$$\lim_{n \rightarrow \infty} \sum_{k=1}^{\infty} x_{nk} = \sum_{k=1}^{\infty} \lim_{n \rightarrow \infty} x_{nk}.$$

1.11 Basic Linear Algebra

1.12 Mathematical Fact

Property 1.11 (Stirling's Approximation). If n is large, then

$$n! \approx \left(\frac{n}{e}\right)^n \sqrt{2\pi n},$$

or

$$\lim_{n \rightarrow \infty} \frac{n!}{\left(\frac{n}{e}\right)^n \sqrt{2\pi n}} = 1.$$

Property 1.12 (Number Theory Lemma). If a set A of positive integers is non-empty, and satisfies additivity and $\gcd(A)$, then $\exists n_0 \in \mathbb{N}$ s.t. $n \in A, \forall n \geq n_0$.

2 Markov Chain Probabilities

2.1 Markov Chain

Definition 2.1. A discrete-time, discrete-space, time-homogeneous **Markov chain** is specified by three ingredients:

- (i) A **state space** S , any non-empty finite or countable set.
- (ii) **Initial probabilities** $\{v_i\}_{i \in S}$, where v_i is the probability of starting at i (at time 0). (So $v_i \geq 0$ and $\sum_i v_i = 1$.)
- (iii) **Transition probabilities** $\{p_{ij}\}_{i,j \in S}$, where p_{ij} is the probability of jumping to j if you start at i . (So, $p_{ij} \geq 0$ and $\sum_j p_{ij} = 1, \forall i$.)

Note. (1) Given any Markov chain, let X_n be the Markov chain's state at time n and thus X_0, X_1, \dots are random variables.

(2) At time 0, we have $P(X_0 = i) = v_i, \forall i \in S$.

(3) p_{ij} can be interpreted as conditional probabilities, i.e., if $P(X_n = i) > 0$, then

$$P(X_{n+1} = j | X_n = i) = p_{ij}, \forall i, j \in S, n = 0, 1, \dots,$$

which does not depend on n because of time-homogeneous property.

(4) The probabilities at time $n + 1$ depend only on the state at time n , i.e.,

$$P(X_{n+1} = j | X_0 = i_0, X_1 = i_1, \dots, X_n = i_n) = P(X_{n+1} = j | X_n = i_n) = p_{i_n j},$$

which is called the Markov property.

(5) The joint probabilities can be computed by relating them to conditional probabilities:

$$\begin{aligned} P(X_0 = i_0, X_1 = i_1, \dots, X_n = i_n) &= P(X_0 = i_0)P(X_1 = i_1 | X_0 = i_0) \cdots P(X_n = i_n | X_{n-1} = i_{n-1}) \\ &= v_{i_0} p_{i_0 i_1} \cdots p_{i_{n-1} i_n}, \end{aligned}$$

which completely defines the probabilities of the sequence $\{X_n\}_{n=0}^\infty$. The random sequence $\{X_n\}_{n=0}^\infty$ is the Markov chain.

Example 2.1 (Bernoulli Process). Let $0 < p < 1$. Suppose repeatedly flip a p -coin at times $1, 2, \dots$. Let X_n be the number of heads on the first n flips, then $\{X_n\}$ is a Markov chain, with $S = \{0, 1, \dots\}$, $X_0 = 0$ (i.e., $v_0 = 1$ and $v_i = 0, \forall i \neq 0$), and

$$p_{ij} = \begin{cases} p, & j = i + 1 \\ 1 - p, & j = i \\ 0, & \text{otherwise} \end{cases}.$$

Example 2.2 (Simple Random Walk). Let $0 < p < 1$. Suppose repeatedly bet \$1. Each time, you have probability p of winning \$1 and probability $1 - p$ of losing \$1. Let X_n be the net gain after n bets, then $\{X_n\}$ is a Markov chain, with $S = \mathbb{Z}$, $X_0 = a$ for some $a \in \mathbb{Z}$ (i.e., $v_a = 1$), and

$$p_{ij} = \begin{cases} p, & j = i + 1 \\ 1 - p, & j = i - 1 \\ 0, & \text{otherwise} \end{cases}.$$

If $p = \frac{1}{2}$, we call it simple symmetric random walk since $p = 1 - p$.

Example 2.3 (Ehrenfest's Urn). Suppose we have d balls, divided into two urns. At each time, we choose one of the d balls uniformly at random, and move it to the other urn. Let X_n be the number of balls in Urn 1 at time n , then $\{X_n\}$ is a Markov chain, with $S = \{0, 1, \dots, d\}$, and

$$p_{ij} = \begin{cases} \frac{i}{d}, & j = i - 1 \\ \frac{d-i}{d}, & j = i + 1 \\ 0, & \text{otherwise} \end{cases}.$$

2.2 Multi-Step Transitions

Let $\mu_i^{(n)} = P(X_n = i)$ be the probabilities at time n : at time 0, $\mu_i^{(0)} = P(X_0 = i) = v_i$; at time 1, $\mu_j^{(1)} = P(X_1 = j) = \sum_{i \in S} P(X_0 = i, X_1 = j) = \sum_{i \in S} v_i p_{ij} = \sum_{i \in S} \mu_i^{(0)} p_{ij}$ by the Law of Total Probability; at time 2, $\mu_k^{(2)} = \sum_{i \in S} \sum_{j \in S} v_i p_{ij} p_{jk}$, etc.

Let $m = |S|$ be the number of elements in S (could be infinity), $v = (v_1, v_2, \dots, v_m)$ be a $1 \times m$ row vector, $\mu^{(n)} = (\mu_1^{(n)}, \mu_2^{(n)}, \dots, \mu_m^{(n)})$ be a $1 \times m$ row vector, and

$$P = (p_{ij}) = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1} & p_{m2} & \cdots & p_{mm} \end{pmatrix}$$

be an $m \times m$ matrix. Therefore, in matrix form: $\mu^{(1)} = vP = \mu^{(0)}P$, $\mu^{(2)} = vPP = vP^2 = \mu^{(0)}P^2$. By induction, we have

$$\mu^{(n)} = vP^n = \mu^{(0)}P^n, n \in \mathbb{N}.$$

By convention, let $P^0 = I$, then $\mu^{(n)} = vP^n$ holds for $n = 0$.

Another way to track the probabilities of a Markov chain is with n -step transitions

$$p_{ij}^{(n)} = P(X_n = j | X_0 = i).$$

Since the chain is time-homogeneous, $p_{ij}^{(n)} = P(X_{m+n} = j | X_m = i), \forall m \in \mathbb{N}$. Note that $p_{ij}^{(n)} \geq 0$ and $\sum_{j \in S} p_{ij}^{(n)} = \sum_{j \in S} P_i(X_n = j) = P_i(X_n \in S) = 1$. We have $p_{ij}^{(1)} = P(X_1 = j | X_0 = i) = p_{ij}$, and $p_{ij}^{(2)} = P(X_2 = j | X_0 = i) = \sum_{k \in S} P(X_2 = j, X_1 = k | X_0 = i) = \sum_{k \in S} p_{ik} p_{kj}$, $p_{ij}^{(3)} = \sum_{k \in S} \sum_{l \in S} p_{ik} p_{kl} p_{lj}$, etc.

Therefore, in matrix form: $P^{(2)} = (p_{ij}^{(2)}) = PP = P^2$, $P^{(3)} = P^3$. By induction we have

$$P^{(n)} = P^n, n \in \mathbb{N}.$$

By convention, let $P^{(0)} = I$, then $P^{(n)} = P^n$ holds for $n = 0$.

Theorem 2.1 (Chapman-Kolmogorov Equations).

$$p_{ij}^{(m+n)} = \sum_{k \in S} p_{ik}^{(m)} p_{kj}^{(n)}, p_{ij}^{m+s+n} = \sum_{k \in S} \sum_{l \in S} p_{ik}^{(m)} p_{kl}^{(s)} p_{lj}^{(n)}, \text{ etc.}$$

Proof. By the Law of Total Probability,

$$p_{ij}^{(m+n)} = P(X_{m+n} = j | X_0 = i) = \sum_{k \in S} P(X_{m+n} = j, X_m = k | X_0 = i) = \sum_{k \in S} p_{ik}^{(m)} p_{kj}^{(n)}.$$

□

In matrix form: $P^{(m+n)} = P^{(m)}P^{(n)}$, $P^{(m+s+n)} = P^{(m)}P^{(s)}P^{(n)}$, etc.

Theorem 2.2 (Chapman-Kolmogorov Inequality).

$$p_{ij}^{(m+n)} \geq p_{ik}^{(m)} p_{kj}^{(n)},$$

for any fixed state $k \in S$, etc.

2.3 Recurrence and Transience

Let $N(i) = |\{n \geq 1 : X_n = i\}|$ be the total number of times that the chain hits i (not counting time 0) and so $N(i)$ is a random variable, possibly infinite. Let f_{ij} be the **return probability** from i to j , i.e., f_{ij} is the probability, starting from i , that the chain will eventually visit j at least once:

$$f_{ij} := P_i(X_n = j \text{ for some } n \geq 1) = P_i(N(j) \geq 1).$$

Thus, we have

$$1 - f_{ij} = P_i(X_n \neq j \text{ for all } n \geq 1).$$

Also, we have

$$P_i(\text{Chain will eventually visit } j, \text{ and then eventually visit } k) = f_{ij}f_{jk}, \text{ etc.}$$

$$\text{Hence, } P_i(N(i) \geq k) = (f_{ii})^k, P_i(N(j) \geq k) = f_{ij}(f_{jj})^{k-1}.$$

Property 2.1. $f_{ik} \geq f_{ij}f_{jk}$, etc.

Definition 2.2. A state i of a Markov chain is **recurrent** or **persistent** if

$$P_i(X_n = i \text{ for some } n \geq 1) = 1, \text{ i.e., } f_{ii} = 1.$$

Otherwise, if $f_{ii} < 1$, then i is **transient**.

Theorem 2.3 (Recurrent State Theorem). State i is recurrent iff $P_i(N(i) = \infty) = 1$ iff $\sum_{n=1}^{\infty} p_{ii}^{(n)} = \infty$.

State i is transient iff $P_i(N(i) = \infty) = 0$ iff $\sum_{n=1}^{\infty} p_{ii}^{(n)} < \infty$.

Proof. We have

$$P_i(N(i) = \infty) = \lim_{k \rightarrow \infty} P_i(N(i) \geq k) = \lim_{k \rightarrow \infty} (f_{ii})^k = \begin{cases} 1, & f_{ii} = 1 \\ 0, & f_{ii} < 1 \end{cases}.$$

Also, we have

$$\begin{aligned} \sum_{n=1}^{\infty} p_{ii}^{(n)} &= \sum_{n=1}^{\infty} P_i(X_n = i) = \sum_{n=1}^{\infty} \mathbb{E}_i[\mathbf{1}_{X_n=i}] = \mathbb{E}_i \left[\sum_{n=1}^{\infty} \mathbf{1}_{X_n=i} \right] \\ &= \mathbb{E}_i[N(i)] = \sum_{k=1}^{\infty} P_i(N(i) \geq k) = \sum_{k=1}^{\infty} (f_{ii})^k = \begin{cases} \infty, & f_{ii} = 1 \\ \frac{f_{ii}}{1-f_{ii}} < \infty, & f_{ii} < 1 \end{cases}. \end{aligned}$$

□

Corollary 2.1.

$$\sum_{n=1}^{\infty} p_{ii}^{(n)} = \sum_{k=1}^{\infty} (f_{ii})^k.$$

Example 2.4 (Simple Random Walk). State 0 is recurrent only if $p = \frac{1}{2}$.

Proof. If n is odd, then $p_{00}^{(n)} = 0$.

If n is even,

$$p_{00}^{(n)} = \binom{n}{\frac{n}{2}} p^{\frac{n}{2}} (1-p)^{\frac{n}{2}} = \frac{n!}{\left[\left(\frac{n}{2}\right)!\right]^2} p^{\frac{n}{2}} (1-p)^{\frac{n}{2}}.$$

By Stirling's approximation, we have

$$n! \approx \left(\frac{n}{e}\right)^n \sqrt{2\pi n}, \quad \left(\frac{n}{2}\right)! \approx \left(\frac{n}{2e}\right)^{\frac{n}{2}} \sqrt{\pi n}$$

and thus

$$p_{00}^{(n)} \approx [4p(1-p)]^{\frac{n}{2}} \sqrt{\frac{2}{\pi n}}.$$

If $p = \frac{1}{2}$, then

$$\sum_{n=1}^{\infty} p_{00}^{(n)} \approx \sqrt{\frac{2}{\pi}} \sum_{n=2,4,6,\dots} n^{-\frac{1}{2}} = \infty.$$

If $p \neq \frac{1}{2}$, then $4p(1-p) < 1$ and thus

$$\sum_{n=1}^{\infty} p_{00}^{(n)} \approx \sum_{n=2,4,6,\dots} [4p(1-p)]^{\frac{n}{2}} \sqrt{\frac{2}{\pi n}} < \sum_{n=2,4,6,\dots} [4p(1-p)]^{\frac{n}{2}} < \infty.$$

Therefore, if $p = \frac{1}{2}$, then state 0 is recurrent and the chain will return to state 0 infinitely often with probability 1; if $p \neq \frac{1}{2}$, then state 0 is transient and the chain will not return to state 0 infinitely often. \square

Property 2.2 (f -Expansion).

$$f_{ij} = p_{ij} + \sum_{k \in S, k \neq j} p_{ik} f_{kj}.$$

Proof. We have

$$\begin{aligned} f_{ij} &= P_i(\exists n \geq 1 : X_n = j) = \sum_{k \in S} P_i(X_1 = k, \exists n \geq 1 : X_n = j) \\ &= P_i(X_1 = j, \exists n \geq 1 : X_n = j) + \sum_{k \neq j} P_i(X_1 = k, \exists n \geq 1 : X_n = j) \\ &= p_{ij}^{(1)} + \sum_{k \neq j} p_{ik} f_{kj} = p_{ij} + \sum_{k \neq j} p_{ik} f_{kj}. \end{aligned}$$

\square

Corollary 2.2. $f_{ij} \geq p_{ij}$.

2.4 Communicating States and Irreducibility

Definition 2.3. State i **communicates** with state j , written $i \rightarrow j$, if $f_{ij} > 0$, i.e., if it is possible to get from i to j . $f_{ij} > 0$ iff $\exists m \geq 1$ s.t. $p_{ij}^{(m)} > 0$, i.e., there is some time m for which it is possible to get from i to j in m steps.

We will write $i \leftrightarrow j$ if both $i \rightarrow j$ and $j \rightarrow i$.

Definition 2.4. A Markov chain is **irreducible** if $i \rightarrow j$ for all $i, j \in S$, i.e., if $f_{ij} > 0, \forall i, j \in S$. Otherwise, it is reducible.

Lemma 2.1 (Sum Lemma). If $i \rightarrow k, l \rightarrow j$, and $\sum_{n=1}^{\infty} p_{kl}^{(n)} = \infty$, then $\sum_{n=1}^{\infty} p_{ij}^{(n)} = \infty$.

Proof. Since $i \rightarrow k, l \rightarrow j$, then $\exists m, r \geq 1$ s.t. $p_{ik}^{(m)}, p_{lj}^{(r)} > 0$. By Chapman-Kolmogorov inequality, we have $p_{ij}^{(m+s+r)} \geq p_{ik}^{(m)} p_{kl}^{(s)} p_{lj}^{(r)}$. Thus,

$$\sum_{n=1}^{\infty} p_{ij}^{(n)} \geq \sum_{n=m+1+r}^{\infty} p_{ij}^{(n)} = \sum_{s=1}^{\infty} p_{ij}^{(m+s+r)} \geq \sum_{s=1}^{\infty} p_{ik}^{(m)} p_{kl}^{(s)} p_{lj}^{(r)} = \infty.$$

□

Corollary 2.3 (Sum Corollary). If $i \leftrightarrow k$, then i is recurrent iff k is recurrent.

Proof. By Sum Lemma, we have if $i \rightarrow k, k \rightarrow i$ and then

$$\sum_{n=1}^{\infty} p_{kk}^{(n)} = \infty \Leftrightarrow \sum_{n=1}^{\infty} p_{ii}^{(n)} = \infty.$$

□

Theorem 2.4 (Cases Theorem). For an irreducible Markov chain, either

(a) $\sum_{n=1}^{\infty} p_{ij}^{(n)} = \infty, \forall i, j \in S$, and all states are recurrent.

or (b) $\sum_{n=1}^{\infty} p_{ij}^{(n)} < \infty, \forall i, j \in S$, and all states are transient.

Example 2.5 (Simple Random Walk). Simple random walk is irreducible. If $p = \frac{1}{2}$, state 0 is recurrent, then all states are recurrent and $\sum_{n=1}^{\infty} p_{ij}^{(n)} = \infty, \forall i, j \in S$. If $p \neq \frac{1}{2}$, state 0 is transient, then all states are transient and $\sum_{n=1}^{\infty} p_{ij}^{(n)} < \infty, \forall i, j \in S$.

Theorem 2.5 (Finite Space Theorem). An irreducible Markov chain on a finite state space always falls into case (a), i.e., $\sum_{n=1}^{\infty} p_{ij}^{(n)} = \infty, \forall i, j \in S$ and all states are recurrent.

Proof. Choose any state $i \in S$. Since $\sum_{j \in S} p_{ij}^{(n)} = 1$, we have

$$\sum_{j \in S} \sum_{n=1}^{\infty} p_{ij}^{(n)} = \sum_{n=1}^{\infty} \sum_{j \in S} p_{ij}^{(n)} = \sum_{n=1}^{\infty} 1 = \infty.$$

Since S is finite, then we have at least one $j \in S$ s.t. $\sum_{n=1}^{\infty} p_{ij}^{(n)} = \infty$ and thus we must be in case (a). □

Lemma 2.2 (Hit Lemma). Let H_{ij} be the event that the chain hits the state i before returning to j , i.e.,

$$H_{ij} = \{\exists n \in \mathbb{N} : X_n = i, X_m \neq j, 1 \leq m \leq n-1\}.$$

If $j \rightarrow i$ with $j \neq i$, then $P_j(H_{ij}) > 0$. In other words, if it is possible to get from j to i at all, then it is possible to get from j to i without first returning to j .

Proof (optional). Since $j \rightarrow i$, there is some possible path R from j to i , i.e., $\exists m \in \mathbb{N}$ and x_0, \dots, x_m s.t. $x_0 = j, x_m = i, p_{x_r x_{r+1}} > 0, \forall 0 \leq r \leq m-1$. Let $S = \max\{r : x_r = j\}$ be the last time path R hits j , then x_S, x_{S+1}, \dots, x_m is a possible path which goes from j to i without first returning to j . So $P_j(H_{ij}) \geq P_j(R) = p_{x_S x_{S+1}} p_{x_{S+1} x_{S+2}} \cdots p_{x_{m-1} x_m} > 0$. \square

Lemma 2.3 (*f*-Lemma). If $j \rightarrow i$ and $f_{jj} = 1$, then $f_{ij} = 1$.

Proof. If $i = j$, it is obvious. Now assume $i \neq j$. Since $j \rightarrow i$, we have $P_j(H_{ij}) > 0$. But one way to never return to j is to first hit i and then from i never return to j , i.e.,

$$P_j(\text{Never return to } j) = 1 - f_{jj} \geq P_j(H_{ij})P_i(\text{Never return to } j) = P_j(H_{ij})(1 - f_{ij}).$$

Since $f_{jj} = 1$, then

$$P_j(H_{ij})(1 - f_{ij}) = 0.$$

Since $P_j(H_{ij}) > 0$, then $1 - f_{ij} = 0 \Rightarrow f_{ij} = 1$. \square

Lemma 2.4 (Infinite Returns Lemma). For an irreducible Markov chain, if it is recurrent then $P_i(N(j) = \infty) = 1, \forall i, j \in S$; if it is transient then $P_i(N(j) = \infty) = 0, \forall i, j \in S$.

Proof. If the chain is recurrent, then $f_{ij} = f_{jj} = 1$. We have

$$P_i(N(j) = \infty) = \lim_{k \rightarrow \infty} P_i(N(j) \geq k) = \lim_{k \rightarrow \infty} f_{ij}(f_{jj})^{k-1} = 1.$$

If the chain is transient, then $f_{jj} < 1$, then similarly, $P_i(N(j) = \infty) = 0$. \square

Theorem 2.6 (Recurrence Equivalences Theorem). If a chain is irreducible, the following are equivalent:

- (1) $\exists k, l \in S$ s.t. $\sum_{n=1}^{\infty} p_{kl}^{(n)} = \infty$.
- (2) $\forall i, j \in S, \sum_{n=1}^{\infty} p_{ij}^{(n)} = \infty$.
- (3) $\exists k \in S$ s.t. $f_{kk} = 1$, i.e., k is recurrent.
- (4) $\forall j \in S, f_{jj} = 1$, i.e., all states are recurrent.
- (5) $\forall i, j \in S, f_{ij} = 1$.
- (6) $\exists k, l \in S$ s.t. $P_k(N(l) = \infty) = 1$.
- (7) $\forall i, j \in S, P_i(N(j) = \infty) = 1$.

Proof. We have (1) \Rightarrow (2) : sum lemma; (2) \Rightarrow (4) : recurrent state theorem; (4) \Rightarrow (5) : *f*-lemma; (5) \Rightarrow (3) : immediate; (3) \Rightarrow (1) : recurrent state theorem with $l = k$; (4) \Rightarrow (7) : infinite returns lemma; (7) \Rightarrow (6) : immediate; (6) \Rightarrow (3) : infinite returns lemma. \square

Theorem 2.7 (Transience Equivalences Theorem). If a chain is irreducible, the following are equivalent:

- (1) $\forall k, l \in S, \sum_{n=1}^{\infty} p_{kl}^{(n)} < \infty$.
- (2) $\exists i, j \in S$ s.t. $\sum_{n=1}^{\infty} p_{ij}^{(n)} < \infty$.
- (3) $\forall k \in S, f_{kk} < 1$, i.e., k is transient.
- (4) $\exists j \in S$ s.t. $f_{jj} < 1$, i.e., some state is transient.
- (5) $\exists i, j \in S$ s.t. $f_{ij} < 1$.
- (6) $\forall k, l \in S, P_k(N(l) = \infty) = 0$.
- (7) $\exists i, j \in S$ s.t. $P_i(N(j) = \infty) = 0$.

Example 2.6 (Simple Symmetric Random Walk). We know the simple symmetric ($p = \frac{1}{2}$) random walk is recurrent and thus $f_{ij} = 1, \forall i, j \in S$, i.e., for any conceivable pattern of values, with probability 1, the chain will eventually hit each of them in sequence. We say the chain has **infinite fluctuations**.

Property 2.3 (Closed Subset). Suppose a chain is reducible, but has a closed subset $C \subseteq S$ (i.e., $p_{ij} = 0, i \in C, j \notin C$), on which it is irreducible (i.e., $i \rightarrow j, \forall i, j \in C$). Then the recurrence equivalences theorem and all results about irreducible chains still apply to the chain restricted to C .

Example 2.7. For simple random walk with $p > \frac{1}{2}$, $f_{ij} = 1$ whenever $j > i$. Or if $p < \frac{1}{2}$ and $j < i$, then $f_{ij} = 1$.

Proof. Let $X_0 = 0$ and $Z_n = X_n - X_{n-1}$ for $n \in \mathbb{N}$. By construction, $X_n = \sum_{i=1}^n Z_i$. $\{Z_n\}$ are i.i.d. with $P(Z_n = +1) = p$ and $P(Z_n = -1) = 1 - p$. Thus by the law of large numbers,

$$\lim_{n \rightarrow \infty} \frac{1}{n} (Z_1 + \cdots + Z_n) = \mathbb{E}[Z_1] = p \cdot 1 + (1 - p) \cdot (-1) = 2p - 1 > 0.$$

Therefore,

$$\lim_{n \rightarrow \infty} Z_1 + \cdots + Z_n = \infty \Rightarrow X_n - X_0 \rightarrow \infty \Rightarrow X_n \rightarrow \infty.$$

So starting from i , the chain will converge to ∞ . If $i < j$, then to go from i to ∞ , the chain must pass through j , i.e., $f_{ij} = 1$. \square

2.5 Application: Gambler's Ruin

Let $0 < a < c$ be integers, and $0 < p < 1$. Suppose player A starts with a dollars, B starts with $c - a$ dollars and they repeatedly bet. At each bet, A wins \$1 from B with probability p or B wins \$1 from A with probability $1 - p$. If X_n is the amount of money that A has at time n , then $X_0 = a$ and $\{X_n\}$ follows a simple random walk. Let $T_i = \inf\{n \geq 0 : X_n = i\}$ be the first time A has i dollars.

First consider what is the probability that A reaches c dollars before losing all their money, i.e., $P_a(T_c < T_0)$.

Write $P_a(T_c < T_0)$ as $s(a)$ and consider it to be a function of the player's initial fortune a . It is obvious that $s(0) = 0, s(c) = 1$. On the first bet, for $1 \leq a \leq c - 1$,

$$\begin{aligned} s(a) &= P_a(T_c < T_0) = P_a(T_c < T_0, X_1 = X_0 + 1) + P_a(T_c < T_0, X_1 = X_0 - 1) \\ &= P(X_1 = X_0 + 1)P_a(T_c < T_0 | X_1 = X_0 + 1) + P(X_1 = X_0 - 1)P_a(T_c < T_0 | X_1 = X_0 - 1) \\ &= ps(a + 1) + (1 - p)p(a - 1). \end{aligned}$$

Therefore,

$$ps(a) + (1 - p)s(a) = ps(a + 1) + (1 - p)p(a - 1) \Rightarrow s(a + 1) - s(a) = \frac{1 - p}{p} [s(a) - s(a - 1)].$$

Let $x = s(1)$, then $s(1) - s(0) = x, s(2) - s(1) = \frac{1 - p}{p}x, \dots, s(a + 1) - s(a) = \left(\frac{1 - p}{p}\right)^a x$ and thus

$$\begin{aligned} s(a) &= s(a) - s(0) = [s(a) - s(a - 1)] + [s(a - 1) - s(a - 2)] + \cdots + [s(1) - s(0)] \\ &= \left[\left(\frac{1 - p}{p}\right)^{a-1} + \left(\frac{1 - p}{p}\right)^{a-2} + \cdots + \left(\frac{1 - p}{p}\right)^0 \right] x = \begin{cases} ax, & p = \frac{1}{2} \\ \frac{\left(\frac{1 - p}{p}\right)^a - 1}{\frac{1 - p}{p} - 1} x, & p \neq \frac{1}{2} \end{cases}. \end{aligned}$$

Since $s(c) = 1$, then

$$x = \begin{cases} \frac{1}{c}, & p = \frac{1}{2} \\ \frac{\frac{1-p}{p} - 1}{\left(\frac{1-p}{p}\right)^c - 1}, & p \neq \frac{1}{2} \end{cases}.$$

We then obtain Gambler's Ruin formula:

$$s(a) = \begin{cases} \frac{a}{c}, & p = \frac{1}{2} \\ \frac{\left(\frac{1-p}{p}\right)^a - 1}{\left(\frac{1-p}{p}\right)^c - 1}, & p \neq \frac{1}{2} \end{cases} := s_{c,p}(a).$$

We can also consider $r_{c,p}(a) = P_a(T_0 < T_c) = P_a(\text{Ruin})$:

$$r_{c,p}(a) = s_{c,1-p}(c-a) = \begin{cases} \frac{c-a}{c}, & p = \frac{1}{2} \\ \frac{\left(\frac{1-p}{p}\right)^{c-a} - 1}{\left(\frac{1-p}{p}\right)^c - 1}, & p \neq \frac{1}{2} \end{cases}.$$

Now we consider $P_a(T_0 < \infty)$, the probability of eventual ruin:

$$\begin{aligned} P_a(T_0 < \infty) &= \lim_{K \rightarrow \infty} P_a(T_0 < K) = \lim_{c \rightarrow \infty} P_a(T_0 < T_c) = \lim_{c \rightarrow \infty} r_{c,p}(a) \\ &= \begin{cases} 1, & p \leq \frac{1}{2} \\ \left(\frac{1-p}{p}\right)^{-a}, & p > \frac{1}{2} \end{cases}. \end{aligned}$$

Thus, eventual ruin is certain if $p \leq \frac{1}{2}$.

Finally, we consider the time $T = \min(T_0, T_c)$ when the Gambler's Ruin game ends.

Property 2.4. $P(T > mc) \leq (1 - p^c)^m$, $P(T = \infty) = 0$, and $\mathbb{E}[T] < \infty$.

Proof. If the player ever wins c bets in a row, then the game must be over. But if $T > mc$, then the player has failed to win c bets in a row, despite having m independent attempts to do so. The probability of winning c bets in a row is p^c and the probability of failing to win c bets in a row is $1 - p^c$. Thus, the probability of failing on m independent attempts is $(1 - p^c)^m$ and

$$P(T > mc) \leq (1 - p^c)^m.$$

Then

$$P(T = \infty) = \lim_{m \rightarrow \infty} P(T > mc) \leq \lim_{m \rightarrow \infty} (1 - p^c)^m = 0.$$

And

$$\begin{aligned} \mathbb{E}[T] &= \sum_{i=1}^{\infty} P(T \geq i) \leq \sum_{i=0}^{\infty} P(T \geq i) \\ &\leq P(T \geq 0) + \cdots + P(T \geq 0) + P(T \geq c) + \cdots = \sum_{j=0}^{\infty} cP(T \geq cj) \\ &\leq \sum_{j=0}^{\infty} c(1 - p^c)^j = \frac{c}{1 - (1 - p^c)} = \frac{c}{p^c} < \infty. \end{aligned}$$

Hence, with probability 1, the Gambler's Ruin game must eventually end, and the time it takes to end has finite expected value. \square

3 Markov Chain Convergence

3.1 Stationary Distributions

Definition 3.1. If π is a probability distribution on S (i.e., $\pi_i \geq 0, \forall i \in S, \sum_{i \in S} \pi_i = 1$), then π is **stationary** for a Markov chain with transition probabilities (p_{ij}) is $\sum_{i \in S} \pi_i p_{ij} = \pi_j, \forall j \in S$. In matrix notation: $\pi P = \pi$ or π is a left eigenvector for the matrix P with eigenvalue 1.

Intuitively, if the chain starts with probabilities $\{\pi_i\}$, then it will keep the same probabilities one time unit later.

Example 3.1. Suppose $|S| < \infty$, we say a chain is doubly stochastic if $\sum_{i \in S} p_{ij} = 1, \forall j \in S$ (in addition to the usual condition that $\sum_{j \in S} p_{ij} = 1, \forall i \in S$). Let π be the uniform distribution on S , i.e., $\pi_i = \frac{1}{|S|}, \forall i \in S$. Then

$$\sum_{i \in S} \pi_i p_{ij} = \frac{1}{|S|} \sum_{i \in S} p_{ij} = \frac{1}{|S|} = \pi_j, \forall j \in S.$$

Thus, $\{\pi_i\}$ is stationary.

3.2 Searching for Stationarity

Definition 3.2. A Markov chain is **reversible** (or **time reversible**, or satisfies **detailed balance**) with respect to a probability distribution $\{\pi_i\}$ if $\pi_i p_{ij} = \pi_j p_{ji}, \forall i, j \in S$.

Property 3.1. If a chain is reversible with respect to π , then π is a stationary distribution.

Proof. Reversibility means $\pi_i p_{ij} = \pi_j p_{ji}$ so for $j \in S$,

$$\sum_{i \in S} \pi_i p_{ij} = \sum_{i \in S} \pi_j p_{ji} = \pi_j \sum_{i \in S} p_{ji} = \pi_j \cdot 1 = \pi_j.$$

□

Note that the converse is false: it is possible for a chain to have a stationary distribution if it is not reversible.

Example 3.2 (Ehrenfest's Urn). Let

$$\pi_i = \binom{d}{i} \left(\frac{1}{2}\right)^i \left(\frac{1}{2}\right)^{d-i} = 2^{-d} \frac{d!}{i!(d-i)!},$$

since $\pi_i \geq 0$ and $\sum_i \pi_i = 1$, then π is a distribution. To check if π stationary, we can check the reversibility, i.e., to check if $\pi_i p_{ij} = \pi_j p_{ji}, \forall i, j \in S$.

Clearly, both sides are 0 unless $j = i + 1$ or $j = i - 1$.

If $j = i + 1$, then

$$\pi_i p_{ij} = 2^{-d} \frac{d!}{i!(d-i)!} \frac{d-i}{d} = 2^{-d} \frac{(d-1)!}{i!(d-i-1)!}$$

and

$$\pi_j p_{ji} = 2^{-d} \frac{d!}{j!(d-j)!} \frac{j}{d} = \frac{(d-1)!}{(j-1)!(d-j)!} = \frac{(d-1)!}{i!(d-i-1)!} = \pi_i p_{ij}.$$

If $j = i - 1$, similarly, we have $\pi_i p_{ij} = \pi_j p_{ji}$. Hence, it is reversible w.r.t. π and thus π is a stationary distribution.

Property 3.2 (Vanishing Probabilities Proposition). IF a Markov chain's transition probabilities have $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0, \forall i, j \in S$, then the chain does not have a stationary distribution.

Proof. If there were a stationary distribution π , then we would have $\pi_j = \sum_{i \in S} \pi_i p_{ij}^{(n)}$ for any n and thus

$$\pi_j = \lim_{n \rightarrow \infty} \pi_j = \lim_{n \rightarrow \infty} \pi_j p_{ij}^{(n)}.$$

By M -test, we have

$$\pi_j = \lim_{n \rightarrow \infty} \sum_{i \in S} \pi_i p_{ij}^{(n)} = \sum_{i \in S} \lim_{n \rightarrow \infty} \pi_i p_{ij}^{(n)} = \sum_{i \in S} 0 = 0, \forall j.$$

But then $\sum_j \pi_j = 0$, which is a contraction. Therefore, there is no stationary distribution. \square

Lemma 3.1 (Vanishing Lemma). If a Markov chain has some $k, l \in S$ with $\lim_{n \rightarrow \infty} p_{kl}^{(n)} = 0$, then for any $i, j \in S$ with $k \rightarrow i$ and $j \rightarrow l$, $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0$.

Proof. Since $k \rightarrow i, j \rightarrow l$, then we can find $r, s \in \mathbb{N}$ s.t. $p_{ki}^{(r)} > 0$ and $p_{jl}^{(s)} > 0$. By Chapman-Kolmogorov, $p_{kl}^{(r+n+s)} \geq p_{ki}^{(r)} p_{ij}^{(n)} p_{jl}^{(s)}$ and hence

$$p_{ij}^{(n)} \leq \frac{p_{kl}^{(r+n+s)}}{p_{ki}^{(r)} p_{jl}^{(s)}} \rightarrow 0$$

Since $p_{ij}^{(n)} \geq 0$, then we have

$$\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0.$$

\square

Corollary 3.1 (Vanishing Together Corollary). For an irreducible Markov chain, either (i) $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0, \forall i, j \in S$, or (ii) $\lim_{n \rightarrow \infty} p_{ij}^{(n)} \neq 0, \forall i, j \in S$.

Corollary 3.2 (Vanishing Probabilities Corollary). If an irreducible Markov chain has $\lim_{n \rightarrow \infty} p_{kl}^{(n)} = 0$ for some $k, l \in S$, then the chain does not have a stationary distribution.

Example 3.3 (Simple Random Walk). Simple random walk is irreducible. It may be recurrent or transient, depending on p . But $\lim_{n \rightarrow \infty} p_{00}^{(n)} = 0$ and thus by the vanishing together corollary, $p_{ij}^{(n)} \rightarrow 0, \forall i, j \in S$ and by vanishing probabilities corollary, simple random walk does not have a stationary distribution.

Note that for all $i, j \in S$, $\sum_{n=1}^{\infty} p_{ij}^{(n)} = \infty$ even though $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0$.

Corollary 3.3 (Transient not Stationary Corollary). An irreducible and transient Markov chain cannot have a stationary distribution.

Proof. If a chain is irreducible and transient, then by the transience equivalence theorem, $\sum_{n=1}^{\infty} p_{ij}^{(n)} < \infty, \forall i, j \in S$. Hence, $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = 0, \forall i, j \in S$. Then by the vanishing probabilities proposition, there is no stationary distribution. \square

3.3 Obstacles to Convergence

Definition 3.3. The **period** of a state i is the greatest common divisor (gcd) of the set $\{n \geq 1 : p_{ii}^{(n)} > 0\}$, i.e., the largest number m s.t. all values of n with $p_{ii}^{(n)} > 0$ are integer multiples of m . If the period of every state is 1, the chain is **aperiodic**; otherwise, the chain is **periodic**.

Property 3.3. If state i has period t and $p_{ii}^{(m)} > 0$, then m is an integer multiple of t , i.e., t divides m .

Property 3.4. If $p_{ii} > 0$, then the period of state i is 1.

Property 3.5. If $p_{ii}^{(n)} > 0$ and $p_{ii}^{(n+1)} > 0$ for some n , then the period of state i is 1.

Lemma 3.2 (Equal Periods Lemma). If $i \leftrightarrow j$, then the periods of i and of j are equal.

Proof. Let the periods of i and j be t_i and t_j respectively. Since $i \leftrightarrow j$, $\exists r, s \in \mathbb{N}$ s.t. $p_{ij}^{(r)} > 0$ and $p_{ji}^{(s)} > 0$. Also, $p_{ii}^{(r+s)} \geq p_{ij}^{(r)} p_{ji}^{(s)} > 0$ and t_i divides $r + s$.

Suppose now that $p_{jj}^{(n)} > 0$, then $p_{ii}^{(r+n+s)} \geq p_{ij}^{(r)} p_{jj}^{(n)} p_{ji}^{(s)} > 0$ so t_i divides $r + n + s$. Thus, t_i divides n . Since it holds for any n with $p_{jj}^{(n)} > 0$, it follows that t_i is a common divisor of $\{n \in \mathbb{N} : p_{jj}^{(n)} > 0\}$. We know t_j is the gcd so that $t_j \geq t_i$.

Similarly, $t_i \geq t_j$ so $t_i = t_j$. □

Corollary 3.4 (Equal Period Corollary). If a chain is irreducible, then all states have the same period.

Corollary 3.5. If a chain is irreducible and $p_{ii} > 0$ for some state i , then the chain is aperiodic.

3.4 Convergence Theorem

Theorem 3.1 (Stationary Recurrence Theorem). If a Markov chain is irreducible and has a stationary distribution, then it is recurrent.

Proof. The transient not stationary corollary says that a chain cannot be irreducible and transient had have a stationary distribution. So if a chain is irreducible and has a stationary distribution, then it cannot be transient, i.e., it must be recurrent. □

Property 3.6. If a state i has $f_{ii} > 0$ and is aperiodic, then $\exists n_0(i) \in \mathbb{N}$ s.t. $p_{ii}^{(n)} > 0, \forall n \geq n_0(i)$.

Proof. Let $A = \{n \geq 1 : p_{ii}^{(n)} > 0\}$. If $m, n \in A$, then $p_{ii}^{(m)} > 0$ and $p_{ii}^{(n)} > 0$. Thus $p_{ii}^{(m+n)} \geq p_{ii}^{(m)} p_{ii}^{(n)} > 0$, so $m + n \in A$, which shows that A satisfies additivity. Since the state i is aperiodic, then $\gcd(A) = 1$. Hence, from the number theory lemma, $\exists n_0 \in \mathbb{N}$ s.t. $\forall n \geq n_0, n \in A$, i.e., $p_{ii}^{(n)} > 0$. □

Corollary 3.6. If a chain is irreducible and aperiodic, then for any states $i, j \in S$, $\exists n_0(i, j) \in \mathbb{N}$ s.t. $p_{ij}^{(n)} > 0, \forall n \geq n_0(i, j)$.

Proof. Let $n_0(i)$ and $m \in \mathbb{N}$ s.t. $p_{ij}^{(m)} > 0$. Let $n_0(i, j) = n_0(i) + m$. If $n \geq n_0(i, j)$, then $n - m \geq n_0(i)$ and thus $p_{ij}^{(n)} \geq p_{ii}^{(n-m)} p_{ij}^{(m)} > 0$. □

Lemma 3.3 (Markov Forgetting Lemma). If a Markov chain is irreducible and aperiodic, and has stationary distribution $\{\pi_i\}$, then

$$\forall i, j, k \in S, \lim_{n \rightarrow \infty} |p_{ik}^{(n)} - p_{jk}^{(n)}| = 0.$$

Intuitively, after a long time n , the chain forgets whether it started from state i or from state j .

Proof. Define a new Markov chain $\{(X_n, Y_n)\}_{n=0}^\infty$ with state space $\bar{S} = S \times S$, and transition probabilities $\bar{p}_{(ij),(kl)} = p_{ik}p_{jl}$. Because of independence, new chain has a joint stationary distribution given by $\bar{\pi}_{(ij)} = \pi_i\pi_j$ for $i, j \in S$. Whenever $n \geq \max[n_0(i, k), n_0(j, l)]$, $\bar{p}_{(ij),(kl)}^{(n)} = p_{ik}^{(n)}p_{jl}^{(n)} > 0$. Hence the new chain is irreducible and is aperiodic. By the stationary recurrence theorem, the new chain is recurrent.

Choose any $i_0 \in S$ and let $\tau = \inf\{n \geq 0 : X_n = Y_n = i_0\}$ be the first time that the new chain hits (i_0, i_0) , i.e., that both chains equal i_0 at the same time. Since the new chain is irreducible and recurrent, the recurrence equivalences theorem says that $\bar{f}_{(ij),(i_0i_0)} = 1$, i.e., starting from (i, j) , the new chain must eventually hit (i_0, i_0) , so $P_{(ij)}(\tau < \infty) = 1$. By the law of total probability,

$$p_{ij} = P_{(ij)}(X_n = k) = \sum_{m=1}^{\infty} P_{(ij)}(X_n = k, \tau = m) = \sum_{m=1}^n P_{(ij)}(X_n = k, \tau = m) + P_{(ij)}(X_n = k, \tau > n).$$

Similarly,

$$p_{jk}^{(n)} = \sum_{m=1}^n P_{(ij)}(Y_n = k, \tau = m) + P_{(ij)}(Y_n = k, \tau > n).$$

If $n \geq m$, then

$$\begin{aligned} P_{(ij)}(X_n = k, \tau = m) &= P_{(ij)}(\tau = m)P_{(ij)}(X_n = k | \tau = m) = P_{(ij)}(\tau = m)P_{(ij)}(X_n = k | X_m = Y_m = i_0) \\ &= P_{(ij)}(\tau = m)P(X_n = k | X_m = i_0) = P_{(ij)}(\tau = m)p_{i_0, k}^{(n-m)}. \end{aligned}$$

Similarly,

$$P_{(ij)}(Y_n = k, \tau = m) = P_{(ij)}(\tau = m)p_{i_0, k}^{(n-m)}.$$

So, $P_{(ij)}(X_n = k, \tau = m) = P_{(ij)}(Y_n = k, \tau = m)$. Hence, for all $i, j, k \in S$,

$$\begin{aligned} |p_{ik}^{(n)} - p_{jk}^{(n)}| &= \left| \sum_{m=1}^n P_{(ij)}(X_n = k, \tau = m) + P_{(ij)}(X_n = k, \tau > n) \right. \\ &\quad \left. - \sum_{m=1}^n P_{(ij)}(Y_n = k, \tau = m) - P_{(ij)}(Y_n = k, \tau > n) \right| \\ &= |P_{(ij)}(X_n = k, \tau > n) - P_{(ij)}(Y_n = k, \tau > n)| \\ &\leq |P_{(ij)}(X_n = k, \tau > n)| + |P_{(ij)}(Y_n = k, \tau > n)| \\ &\leq P_{(ij)}(\tau > n) + P_{(ij)}(\tau > n) = 2P_{(ij)}(\tau > n). \end{aligned}$$

Thus, $\lim_{n \rightarrow \infty} 2P_{(ij)}(\tau > n) = 2P_{(ij)}(\tau = \infty) = 0$ since $P_{(ij)}(\tau < \infty) = 1$. \square

Theorem 3.2 (Markov Chain Convergence Theorem). If a Markov chain is irreducible, aperiodic and has a stationary distribution $\{\pi_i\}$, then $\lim_{n \rightarrow \infty} p_{ij}^{(n)} = \pi_j, \forall i, j \in S$ and $\lim_{n \rightarrow \infty} P(X_n = j) = \pi_j$ for any initial distribution $\{v_i\}$.

Corollary 3.7. If a Markov chain is irreducible and aperiodic, then it has at most one stationary distribution.

Proof. By Markov chain convergence theorem, any stationary distribution must be equal to $\lim_{n \rightarrow \infty} P(X_n = j)$, so they are all equal. \square