STA447: Stochastic Processes

Tianyu Du

March 18, 2020

Contents

1 Markov Chain Probabilities

Definition 1.1. A **discrete-time, discrete-space, and time-homogenous Markov chain** is a triple of $\mathcal{S} = (S, v, P)$ in which

(i) S represents the *state space*, which is nonempty and countable;

(ii) *initial probability* v, which is a distribution on S;

(iii) and *transition probability* (p_{ij}) satisfying

$$\sum_{j \in S} p_{ij} = 1 \quad \forall i \in S \tag{1.1}$$

Definition 1.2. A Markov chain satisfies the **time-homogenous property** if

$$P(X_{n+1} = j | X_n = i) = P(X_1 = j | X_0 = i) = p_{ij} \quad \forall n \in \mathbb{N} \tag{1.2}$$

Definition 1.3. A Markov chain satisfies the **Markov property** if

$$P(X_{n+1} = j | X_n = i_n, X_{n-1} = i_{n-1}, \cdots, X_0 = i_0) = P(X_{n+1} = j | X_n = i_n) \tag{1.3}$$

That is, the chain is *memoryless.*

Proposition 1.1. As an immediate result from the Markov property, the joint probability

$$\begin{align}
P(X_0 = i_0, X_1 = i_1, X_2 = i_2, \cdots, X_n = i_n) &= P(X_0 = i_0) P(X_1 = i_1, X_2 = i_2, \cdots, X_n = i_n | X_0 = i_0) \tag{1.4} \\
&= v_{i_0} P(X_1 = i_1 | X_0 = i_0) P(X_2 = i_2, \cdots, X_n = i_n | X_0 = i_0, X_1 = i_1) \tag{1.5} \\
&= v_{i_0} P(X_1 = i_1 | X_0 = i_0) P(X_2 = i_2, \cdots, X_n = i_n | X_1 = i_1) \text{ (Markov property)} \tag{1.6} \\
&= v_{i_0} p_{i_0 i_1} \cdots p_{i_{n-1} i_n} \tag{1.7}
\end{align}$$

Definition 1.4 (n-step Arrival Probability)**.** Let $m = |S|$ and $\mu_i^{(n)} := P(X_n = i)$ denote the probability that the state ends up at i after n step (starting point follows v).

Proposition 1.2.

$$\mu^{(n)} = vP^n \tag{1.8}$$

Proof. By the law of total expectation,

$$\begin{align}
P(X_n = i) &= \sum_{j \in S} P(X_n = i, X_{n-1} = j) \tag{1.9} \\
&= \sum_{j \in S} P(X_n = i | X_{n-1} = j) P(X_{n-1} = j) \tag{1.10} \\
&= \sum_{j \in S} P(X_{n-1} = j) p_{ij} \tag{1.11} \\
&= \sum_{j \in S} \mu_j^{(n-1)} p_{ij} \tag{1.12}
\end{align}$$

Let $\mu^{(n)} := \left[\mu_1^{(n)}, \mu_2^{(n)}, \cdots, \mu_m^{(n)}\right] \in \mathbb{R}^{1\times m}$ and $P = [p_{ij}] \in \mathbb{R}^{m\times m}$. The recurrence relation can be expressed in matrix notation as:

$$\mu^{(n)} = \mu^{(n-1)} P \tag{1.13}$$

where $\mu^{(0)} = v = [v_1, v_2, \cdots, v_m]$ by construction. Define P^0 to be the identity matrix I_m, then

$$\begin{align}
\mu^{(0)} &= v = vP^0 \tag{1.14}\\
\mu^{(1)} &= \mu^{(0)} P = vP^1 \tag{1.15}\\
&\vdots \tag{1.16}\\
\mu^{(n)} &= vP^n \tag{1.17}
\end{align}$$

■

Definition 1.5 (n-step Transition Probability)**.** Define

$$p_{ij}^{(n)} := P(X_{m+n} = j | X_m = i) \tag{1.18}$$

to be the probability of arriving state j after n steps, starting from state i[1]. By the time-homogenous property,

$$p_{ij}^{(n)} = P(X_{m+n} = j | X_m = i) \quad \forall m \in \mathbb{N} \tag{1.19}$$

Proposition 1.3. Let $P^{(n)} := [p_{ij}^{(n)}] \in \mathbb{R}^{m\times m}$, then

$$P^{(n)} = P^n \tag{1.20}$$

Proof. Initial Step: for $n = 1$, $P^{(1)} = P$ by definition.
Inductive Step: for $n \in \mathbb{N}$,

$$\begin{align}
p_{ij}^{(n+1)} &= P(X_{n+1} = j | X_0 = i) \tag{1.21}\\
&= \sum_{k\in S} P(X_{n+1} = j | X_n = k, X_0 = i) P(X_n = k | X_0 = i) \tag{1.22}\\
&= \sum_{k\in S} P(X_{n+1} = j | X_n = k) p_{ik}^{(n)} \tag{1.23}\\
&= \sum_{k\in S} p_{ik}^{(n)} p_{kj} \tag{1.24}\\
&= [P^{(n)} P]_{ij} \tag{1.25}
\end{align}$$

Therefore,

$$P^{(n+1)} = P^{(n)} P \tag{1.26}$$

and

$$P^{(n)} = P^n \tag{1.27}$$

[1] In the definition of $\mu_j^{(n)}$, the starting state is random following distribution v. While defining $p_{ij}^{(n)}$ the initial state is fixed to be i.

■

Theorem 1.1 (Chapman-Kolmogorov Equation)**.** For every $k \in S$,

$$p_{ij}^{(m+n)} = p_{ik}^{(m)} p_{kj}^{(n)} \tag{1.28}$$

For $k, \ell \in S$,

$$p_{ij}^{(m+s+n)} = p_{ik}^{(m)} p_{k\ell}^{(s)} p_{\ell j}^{(n)} \tag{1.29}$$

Theorem 1.2 (Chapman-Kolmogorov Equations (Generalization))**.** Let $n = (n_1, n_2, \cdots, n_k)$ be a multi-set of non-negative integers, then

$$P^{(\sum_{i=1}^{k} n_i)} = \prod_{i=1}^{k} P^{(n_i)} \quad (\dagger) \tag{1.30}$$

Proof. Prove by induction on the size of multi-set:
Base case is trivial for $k = 1$.
Inductive step for $k > 1$, suppose ($\dagger$) holds for every set of length k, consider another multi-set with length $k+1$: $n' = (n_1, n_2, \cdots, n_k, n_{k+1})$. Let $\delta := \sum_{i=1}^{k} n_i$.

$$\begin{aligned}
P_{ij}^{(\delta+n_{k+1})} &= P(X_{\delta+n_{k+1}} = j | X_0 = i) && (1.31)\\
&= \sum_{k \in S} P(X_{\delta+n_{k+1}} = j | X_\delta = k, X_0 = i) P(X_\delta | X_0 = i) && (1.32)\\
&= \sum_{k \in S} P(X_{\delta+n_{k+1}} = j | X_\delta = k) P(X_\delta | X_0 = i) && (1.33)\\
&= \sum_{k \in S} P(X_{n_{k+1}} = j | X_0 = k) P(X_\delta = k | X_0 = i) && (1.34)\\
&= \sum_{k \in S} p_{kj}^{n_{k+1}} p_{ik}^{(\delta)} && (1.35)\\
&= [P^{(\delta)} P^{(n_{k+1})}]_{ij} && (1.36)\\
\implies P^{(\delta+n_{k+1})} &= P^{(\delta)} P^{(n_{k+1})} && (1.37)
\end{aligned}$$

■

Corollary 1.1 (Chapman-Kolmogorov Inequality)**.** For every $k \in S$,

$$p_{ij}^{(m+n)} \geq p_{ik}^{(m)} p_{kj}^{(n)} \tag{1.38}$$

For $k, \ell \in S$,

$$p_{ij}^{(m+s+n)} \geq p_{ik}^{(m)} p_{k\ell}^{(s)} p_{\ell j}^{(n)} \tag{1.39}$$

Informal Proof. Note that $p_{ik}^{(m)} p_{kj}^{(n)}$ is exactly the probability of arriving j from i in $m+n$ steps (say, event E), conditioned on passing state k at m steps. And $p_{ij}^{(m+n)}$ is the unconditional probability of event E, which is no less than the ■

1.1 Recurrent and Transience

Notation 1.1. For an arbitrary event E,

$$P_i(E) := P(E|X_0 = i) \tag{1.40}$$
$$\mathbb{E}_i(E) := \mathbb{E}\left[E|X_0 = i\right] \tag{1.41}$$

Notation 1.2. Let $N(i) := |\{n \geq 1 : X_n = i\}|$ denote the number of times the Markov chain arrives state i. Note that $N(i)$ does not count the initial state.

Definition 1.6. Define the **return probability** from state i to j, f_{ij}, as the probability of arriving state j starting from state i. That is,

$$f_{ij} = P(\exists n \geq 1 \ s.t. \ X_n = j|X_0 = i) \tag{1.42}$$
$$= P_i(N(j) \geq 1) \tag{1.43}$$

Proposition 1.4. The probability of firstly arriving j, then arriving k (denoted as event E) starting from i equals

$$P_i(E) = f_{ij} f_{jk} \tag{1.44}$$

Proof.

$$P_i(E) = P(\exists 1 \leq m \leq n \ s.t. \ X_m = j, \ X_n = k) \tag{1.45}$$
$$= P_i(\exists 1 \leq m \leq n \ s.t. \ X_n = k|\exists m \geq 1 \ s.t. \ X_m = j) P_i(\exists m \geq 1 \ s.t. \ X_m = j) \tag{1.46}$$
$$= P_i(\exists 1 \leq m \leq n \ s.t. \ X_n = k|\exists m \geq 1 \ s.t. \ X_m = j) f_{ij} \tag{1.47}$$
$$= P(\exists 1 \leq m \leq n \ s.t. \ X_n = k|X_m = j) f_{ij} \text{ (Markov property)} \tag{1.48}$$
$$= P(\exists 1 \leq n \ s.t. \ X_n = k|X_0 = j) f_{ij} \text{ (time homogenous property)} \tag{1.49}$$
$$= f_{ij} f_{jk} \tag{1.50}$$

■

Corollary 1.2.

$$P_i(N(i) \geq k) = (f_{ii})^k \tag{1.51}$$
$$P_i(N(j) \geq k) = f_{ij}(f_{jj})^{k-1} \tag{1.52}$$

Corollary 1.3.

$$f_{ij} \geq f_{ik} f_{kj} \tag{1.53}$$

Proposition 1.5. $1 - f_{ij}$ captures the probability that the Markov chain does not return to j from i.

$$1 - f_{ij} = P_i\left(X_n \neq j \text{ for all } n \geq 1\right) \tag{1.54}$$

Definition 1.7. A state i in a Markov chain is **recurrent** if $f_{ii} = 1$. That is, starting from state i, the chain returns state i for sure. Otherwise, state i is **transient**.

Theorem 1.3 (Recurrent State Theorem)**.** The following statements are equivalent:

(i) State i is recurrent (i.e., $f_{ii} = 1$);

(ii) $P_i(N(i) = \infty) = 1$, that is, starting from state i, state i will be visited infinitely often;

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

Proof. $(i) \iff (ii)$:

$$\begin{aligned} P(N(i) = \infty | X_0 = i) &= P(\lim_{k\to\infty} N(i) \geq k | X_0 = i) && (1.55)\\ &= \lim_{k\to\infty} P(N(i) \geq k | X_0 = i) && (1.56)\\ &= \lim_{k\to\infty} (f_{ii})^k = 1 \text{ if and only if } f_{ii} = 1 && (1.57)\end{aligned}$$

$(i) \iff (iii)$:

$$\begin{aligned} \sum_{n=1}^{\infty} p_{ii}^{(n)} &= \sum_{n=1}^{\infty} P(X_n = i | X_0 = i) && (1.58)\\ &= \sum_{n=1}^{\infty} \mathbb{E}(\mathbb{1}\{X_n = i\} | X_0 = i) && (1.59)\\ &= \mathbb{E}\left(\sum_{n=1}^{\infty} \mathbb{1}\{X_n = i\} \Big| X_0 = i \right) && (1.60)\\ &= \mathbb{E}(N(i) | X_0 = i) && (1.61)\\ &= \sum_{k=1}^{\infty} k P(N(i) = k | X_0 = i) && (1.62)\\ &= \sum_{k=1}^{\infty} P(N(i) \geq k | X_0 = i) && (1.63)\\ &= \sum_{k=1}^{\infty} (f_{ii})^k && (1.64)\\ &= \infty \text{ if and only if } f_{ii} = 1 && (1.65)\end{aligned}$$

■

Theorem 1.4 (Transient State Theorem)**.** The following statements are equivalent:

(i) State i is transient;

(ii) $P_i(N(i) = \infty) = 0$, that is, state i will only be visited finitely many times;

(iii) $\sum_{n=1}^{\infty} p_{ii}^{(n)} < \infty$.

Proof. Take negation of the recurrent state theorem. ■

Lemma 1.1 (Stirling's Approximation)**.**

$$n! \approx (n/e)^n \sqrt{2\pi n} \tag{1.66}$$

Proposition 1.6. For simple random walk, if $p = 1/2$, then $f_{ii} = 1$ $\forall i \in S$. Otherwise, all states are transient.

$$\forall i \in S,\ f_{ii} = 1 \iff p = \frac{1}{2} \tag{1.67}$$

Proof. For simplicity, consider state 0 and the series $\sum_{n=1}^{\infty} p_{00}^{(n)}$.
Note that for odd n's, $p_{00}^{(n)} = 0$.
For all even n's such that $n = 2k$,

$$\begin{aligned}
\sum_{n=1}^{\infty} p_{00}^{(n)} &= \sum_{k=1}^{\infty} p_{00}^{(2k)} && (1.68)\\
&= \sum_{k=1}^{\infty} \binom{2k}{k} p^k (1-p)^k && (1.69)\\
&= \sum_{k=1}^{\infty} \frac{2k!}{(k!)^2} p^k (1-p)^k && (1.70)\\
&\approx \sum_{k=1}^{\infty} \frac{(2k/e)^{2k}\sqrt{4\pi k}}{(k^k e^{-k}\sqrt{2\pi k})^2} p^k (1-p)^k && (1.71)\\
&= \sum_{k=1}^{\infty} \frac{2^{2k} k^{2k} e^{-2k} 2\sqrt{\pi k}}{k^{2k} e^{-2k} 2\pi k} p^k (1-p)^k && (1.72)\\
&= \sum_{k=1}^{\infty} \frac{2^{2k}}{\sqrt{\pi k}} p^k (1-p)^k && (1.73)\\
&= \sum_{k=1}^{\infty} \frac{4^k}{\sqrt{\pi k}} p^k (1-p)^k && (1.74)\\
&= \sum_{k=1}^{\infty} \frac{1}{\sqrt{\pi k}} [4p(1-p)]^k && (1.75)
\end{aligned}$$

When $p = \frac{1}{2}$,

$$\begin{aligned}
\sum_{n=1}^{\infty} p_{00}^{(n)} &= \frac{1}{\sqrt{\pi}} \sum_{k=1}^{\infty} k^{-1/2} && (1.76)\\
&= \infty && (1.77)
\end{aligned}$$

When $p \neq \frac{1}{2}$,

$$\begin{aligned}
\sum_{n=1}^{\infty} p_{00}^{(n)} &< \frac{1}{\sqrt{\pi}} \sum_{k=1}^{\infty} [4\pi(1-p)]^k && (1.78)\\
&< \infty && (1.79)
\end{aligned}$$

By the recurrent state theorem, $f_{ii} = 1 \iff p = 1/2$.
For other $i \neq 0$, the prove is similar. ∎

Theorem 1.5 (f-Expansion)**.**

$$f_{ij} = p_{ij} + \sum_{k \in S \setminus \{j\}} p_{ik} f_{kj} \tag{1.80}$$

Proof.

$$
\begin{aligned}
f_{ij} &= P(\exists n \in \mathbb{Z}_{++} \ s.t. \ X_n = j | X_0 = i) && (1.81)\\
&= \sum_{k \in S} P(\exists n \in \mathbb{Z}_{++} \ s.t. \ X_n = j | X_0 = i, X_1 = k) P(X_1 = k | X_0 = i) && (1.82)\\
&= \sum_{k \in S} P(\exists n \in \mathbb{Z}_{++} \ s.t. \ X_n = j | X_1 = k) P(X_1 = k | X_0 = i) \text{ (Markov Property)} && (1.83)\\
&= \underbrace{P(\exists n \in \mathbb{Z}_{++} \ s.t. \ X_n = j | X_1 = j)}_{=1} P(X_1 = j | X_0 = i) + \sum_{k \neq j} f_{kj} P(X_1 = k | X_0 = i) && (1.84)\\
&= p_{ij} + \sum_{k \neq j} f_{kj} p_{ik} && (1.85)
\end{aligned}
$$

■

1.2 Communicating States

Definition 1.8. State i is said to **communicate** with state j, denoted as $i \to j$, if $f_{ij} > 0$. That is, it is possible to get from state i to state j given arbitrarily long period of time.

Proposition 1.7 (Equivalent Defintion)**.** The following statements are equivalent:

(i) $i \to j$;

(ii) $\exists m \geq 1, \ s.t. \ p_{ij}^{(m)} > 0$.

Proof. (Proving the negation) If $p_{ij}^{(m)} = 0$ for every $m \geq 1$, then it's impossible to get state j from state i, that's, $f_{ij} = 0$. ■

Definition 1.9. A Markov chain s **irreducible** if $i \to j \ \forall i, j \in S$.

1.3 Recurrence and Transience Equivalence Theorem

Lemma 1.2 (Sum Lemma)**.** If

(i) $i \to k$;

(ii) $\ell \to j$;

(iii) $\sum_{n=1}^{\infty} p_{k\ell}^{(n)} = \infty$.

Then, $\sum_{n=1}^{\infty} p_{ij}^{(n)} = \infty$.

Proof. Suppose $i \to k$ and $\ell \to j$, then there exists m and r such that $p_{ik}^{(m)} > 0$ and $p_{\ell j}^{(r)} > 0$.
By the Chapman-Kolmogorov inequality, $p_{ij}^{(m+n+r)} \geq p_{ik}^{(m)} p_{k\ell}^{(n)} p_{\ell j}^{(r)}$.

Then,

$$\sum_{n=1}^{\infty} p_{ij}^{(n)} \geq \sum_{n=m+r+1}^{\infty} p_{ij}^{(n)} \tag{1.86}$$

$$= \sum_{s=1}^{\infty} p_{ij}^{(m+s+r)} \tag{1.87}$$

$$\geq \sum_{s=1}^{\infty} p_{ik}^{(m)} p_{k\ell}^{(s)} p_{\ell j}^{(r)} \tag{1.88}$$

$$= p_{ik}^{(m)} p_{\ell j}^{(r)} \sum_{s=1}^{\infty} p_{k\ell}^{(s)} = \infty \tag{1.89}$$

■

Remark 1.1. Note that sum lemma is still applicable when $k = \ell$ or $i = j$.

Corollary 1.4 (Sum Corollary)**.** If $i \leftrightarrow k$, then

$$f_{ii} = 1 \iff f_{kk} = 1 \tag{1.90}$$

Proof. Provided $i \leftrightarrow k$, there exists $m, r \in \mathbb{N}$ such that

$$p_{ik}^{(m)} > 0 \tag{1.91}$$

$$p_{kj}^{(r)} > 0 \tag{1.92}$$

Suppose $f_{ii} = 1$,

$$\sum_{i=1}^{\infty} p_{kk}^{(n)} \geq \sum_{i=1}^{\infty} p_{ik}^{(m)} p_{ii}^{(s)} p_{kj}^{(r)} \tag{1.93}$$

$$\geq \sum_{s=1}^{\infty} p_{ik}^{(m)} p_{ii}^{(s)} p_{kj}^{(r)} \tag{1.94}$$

$$= p_{ik}^{(m)} p_{kj}^{(r)} \sum_{s=1}^{\infty} p_{ii}^{(s)} \tag{1.95}$$

$$= \infty \tag{1.96}$$

$$\iff f_{kk} = 1 \tag{1.97}$$

■

Theorem 1.6 (Case Theorem)**.** For an irreducible Markov chain, it is either

(a) a **recurrent** Markov chain: $\forall i \in S,\ f_{ii} = 1$ and $\sum_{n=1}^{\infty} p_{ij}^{(n)} = \infty\ \forall i, j \in S$;

(b) or a **transient** Markov chain: $\forall i \in S,\ f_{ii} < 1$ and $\sum_{n=1}^{\infty} p_{ij}^{(n)} < \infty\ \forall i, j \in S$.

Proof. Let $\mathcal{M}$ be an irreducible Markov chain, if there exists $i \neq j \in S$ such that $f_{ii} = 1$ but $f_{jj} < 1$, this leads to a contradiction to the sum corollary because irreducibility of $\mathcal{M}$ implies $i \leftrightarrow j$. Also, if there exists. some $i, j \in S$ such that $\sum_{n=1}^{\infty} p_{ij}^{(n)} = \infty$. Then for every other $k, \ell \in S$, $k \leftrightarrow i$ and $j \leftrightarrow \ell$ by the irreducibility of $\mathcal{M}$. Then $\sum_{n=1}^{\infty} p_{k\ell}^{(n)} = \infty$ by sum lemma. ■

Theorem 1.7 (Finite Space Theorem)**.** An irreducible Markov chain on a finite state space is always recurrent.

Proof. Let $i \in S$ (u.i.),

$$\sum_{j\in S}\sum_{n=1}^{\infty} p_{ij}^{(n)} = \sum_{n=1}^{\infty}\sum_{j\in S} p_{ij}^{(n)} \tag{1.98}$$

$$= \sum_{n=1}^{\infty} 1 = \infty \tag{1.99}$$

Because S is finite, $\exists k \in S$ such that $\sum_{n=1}^{\infty} p_{ik}^{(n)} = \infty$. Therefore, all states are recurrent. ∎

Theorem 1.8 (Hit-Lemma)**.** Define H_{ij} as the event in which the chain starts from j and visits i without firstly returning to j (*direct path from j to i*) [2]:

$$H_{ij} := \{\exists n \in \mathbb{N}\ s.t.\ X_n = i \wedge X_m \neq j\ \forall m < n\} \tag{1.100}$$

If $j \to i$ with $j \neq i$, then $P(H_{ij}|X_0 = j) > 0$.

Theorem 1.9 (f-Lemma)**.** For all $i, j \in S$, if $j \to i$ and $f_{jj} = 1$, then $f_{ij} = 1$.

Proof. For $i = j$, trivial.
Suppose $i \neq j$, since $j \to i$, then $P(H_{ij}|X_0 = j) > 0$.
Further,

$$P(X_n \neq j\ \forall n \in \mathbb{Z}_{++}|X_0 = j) \geq P(H_{ij}|X_0 = j)P(X_n \neq j\ \forall n \in \mathbb{Z}_{++}|X_0 = i) \tag{1.101}$$

$$\implies 0 = 1 - f_{jj} \geq P(H_{ij}|X_0 = j)(1 - f_{ij}) \tag{1.102}$$

$$\implies f_{ij} = 1 \tag{1.103}$$

∎

Lemma 1.3 (Infinite Returns Lemma)**.** For an irreducible Markov chain,

(i) if this chain is recurrent, then $P(N(j) = \infty|X_0 = i) = 1\ \forall i, j \in S$;

(ii) if this chain is transient, then $P(N(j) = \infty|X_0 = i) = 0\ \forall i, j \in S$.

Proof. Let $i, j \in S$.
Suppose the chain is irreducible and recurrent, if $i = j$, then $f_{ii} = f_{jj} = 1$.
Otherwise, $i \neq j$. Since $j \to i$, by the f-Lemma, $f_{jj} = f_{ii} = f_{ij} = f_{ji} = 1$.

$$P(N(j) = \infty|X_0 = i) = \lim_{k\to\infty} P(N(j) \geq k|X_0 = i) \tag{1.104}$$

$$= \lim_{k\to\infty} f_{ij} f_{jj}^{k-1} \tag{1.105}$$

$$= 1 \tag{1.106}$$

When the chain is transient, $f_{jj} < 1$, and $\lim_{k\to\infty} f_{ij} f_{jj}^{k-1} = 0$. ∎

Theorem 1.10 (Recurrent Equivalences Theorem)**.** For a irreducible Markov chain (so that $i \to j$ for all $i, j \in S$), the following statements are equivalent:

[2] Notation abuse: H_{ij} describes the event starting from j and ending at i, instead of the other way round.

(1) $\exists k, \ell \in S$ such that $\sum_{n=1}^{\infty} p_{k\ell}^{(n)} = \infty$;

(2) $\forall i, j \in S,\ \sum_{n=1}^{\infty} p_{ij}^{(n)} = \infty$;

(3) $\exists k \in S$ $s.t.$ $f_{kk} = 1$ (need two nodes to be the same to form a strong condition);

(4) $\forall j \in S,\ f_{jj} = 1$;

(5) $\forall i, j \in S,\ f_{ij} = 1$;

(6) $\exists k, \ell \in S$ such that $P_k(N(\ell) = \infty) = 1$;

(7) $\forall i, j \in S,\ P_i(N(j) = \infty) = 1$.

Proof. (1) $\implies$ (2) by sum lemma;
(2) $\implies$ (3) take the special case when $i = j$, use recurrent state theorem;
(3) $\implies$ (4) by sum corollary;
(4) $\implies$ (5) by f-lemma;
(5) $\implies$ (6) by infinite returns lemma;
(6) $\implies$ (7)
(7) $\implies$ (1) ■

Theorem 1.11 (Transience Equivalences Theorem)**.** For a irreducible Markov chain (so that $i \to j$ for all $i, j \in S$), the following statements are equivalent:

(1) $\forall k, \ell \in S \sum_{n=1}^{\infty} p_{k\ell}^{(n)} < \infty$;

(2) $\exists i, j \in S,\ s.t.\ \sum_{n=1}^{\infty} p_{ij}^{(n)} < \infty$;

(3) $\forall k \in S\ f_{kk} < 1$;

(4) $\exists j \in S,\ s.t.\ f_{jj} < 1$;

(5) $\exists i, j \in S,\ s.t.\ f_{ij} < 1$;

(6) $\forall k, \ell \in S,\ P_k(N(\ell) = \infty) = 0$;

(7) $\exists i, j \in S,\ s.t.\ P_i(N(j) = \infty) = 0$.

1.4 Closed Subset of a Markov Chain

Definition 1.10. For a Markov chain with state space S, then any $C \subseteq S$ satisfies

$$p_{ij} = 0 \quad \forall i \in C,\ j \notin C \tag{1.107}$$

is a **closed subset** of the original Markov chain. That is, the chain will stay in the closed subset once enters it.

Remark 1.2. All theorems hold on the closed subset as well.

Proposition 1.8. For a simple random walk, if $p \geq \frac{1}{2}$, then $f_{ij} = 1$ for every $j > i$.

2 Markov Chain Convergence

2.1 Stationary Distributions

Definition 2.1. Let $\pi \in \Delta(S)$, π is **stationary** for a Markov chain if

$$\pi_j = \sum_{i \in S} \pi_i p_{ij} \quad \forall j \in S \tag{2.1}$$

In matrix notation

$$\pi = \pi P \tag{2.2}$$

Proposition 2.1. Let π be a stationary distribution of $\mathcal{M}$, then

$$\pi_j = \sum_{i \in S} \pi_i p_{ij}^{(n)} \tag{2.3}$$

In matrix notation,

$$\pi = \pi P^n \tag{2.4}$$

Proof. Using the matrix notation, it can be shown that $\pi = \pi P^n$ for every $n \in \mathbb{N}$. Therefore,

$$\begin{aligned} \pi_j &= \sum_{i \in S} \pi_i [P^n]_{ij} && (2.5) \\ &= \pi_j = \sum_{i \in S} \pi_i p_{ij}^{(n)} \text{ since } P^{(n)} = P^n && (2.6) \end{aligned}$$

■

Definition 2.2. A chain is **doubly stochastic** if

$$\forall j \in S \ \sum_{i \in S} p_{ij} = 1 \tag{2.7}$$

That is, for every state j, the arrival probability is one.

Proposition 2.2. Uniform distribution is stationary for all finite state doubly stochastic Markov chains.

Proof. Let $\pi_i = \frac{1}{|S|}$ for all $i \in S$, then

$$\begin{aligned} \sum_{i \in S} \pi_i p_{ij} &= \sum_{i \in S} \frac{1}{|S|} p_{ij} && (2.8) \\ &= \frac{1}{|S|} \sum_{i \in S} p_{ij} && (2.9) \\ &= \frac{1}{|S|} \text{ (doubly stochastic)} && (2.10) \\ &= \pi_j && (2.11) \end{aligned}$$

■

2.2 Searching for Stationarity

Definition 2.3. A Markov chain is **reversible** with respect to a distribution π if

$$\pi_i p_{ij} = \pi_j p_{ji} \quad \forall i, j \in S \tag{2.12}$$

Theorem 2.1. If a chain is reversible with respect to π, then π is a stationary distribution.

Proof.

$$\begin{aligned}
\sum_{i \in S} \pi_i p_{ij} &= \sum_{i \in S} \pi_j p_{ji} && (2.13)\\
&= \pi_j \sum_{i \in S} p_{ji} \text{ (reverse the chain)} && (2.14)\\
&= \pi_j && (2.15)
\end{aligned}$$

■

Proposition 2.3 (Vanishing Probability Proposition)**.** For a Markov chain $\mathcal{M}$, if

$$\forall i, j \in S,\ \lim_{n\to\infty} p_{ij}^{(n)} = 0 \tag{2.16}$$

that is, the chain moves chaotically, then $\mathcal{M}$ cannot have a stationary distribution.

Proof. Suppose, for contradiction, there is a stationary distribution π. Then,

$$\begin{aligned}
\pi_j &= \lim_{n\to\infty} \pi_j && (2.17)\\
&= \lim_{n\to\infty} \sum_{i \in S} \pi_i p_{ij}^{(n)} && (2.18)\\
&= \sum_{i \in S} \lim_{n\to\infty} \pi_i p_{ij}^{(n)} && (2.19)\\
&= \sum_{i \in S} \pi_i \lim_{n\to\infty} p_{ij}^{(n)} && (2.20)\\
&= 0 \neq 1 && (2.21)
\end{aligned}$$

$\rightarrow\leftarrow$ ■

Lemma 2.1 (Vanishing Lemma)**.** If $\mathcal{M}$ has some k, ℓ such that $\lim_{n\to\infty} p_{k\ell}^{(n)} = 0$, then for all $i, j \in S$ such that $k \to i$ and $j \to \ell$, $\lim_{n\to\infty} p_{ij}^{(n)} = 0$.

Proof. Because $k \to i$ and $j \to \ell$, there exists $r, s \in \mathbb{N}$ such that

$$p_{ki}^{(r)} > 0,\ p_{j\ell}^{(s)} > 0 \tag{2.22}$$

Note that for arbitrary $n \in \mathbb{N}$,

$$\begin{aligned}
p_{k\ell}^{(r+n+s)} &\geq p_{ki}^{(r)} p_{ij}^{(n)} p_{j\ell}^{(s)} && (2.23)\\
\implies p_{ij}^{(n)} &\leq \frac{p_{k\ell}^{(r+n+s)}}{p_{ki}^{(r)} p_{j\ell}^{(s)}} && (2.24)
\end{aligned}$$

Therefore,

$$0 \geq \lim_{n\to\infty} p_{ij}^{(n)} \leq \lim_{n\to\infty} \frac{p_{k\ell}^{(r+n+s)}}{p_{ki}^{(r)} p_{j\ell}^{(s)}} \tag{2.25}$$

$$= \frac{1}{p_{ki}^{(r)} p_{j\ell}^{(s)}} \lim_{n\to\infty} p_{k\ell}^{(r+n+s)} \tag{2.26}$$

$$= \frac{1}{p_{ki}^{(r)} p_{j\ell}^{(s)}} 0 = 0 \tag{2.27}$$

Therefore,

$$\lim_{n\to\infty} p_{ij}^{(n)} = 0 \tag{2.28}$$

■

Corollary 2.1 (Vanishing Together Corollary)**.** For an irreducible Markov chain, either

(i) $\lim_{n\to\infty} p_{ij}^{(n)} = 0$ for all $i, j \in S$;

(ii) $\lim_{n\to\infty} p_{ij}^{(n)} \neq 0$ for all $i, j \in S$.

Proof. Immediate result from vanishing lemma. ■

Corollary 2.2 (Vanishing Probabilities Corollary)**.** If there exists $i, j \in S$, $\lim_{n\to\infty} p_{ij}^{(n)} = 0$, then $\mathcal{M}$ cannot have a stationary distribution.

Proof. Omitted. ■

Corollary 2.3 (Transient Not Stationary Corollary)**.** A Markov chain which is irreducible and transient cannot have a stationary distribution.

Proof.

$$\forall i, j \in S \ \sum_{n=1}^{\infty} f_{ij}^{(n)} < \infty \implies \lim_{n\to\infty} p_{ij}^{(n)} = 0 \tag{2.29}$$

■

Definition 2.4. The **period** of a state i is the greatest common divisor of the set

$$\Phi_i = \{n \geq 1 : p_{ii}^{(n)} > 0\} \tag{2.30}$$

Note that if $f_{ii} = 0$, then $\Phi = \varnothing$, and period is not well-defined.

Definition 2.5. If all states in $\mathcal{M}$ has period of 1, then $\mathcal{M}$ is said to be **aperiodic**.

Lemma 2.2 (Equal Period Lemma)**.** If $i \leftrightarrow j$, then the periods of i and j are equal.

Proof. Let t_i and t_j be the periods of i and j.
Because $i \leftrightarrow j$, there exists $r, s \in \mathbb{N}$ such that $p_{ij}^{(r)}, p_{ji}^{(s)} > 0$.
For any $n \in \mathbb{N}$ such that $p_{jj}^{(n)} > 0$ (i.e., $n \in \Phi_j$), it must be the case that

$$p_{ii}^{(r+n+s)} \geq p_{ij}^{(r)} p_{jj}^{(n)} p_{ji}^{(s)} > 0 \tag{2.31}$$

$$p_{ii}^{(r+s)} \geq p_{ij}^{(r)} p_{ji}^{(s)} > 0 \tag{2.32}$$

Therefore, $r+n+s$ and $r+s \in \Phi_i$, and $t_i | r+n+s$ and $t_i | r+s$.
Hence $t_i | n$.
Because n is chosen to be an arbitrary element in Φ_j, therefore, $t_i \leq t_j$.
Proving $t_i \geq t_j$ is similar. ∎

Corollary 2.4. If $\mathcal{M}$ is irreducible, then all states have the same period.

Proof. Follows the equal period lemma directly. ∎

Corollary 2.5. If $\mathcal{M}$ is irreducible, and $p_{ii} > 0$ for some $i \in S$ (so that state i has period 1), then the whole chain $\mathcal{M}$ is aperiodic.

Proof. Follows the equal period corollary directly. ∎

2.3 Convergence Theorem

Theorem 2.2 (Markov Chain Convergence Theorem)**.** If a Markov chain $\mathcal{M}$ is

(i) irreducible;

(ii) aperiodic;

(iii) with a stationary distribution π

(i. conditioned on initial state) then

$$\lim_{n\to\infty} p_{ij}^{(n)} = \pi_j \quad \forall i, j \in S \tag{2.33}$$

In fact, the limiting probability does not depend on initial state i.
(ii. unconditional) and for any initial probability v,

$$\lim_{n\to\infty} P(X_n = j) = \lim_{n\to\infty} \mu_j^{(n)} = \pi_j \tag{2.34}$$

Theorem 2.3 (Stationary Recurrence Theorem)**.** For an irreducible chain $\mathcal{M}$ with a stationary distribution, $\mathcal{M}$ is always recurrent.

Proof. Suppose not, this contradicts the previous result *irreducible transient Markov chain cannot have stationary distribution.* ∎

Proposition 2.4. If a state i has $f_{ii} > 0$ and is aperiodic, then there is $n_0(i) \in \mathbb{N}$ such that

$$p_{ii}^{(n)} > 0 \quad \forall n \geq n_0(i) \tag{2.35}$$

Proof. Because $f_{ii} > 0$, $\Phi_i := \{n \geq 1 : p_{ii}^{(n)} > 0\} \neq \varnothing$.
Let $m, n \in \Phi_i$, then $p_{ii}^{(m+n)} \geq p_{ii}^{(m)} p_{jj}^{(n)} > 0$, so that $m+n \in \Phi_i$.
Therefore, Φ_i satisfies additivity property.
Also, $\gcd(\Phi_i) = 1$.
Lemma show that $n \in \Phi_i$ implies $n' \in \Phi_i \ \forall n' \geq n$.
Let $n(i) \in \Phi_i$, then for all $n' \geq n(i)$, $n' \in \Phi_i$. ∎

Corollary 2.6. If a chain is irreducible and aperiodic, then for any states $i, j \in S$, there is $n_0(i, j) \in \mathbb{N}$ such that

$$p_{ij}^{(n)} > 0 \quad \forall n \geq n_0(i, j) \tag{2.36}$$

Proof. Let $n_0(i) \in \mathbb{N}$ such that for all $n' \geq n_0(i)$, $n' \in \Phi_i$.
Provided $i \to j$, there exists $m \in \mathbb{N}$ such that $p_{ij}^{(m)} > 0$.
Let $n_0(i,j) = n_0(i) + m$.
For every $n \geq n_0(i,j)$, n can be written as $n = n' + m$ for some $n' \geq n_0(i)$,

$$n' \geq n_0(i) \implies p_{ii}^{(n')} > 0 \tag{2.37}$$

Then

$$p_{ij}^{(n)} = p_{ij}^{(n'+m)} \tag{2.38}$$
$$\geq p_{ii}^{(n')} p_{ij}^{(m)} > 0 \tag{2.39}$$

■

Lemma 2.3 (Markov Forgetting Lemma). If a Markov chain $\mathcal{M}$ is

(i) irreducible;

(ii) aperiodic;

(iii) with a stationary distribution π

then for all $i, j, k \in S$, then

$$\lim_{n\to\infty} \left| p_{ik}^{(n)} - p_{jk}^{(n)} \right| = 0 \tag{2.40}$$

Proof. Omiited ■

Corollary 2.7. If $\mathcal{M}$ is irreducible and aperiodic then it has at most one stationary distribution.

Proof. Suppose $\mathcal{M}$ has a stationary distribution, then by the Markov chain convergence theorem, π_j is the limit of

$$\lim_{n\to\infty} P(X_n = j) \tag{2.41}$$

and such limit must be unique if it exists. ■

Corollary 2.8 (Generalized Version). If $\mathcal{M}$ is irreducible then it has at most one stationary distribution.

2.4 Periodic Convergence

Theorem 2.4 (Periodic Convergence Theorem). Suppose a Markov chain is irreducible, with period $b \geq 2$, and has stationary distribution π, then for all $i, j \in S$,

$$\lim_{n\to\infty} \frac{1}{b} \left[p_{ij}^{(n)} + \ldots + p_{ij}^{(n+b-1)} \right] = \pi_j \tag{2.42}$$

and

$$\lim_{n\to\infty} \frac{1}{b} \left(\mathbb{P}[X_n = j] + \mathbb{P}[X_{n+1} = j] + \ldots + \mathbb{P}[X_{n+b-1} = j] \right) = \pi_j \tag{2.43}$$

Moreover,

$$\lim_{n\to\infty} \mathbb{P}[X_n = j \text{ or } X_{n+1} = j \text{ or } \cdots \text{ or } X_{n+b-1} = j] \tag{2.44}$$

Proof. For the last equality, note that since the period is b, it must take at least b steps for the chain to return to j from j. Therefore, all events in the last equality are disjoint. ■

Lemma 2.4 (Cesaro Sum)**.** If $\lim_{n\to\infty} x_n = r$, then $\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n x_i = r$.

Theorem 2.5 (Average Probability Convergence)**.** If a Markov chain is irreducible with stationary distribution π, then for all $i, j \in S$,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{\ell=1}^{n} p_{ij}^{(\ell)} = \pi_j \tag{2.45}$$

Proof. If the chain is aperiodic, apply the Markov chain convergence theorem and Cesaro sum. Otherwise, suppose the chain has period $b \geq 2$, then by periodic convergence theorem,

$$x_n := \frac{1}{b}\left[p_{ij}^{(n)} + \ldots + p_{ij}^{(n+b-1)}\right] \to \pi_j \tag{2.46}$$

and Cesaro sum,

$$\lim_{n\to\infty} \frac{1}{n} \sum_{\ell=1}^{n} x_\ell = \pi_j \tag{2.47}$$

■

3 Random Walk on Graphs

Definition 3.1. A **graph** consists of a non-empty finite or countable set V of vertices and a symmetric weight function $w : V \times V \to \mathbb{R}_+$.

Definition 3.2. A graph is **unweighted** if for every $w, v \in V$,

(i) $d(w, v) = 1$ if and only if there is an edge between w and v;

(ii) and $d(w, v) = 0$ if and only if there is no edge in between.

Definition 3.3. For a vertex $u \in V$, the **degree** of vertex u is defined as

$$d(u) := \sum_{v\in V} w(u, v) \tag{3.1}$$

Definition 3.4. Given a vertex set V with symmetric weights w the **(simple) random walk on the (undirected) graph** (V, w) is the Markov chain with state space $S = V$ and transition probabilities $p_{uv} = \frac{w(u,v)}{d(u)}$ for all $u, v \in V$.

Definition 3.5. Consider a random walk on a graph V with degree $d(u)$. Assume

$$Z = \sum_{u\in V} d(u) = \sum_{u,v\in V} w(u, v) \tag{3.2}$$

is finite, then

$$\pi_u = \frac{d(u)}{Z} \tag{3.3}$$

is a stationary distribution for this walk.

Proof. We are going to show this random walk is reversible with respect to π. Let $u, v \in V$,

$$\pi_u p_{uv} = \frac{d(u)}{Z}\frac{w(u,v)}{d(u)} = \frac{w(u,v)}{Z} \tag{3.4}$$

$$\pi_v p_{vu} = \frac{d(v)}{Z}\frac{w(v,u)}{d(v)} = \frac{w(v,u)}{Z} \tag{3.5}$$

These two products are in fact equal because weight is symmetric. ■

Proposition 3.1. A random walk on graph is irreducible if and only if the graph is connected.

Proposition 3.2. The period of a random walk on graph is either 1 or 2, since $p_{uu}^{(2)} > 0$.

Proposition 3.3. A random walk on graph is aperiodic if and only if it's non-bipartite.

Proposition 3.4. Any cycle with odd number of vertices is non-bipartite, therefore, aperiodic.

Theorem 3.1 (Graph Convergence Theorem)**.** For a random walk on a connected non-bipartite graph, if $Z < \infty$, then

$$\lim_{n\to\infty} p_{uv}^{(n)} = \frac{d(v)}{Z} \equiv \pi_v \tag{3.6}$$

for all $u, v \in V$, and

$$\lim_{n\to\infty} \mathbf{P}\left[X_n = v\right] = \frac{d(v)}{Z} \tag{3.7}$$

for any initial probabilities.

Proof. Since the graph is irreducible and aperiodic, further it possesses a stationary distribution. By Markov chain convergence theorem, it converges to its stationary distribution. ■

Theorem 3.2 (Graph Average Convergence)**.** For a random walk on any connected graph with $Z < \infty$ (whether bipartite or not), for all $u, v \in V$,

$$\lim_{n\to\infty} \frac{1}{2}\left[p_{uv}^{(n)} + p_{uv}^{(n+1)}\right] = \frac{d(v)}{Z} \tag{3.8}$$

and

$$\lim_{n\to\infty} \frac{1}{n}\sum_{\ell=1}^{n} p_{uv}^{(\ell)} = \frac{d(v)}{Z} \tag{3.9}$$

Proof. By periodic convergence theorem. ■

4 Mean Recurrence Times

Definition 4.1. Given a Markov chain with states S, the **mean recurrence time** of a state $i \in S$ is the expected value of the time of returning state i from state i. That is,

$$m_i = \mathbb{E}_i[\inf\{n \geq 1 : X_n = i\}] = \mathbb{E}\left[\inf\{n \geq 1 : X_n = i | X_0 = i\}\right] \tag{4.1}$$

Let $\tau_i := \inf\{n \geq 1 : X_n = i\}$.

Remark 4.1. Even if state i is recurrent, it is still possible that $m_i = \infty$.

Definition 4.2. A state is **positive recurrent** if $m_i < \infty$. It is **null recurrent** if $m_i = \infty$.

Theorem 4.1 (Recurrent Time Theorem)**.** For an irreducible Markov chain, either

(i) $m_i < \infty$ for all $i \in S$, and there is a unique stationary distribution $\pi_i = \frac{1}{m_i}$;

(ii) or $m_i = \infty$ for all $i \in S$, and there is no stationary distribution.

Proposition 4.1. An irreducible Markov chain on a finite state space S always have $m_i < \infty$ for all $i \in S$, and this chain has stationary distribution $\pi_i = \frac{1}{m_i}$.

Remark 4.2. The converse to above proposition is false. There are Markov chains with stationary distribution, but has infinite state space.

Proposition 4.2. For a symmetric random walk starting from state i,

$$\begin{aligned}\infty = m_i = \mathbb{E}_i[\tau_i] &= 1 + p\mathbb{E}_{i+1}[\tau_i] + (1-p)\mathbb{E}_{i-1}[\tau_i] && (4.2)\\ &= 1 + p\mathbb{E}_i[\tau_{i-1}] + (1-p)\mathbb{E}_i[\tau_{i+1}] && (4.3)\\ &\implies \mathbb{E}_i[\tau_{i-1}] = \mathbb{E}_i[\tau_{i+1}] = \infty && (4.4)\end{aligned}$$

Therefore, on average, it takes infinite steps for the simple symmetric walk to progress one step.

5 Martingales

Assumption 5.1. We assume throughout that the random variable of consideration, X_n, has finite expectation:

$$\mathbb{E}|X_n| < \infty \tag{5.1}$$

Definition 5.1. A stochastic process $\{X_n\}_{n=0}^{\infty}$ is a **martingale** if for all n:

$$\mathbb{E}[X_{n+1}|X_{0:n}] = X_n \tag{5.2}$$

Proposition 5.1. A Markov chain with discrete space S and $X_t = i_t \in S$ is a martingale if

$$\mathbb{E}[X_{n+1}|X_{0:n}] = \sum_{j \in S} p_{i_n j} j = i_n \tag{5.3}$$

Example 5.1. A simple symmetric random walk is a Markov chain martingale.

Lemma 5.1 (Law of Total Expectation)**.**

$$\mathbb{E}\left[\mathbb{E}\left[X|Y\right]\right] = \mathbb{E}\left[X\right] \tag{5.4}$$

Proposition 5.2 (Necessary Condition of a Martingale)**.** Let $\{X_n\}_{n=0}^{\infty}$ be a martingale stochastic process, then

$$\mathbb{E}\left[X_{n+1}\right] = \mathbb{E}\left[\mathbb{E}\left[X_{n+1}|X_{0:n}\right]\right] = \mathbb{E}\left[X_n\right] \tag{5.5}$$

By induction,

$$\mathbb{E}\left[X_n\right] = \mathbb{E}\left[X_0\right] \quad \forall n \in \mathbb{N} \tag{5.6}$$

6 Stopping Times

Definition 6.1. A non-negative-inter-valued random variable T is a **stopping time** for $\{X_n\}$ if for every n, the event $\{T = n\}$ is determined only by $X_{0:n}$. That is, whether the stopping time T is reached at step n is determined completely by the history up to time n.

Motivation Given a stopping time random variable T, whether

$$\mathbb{E}[X_T] = \mathbb{E}[X_0] \tag{6.1}$$

Lemma 6.1 (Optional Stopping Lemma)**.** If $\{X_n\}$ is a martingale, and T is a bounded stopping time:

$$\exists M < \infty \ s.t. \ P(T \leq M) = 1 \tag{6.2}$$

then

$$\mathbb{E}[X_T] = \mathbb{E}[X_0] \tag{6.3}$$

Proof.

$$\begin{aligned}
\mathbb{E}[X_T] - \mathbb{E}[X_0] &= \mathbb{E}[X_T - X_0] && (6.4)\\
&= \mathbb{E}\left[\sum_{t=1}^{T} X_t - X_{t-1}\right] && (6.5)\\
&= \mathbb{E}\left[\sum_{t=1}^{M} (X_t - X_{t-1})\mathbb{1}\{t \leq T\}\right] && (6.6)\\
&= \mathbb{E}\left[\sum_{t=1}^{M} (X_t - X_{t-1})(1 - \mathbb{1}\{t > T\})\right] && (6.7)\\
&= \sum_{t=1}^{M} \mathbb{E}\left[(X_t - X_{t-1})(1 - \mathbb{1}\{t > T\})\right] && (6.8)\\
&= \sum_{t=1}^{M} \mathbb{E}\left[\mathbb{E}\left[(X_t - X_{t-1})(1 - \mathbb{1}\{T \leq t-1\})|X_{0:t-1}\right]\right] && (6.9)\\
&= \sum_{t=1}^{M} \mathbb{E}\left[(1 - \mathbb{1}\{T \leq t-1\})\mathbb{E}\left[(X_t - X_{t-1})|X_{0:t-1}\right]\right] && (6.10)\\
&= \sum_{t=1}^{M} \mathbb{E}\left[(1 - \mathbb{1}\{T \leq t-1\})(\mathbb{E}\left[X_t|X_{0:t-1}\right] - X_{t-1})\right] && (6.11)\\
&= \sum_{t=1}^{M} \mathbb{E}\left[(1 - \mathbb{1}\{T \leq t-1\})(0)\right] && (6.12)\\
&= 0 && (6.13)
\end{aligned}$$

■

Remark 6.1. If $M = \infty$, we may not exchange the summation and expectation at step (6.8).

Theorem 6.1 (Optional Stopping Theorem)**.** If $\{X_n\}_{n=0}^{\infty}$ is martingale with stopping time T such that

$$P(T < \infty) = 1 \tag{6.14}$$
$$\mathbb{E}[X_T] < \infty \tag{6.15}$$
$$\lim_{n\to\infty} \mathbb{E}[X_n \mathbb{1}\{T > n\}] = 0 \tag{6.16}$$

Then

$$\mathbb{E}[X_T] = \mathbb{E}[X_0] \tag{6.17}$$

Proof. For each $m \in \mathbb{N}$, define another random variable $S_m = \min(T, m)$. So that S_m is a bounded stopping time. And by the optimal stopping lemma,

$$\mathbb{E}[X_{S_m}] = \mathbb{E}[X_T] \tag{6.18}$$

Moreover,

$$X_{S_m} = S_{\min(T,m)} = \mathbb{1}\{T > m\}X_m + (1 - \mathbb{1}\{T > m\})X_T \tag{6.19}$$
$$\implies X_T = X_{S_m} + \mathbb{1}\{T > m\}S_T - \mathbb{1}\{T > m\}S_m \tag{6.20}$$
$$\implies \mathbb{E}[X_T] = \mathbb{E}[X_{S_m}] + \mathbb{E}[\mathbb{1}\{T > m\}S_T] - \mathbb{E}[\mathbb{1}\{T > m\}S_m] \tag{6.21}$$
$$\implies \mathbb{E}[X_T] = \mathbb{E}[X_0] + \mathbb{E}[\mathbb{1}\{T > m\}S_T] - \mathbb{E}[\mathbb{1}\{T > m\}S_m] \tag{6.22}$$

By (6.14) and (6.15), as $m \to \infty$, $\mathbb{E}[\mathbb{1}\{T > m\}S_T]$ approaches zero.
By (6.16), as $m \to \infty$, $\mathbb{E}[\mathbb{1}\{T > m\}S_m] \to 0$. Therefore,

$$\lim_{m\to\infty} \mathbb{E}[X_T] = \mathbb{E}[X_0] \tag{6.23}$$

However, $\mathbb{E}[X_T] \perp m$, it must be $\mathbb{E}[X_T] = \mathbb{E}[X_0]$. ■

Corollary 6.1 (Optional Stopping Corollary)**.** If $\{X_n\}_{n=0}^{\infty}$ is martingale with stopping time T such that

$$\exists M < \infty \; s.t. \; P(|X_n|\mathbb{1}\{n < T\} \leq M) = 1 \; \forall n \in \mathbb{N} \tag{6.24}$$
$$P(T < \infty) = 1 \tag{6.25}$$

7 Wald's Theorem

Theorem 7.1 (Wald's Theorem)**.** Consider a stochastic process $X_n := a + Z_1 + \cdots + Z_n$ where $Z_i \overset{i.i.d.}{\sim} f_Z$ with finite mean m. Let T be a stopping time for $\{X_n\}$ such that $\mathbb{E}[T] < \infty$. Then,

$$\mathbb{E}[X_T] = a + m\mathbb{E}[T] \tag{7.1}$$

8 Sequence Waiting Times

Motivation Converting the waiting time for a particular pattern in stochastic process $\{X_n\}$ to a Markov chain.

fill this part

9 Martingale Convergence Theorem

Theorem 9.1 (Martingale Convergence Theorem)**.** Any martingale $\{X_n\}_{n=0}^{\infty}$ which is either

(i) bounded below

(ii) <u>or</u> bounded above

converges to some random variable X with probability 1.

Proof. Omitted. ■

10 Branching Processes

Definition A branching process consists of a **offspring distribution** $\mu \in \Delta(\mathbb{Z}_+)$. Let $X_0 = a$ and

$$X_{n+1} = Z_{n,1} + Z_{n,2} + \cdots + Z_{n,X_n} \tag{10.1}$$

$$\text{where } Z_{n,j} \overset{i.i.d.}{\sim} \mu \tag{10.2}$$

The stochastic process $\{X_n\}_{n=0}^{\infty}$ is called a **branching process**. Note that a branching process is in fact a Markov chain with

$$p_{00} = 1,\ p_{0j} = 0\ \forall j \geq 1 \tag{10.3}$$

In general, the transition probability can be written using the convolution of offspring distributions

$$\forall i, j \geq 1,\ p_{ij} = \underbrace{(\mu * \mu * \cdots * \mu)}_{\times i}(j) \tag{10.4}$$

Define $m := \mathbb{E}[\mu]$ to be the **reproductive number**. Assume $0 < m < \infty$. Then, by induction

$$\mathbb{E}[X_{n+1} | X_0, \cdots, X_n] = mX_n \tag{10.5}$$

By applying law of total expectation and induction,

$$\mathbb{E}[X_n] = m^n \mathbb{E}[X_0] \tag{10.6}$$

Proposition 10.1. When $m < 1$, with probability 1 $X_n \to 0$.

Proposition 10.2. Assuming $\mu(0) > 0$, when $m > 1$, both $P(X_n \to \infty) > 0$ and $P(X_n \to 0) > 0$.

Definition 10.1. A branching process is **degenerate** if $\mu(1) = 1$. That is, $X_n = X_0 = a$ w.p. 1 for all $n \in \mathbb{N}$.

Proposition 10.3. When $m = 1$, the branching process is a non-negative martingale. Therefore, the martingale branching process converges to some random variable X w.p. 1. If the branching process is non-degenerate, then it converges to $X \equiv 0$ w.p. 1.

Proposition 10.4. Given $\mu(0) > 0$, then for $m \leq 1$, $X_n \to 0$ w.p. 1.

11 Discrete Stock Options

Definition 11.1. A (European call) **stock option** is the option to buy one share of the stock for some fixed strike price K at some fixed future strike date (time) $S > 0$. Hence, at time S, the option worth $\max(0, X_S - K)$.

Definition 11.2. The **fair price** of an option is the price such that no profitable arbitrage is possible.

Definition 11.3. Let V denote the set of possible values X_S could take, then the **martingale probability** is defined as $p_M \in \Delta(V)$ such that

$$\mathbb{E}_{X_S \sim p_M}[X_S] = X_0 \tag{11.1}$$

That is, p_M is the transition probability from X_0 to X_S that makes $\{X_n\}_{n=0}^{\infty}$ a martingale.

Theorem 11.1 (Martingale Pricing Principle)**.** The fair price of an option is equal to its expected value under the martingale probability. That is,

$$p^* = \mathbb{E}_{X_S \sim p_M}[\max\{0, X_S - K\}] \tag{11.2}$$

Proposition 11.1. Suppose a stock price at time 0 equals $X_0 = a$, and at time $S > 0$ equals either $X_S = d$ (down) or $X_S = u$ (up), where $d < a < u$ Then if $d < K < u$, then at time 0, the fair (no-arbitrage) price of an option to buy the stock at time S for K is equal to

$$\frac{(a-d)(u-K)}{u-d} \tag{11.3}$$