STA447: Stochastic Processes

Tianyu Du

February 1, 2020

Contents

1 Markov Chain Probabilities **2**

1.1 Recurrent and Transience 5

1.2 Communicating States 9

1.3 Recurrence and Transience Equivalence Theorem 9

1.4 Closed Subset of a Markov Chain 11

1 Markov Chain Probabilities

Definition 1.1. A **discrete-time, discrete-space, and time-homogenous Markov chain** is a triple of (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

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

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,

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

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

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

■

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.

[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.

Inductive Step: for $n \in \mathbb{N}$,

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

Therefore,

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

and

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

■

Theorem 1.1.

$$\begin{aligned}
p_{ij}^{(m+n)} &= \sum_{k \in S} p_{ik}^{(m)} p_{kj}^{(n)} && (1.28)\\
p_{ij}^{(m+s+n)} &= \sum_{k \in S} \sum_{\ell \in S} p_{ik}^{(m)} p_{k\ell}^{(s)} p_{\ell j}^{(n)} && (1.29)
\end{aligned}$$

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,

$$P_i(E) := P(E|X_0 = i) \tag{1.40}$$

$$\mathbb{E}_i(E) := \mathbb{E}[E|X_0 = i] \tag{1.41}$$

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,

$$\begin{aligned}
f_{ij} &= P(\exists n \geq 1 \ s.t. \ X_n = j|X_0 = i) && (1.42)\\
&= P(N(j) \geq 1|X_0 = i) && (1.43)\\
&= P_i(N(j) \geq 1) && (1.44)
\end{aligned}$$

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

$$P_i(E) = f_{ij} f_{jk} \tag{1.45}$$

Proof.

$$\begin{aligned}
P_i(E) &= P(\exists 1 \le m \le n \ s.t. \ X_m = j, \ X_n = k) && (1.46)\\
&= P_i(\exists 1 \le m \le n \ s.t. \ X_n = k | \exists m \ge 1 \ s.t. \ X_m = j) P_i(\exists m \ge 1 \ s.t. \ X_m = j) && (1.47)\\
&= P_i(\exists 1 \le m \le n \ s.t. \ X_n = k | \exists m \ge 1 \ s.t. \ X_m = j) f_{ij} && (1.48)\\
&= P(\exists 1 \le m \le n \ s.t. \ X_n = k | X_m = j) f_{ij} \text{ (Markov property)} && (1.49)\\
&= P(\exists 1 \le n \ s.t. \ X_n = k | X_0 = j) f_{ij} \text{ (time homogenous property)} && (1.50)\\
&= f_{ij} f_{jk} && (1.51)
\end{aligned}$$

■

Corollary 1.2.

$$P_i(N(i) \ge k) = (f_{ii})^k \tag{1.52}$$
$$P_i(N(j) \ge k) = f_{ij}(f_{jj})^{k-1} \tag{1.53}$$

Corollary 1.3.

$$f_{ij} \ge f_{ik} f_{kj} \tag{1.54}$$

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

$$1 - f_{ij} = P_i \left(X_n \ne j \text{ for all } n \ge 1 \right) \tag{1.55}$$

Definition 1.7. A state i in a Markov chain is **recurrent** if $f_{ii} = 1$. Otherwise, this state 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. $(i) \iff (ii)$:

$$\begin{aligned}
P(N(i) = \infty | X_0 = i) &= P(\lim_{k \to \infty} N(i) \ge k | X_0 = i) && (1.56)\\
&= \lim_{k \to \infty} P(N(i) \ge k | X_0 = i) && (1.57)\\
&= \lim_{k \to \infty} (f_{ii})^k = 1 \text{ if and only if } f_{ii} = 1 && (1.58)
\end{aligned}$$

$(i) \iff (iii)$:

$$\sum_{n=1}^{\infty} p_{ii}^{(n)} = \sum_{n=1}^{\infty} P(X_n = i | X_0 = i) \tag{1.59}$$

$$= \sum_{n=1}^{\infty} \mathbb{E}(1_{X_n=i} | X_0 = i) \tag{1.60}$$

$$= \mathbb{E}\left(\sum_{n=1}^{\infty} 1_{X_n=i} \middle| X_0 = i \right) \tag{1.61}$$

$$= \mathbb{E}(N(i) | X_0 = i) \tag{1.62}$$

$$= \sum_{n=k}^{\infty} kP(N(i) = k | X_0 = i) \tag{1.63}$$

$$= \sum_{n=k}^{\infty} P(N(i) \geq k | X_0 = i) \tag{1.64}$$

$$= \sum_{n=k}^{\infty} (f_{ii})^k \tag{1.65}$$

$$= \infty \text{ if and only if } f_{ii} = 1 \tag{1.66}$$

■

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)**.**

$$n! \approx (n/e)^n \sqrt{2\pi n} \tag{1.67}$$

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

$$\forall i \in S,\ f_{ii} = 1 \iff p = \frac{1}{2} \tag{1.68}$$

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 $n = 2k$,

$$\sum_{n=1}^{\infty} p_{00}^{(n)} = \sum_{k=1}^{\infty} p_{00}^{(2k)} \tag{1.69}$$

$$= \sum_{k=1}^{\infty} \binom{2k}{k} p^k (1-p)^k \tag{1.70}$$

$$= \sum_{k=1}^{\infty} \frac{2k!}{(k!)^2} p^k (1-p)^k \tag{1.71}$$

$$\approx \sum_{k=1}^{\infty} \frac{(2k/e)^{2k}\sqrt{4\pi k}}{(k^k e^{-k}\sqrt{2\pi k})^2} p^k (1-p)^k \tag{1.72}$$

$$= \sum_{k=1}^{\infty} \frac{2^{2k} k^{2k} e^{-2k} 2\sqrt{\pi k}}{k^{2k} e^{-2k} 2\pi k} p^k (1-p)^k \tag{1.73}$$

$$= \sum_{k=1}^{\infty} \frac{2^{2k}}{\sqrt{\pi k}} p^k (1-p)^k \tag{1.74}$$

$$= \sum_{k=1}^{\infty} \frac{4^k}{\sqrt{\pi k}} p^k (1-p)^k \tag{1.75}$$

$$= \sum_{k=1}^{\infty} \frac{1}{\sqrt{\pi k}} [4p(1-p)]^k \tag{1.76}$$

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

$$\sum_{n=1}^{\infty} p_{00}^{(n)} = \frac{1}{\sqrt{\pi}} \sum_{k=1}^{\infty} k^{-1/2} \tag{1.77}$$

$$= \infty \tag{1.78}$$

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

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

$$< \infty \tag{1.80}$$

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)**.**

$$f_{ij} = p_{ij} + \sum_{k \in S \setminus \{j\}} p_{ik} f_{kj} \tag{1.81}$$

Proof.

$$f_{ij} = P(\exists n \in \mathbb{Z}_{++} \ s.t. \ X_n = j | X_0 = i) \tag{1.82}$$
$$= \sum_{k \in S} P(\exists n \in \mathbb{Z}_{++} \ s.t. \ X_n = j | X_0 = i, X_1 = k) P(X_1 = k | X_0 = i) \tag{1.83}$$
$$= \sum_{k \in S} P(\exists n \in \mathbb{Z}_{++} \ s.t. \ X_n = j | X_1 = k) P(X_1 = k | X_0 = i) \text{ (Markov Property)} \tag{1.84}$$
$$= \underbrace{P(\exists n \in \mathbb{Z}_{++} \ s.t. \ X_n = j | X_1 = j)}_{=1} P(X_1 = j | X_0 = i) + \sum_{k \neq j} f_{kj} P(X_1 = k | X_0 = i) \tag{1.85}$$
$$= p_{ij} + \sum_{k \neq j} f_{kj} p_{ik} \tag{1.86}$$

■

1.2 Communicating States

Definition 1.8. State i is said to **communicate** with state j, denoted as $i \to j$, if $f_{ij} > 0$.

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

(i) $i \to j$;

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

Proof. 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 $i \to j \ \forall i, j \in S$. That is, all states are attainable regardless of the starting point.

1.3 Recurrence and Transience Equivalence Theorem

Theorem 1.6 (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,

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

$$= \sum_{s=1}^{\infty} p_{ij}^{(m+s+r)} \tag{1.88}$$

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

$$= p_{ik}^{(m)} p_{\ell j}^{(r)} \sum_{s=1}^{\infty} p_{k\ell}^{(s)} = \infty \tag{1.90}$$

■

Theorem 1.7. If $i \leftrightarrow k$, then

$$f_{ii} = 1 \iff f_{kk} = 1 \tag{1.91}$$

Proof. TODO: *Proof.* ■

Theorem 1.8 (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. TODO: *Proof.* ■

Theorem 1.9 (Finite Space Theorem)**.** An *irreducible* Markov chain on a *finite* state space is always recurrent.

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

$$\sum_{j \in S} \sum_{n=1}^{\infty} p_{ij}^{(n)} = \sum_{n=1}^{\infty} \sum_{j \in S} p_{ij}^{(n)} \tag{1.92}$$

$$= \sum_{n=1}^{\infty} 1 = \infty \tag{1.93}$$

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

Theorem 1.10 (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]:

$$H_{ij} := \{\exists n \in \mathbb{N}\ s.t.\ X_n = i \land X_m \neq j\ \forall m < n\} \tag{1.94}$$

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

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

Theorem 1.11 (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,

$$P(X_n \neq j \ \forall n \in \mathbb{Z}_{++}|X_0 = j) \geq P(H_{ij}|X_0 = j)P(X_n \neq j \ \forall n \in \mathbb{Z}_{++}|X_0 = i) \tag{1.95}$$
$$\implies 0 = 1 - f_{jj} \geq P(H_{ij}|X_0 = j)(1 - f_{ij}) \tag{1.96}$$
$$\implies f_{ij} = 1 \tag{1.97}$$

■

Theorem 1.12 (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, $f_{jj} = f_{ii} = f_{ij} = f_{ji} = 1$.

$$P(N(j) = \infty|X_0 = i) = \lim_{k\to\infty} P(N(j) \geq k|X_0 = i) \tag{1.98}$$
$$= \lim_{k\to\infty} f_{ij} f_{jj}^{k-1} \tag{1.99}$$
$$= 1 \tag{1.100}$$

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

Theorem 1.13 (Recurrent Equivalences Theorem)**.** For a irreducible Markov chain (so that $i \to j$ for all $i, j \in S$), the following statements are equivalent:

(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) $\exists k \in S \ s.t. \ f_{kk} = 1$;

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

1.4 Closed Subset of a Markov Chain