Lecture-15: Class Properties

1 Communicating classes

Definition 1.1. Let $x,y \in \mathcal{X}$. If $p_{xy}^{(n)} > 0$ for some $n \in \mathbb{Z}_+$, then we say that state y is **accessible** from state x and denote it by $x \to y$. If two states $x,y \in \mathcal{X}$ are accessible to each other, they are said to **communicate** with each other and denoted by $x \leftrightarrow y$. A set of states that communicate are called a **communicating class**.

Definition 1.2. A relation R on a set $\mathcal{X}$ is a subset of $\mathcal{X} \times \mathcal{X}$.

Definition 1.3. An equivalence relation $R \subseteq \mathcal{X} \times \mathcal{X}$ has following three properties.
Reflexivity: If $x \in \mathcal{X}$, then $(x,x) \in R$.
Symmetry: If $(x,y) \in R$, then $(y,x) \in R$.
Transitivity: If $(x,y),(y,z) \in R$, then $(x,z) \in R$.

Remark 1. Equivalence relations partition a set $\mathcal{X}$.

Proposition 1.4. *Communication is an equivalence relation.*

Proof. Reflexivity follows from zero-step transition, and symmetry follows from the definition of communicating class. For transitivity, suppose $x \leftrightarrow y$ and $y \leftrightarrow z$. Then we can find $m,n \in \mathbb{N}$ such that $p_{xy}^{(m)} > 0$ and $p_{yz}^{(n)} > 0$. From Chapman Kolmogorov equations, we have $m+n \in \mathbb{N}$ such that $p_{xz}^{(m+n)} = \sum_{w \in \mathbb{Z}_+} p_{xw}^{(m)} p_{wz}^{(n)} \geqslant p_{xy}^{(m)} p_{yz}^{(n)} > 0$. □

1.1 Irreducibility and periodicity

A consequence of the previous result is that communicating classes are disjoint or identical.

Definition 1.5. A Markov chain with a single communicating class is called **irreducible**.

Definition 1.6. A **class property** is the one that is satisfied by all states in the communicating class.

Remark 2. We will see many examples of class properties. Once we have shown that a property is a class property, then one only needs to check that one of the states in the communicating class has the property for the entire class to have that.

Definition 1.7. We denote the set of recurrence times for a Markov chain with transition probability matrix $P : \mathcal{X} \times \mathcal{X} \to [0,1]$ as $A_x \triangleq \left\{ n \in \mathbb{N} : p_{xx}^{(n)} > 0 \right\}$.

Remark 3. If one can re-visit a state x in m and n steps, then also in $m+n$ steps, since $p_{xx}^{(m+n)} \geqslant p_{xx}^{(m)} p_{xx}^{(n)}$. It follows that set A_x is closed under addition for all $x \in \mathcal{X}$.

Definition 1.8. The **period** of state x is defined as $d(x) \triangleq \gcd(A_x)$. If the period is 1, we say the state is **aperiodic**.

Proposition 1.9. *Periodicity is a class property.*

Proof. We will show that for two communicating states $x \leftrightarrow y$, the periodicities are identical. We will show that $d(x)|d(y)$ and $d(y)|d(x)$. We choose $m,n \in \mathbb{N}$ such that

$$p_{xx}^{(m+n)} \geqslant p_{xy}^{(m)} p_{yx}^{(n)} > 0, \qquad p_{yy}^{(m+n)} \geqslant p_{yx}^{(n)} p_{xy}^{(m)} > 0.$$

It follows that $m+n \in A_x \cap A_y$. Let $s \in A_x$, then it follows that $m+n+s \in A_y$, since $p_{yy}^{(n+s+m)} \geqslant p_{yx}^{(n)} p_{xx}^{(s)} p_{xy}^{(m)} > 0$. Hence $d(y)|n+m$ and $d(y)|n+s+m$ which implies $d(y)|s$. Since the choice of $s \in A_x$ was arbitrary, it follows that $d(y)|d(x)$. Similarly, we can show that $d(x)|d(y)$. □

Example 1.10 (Random walk on a ring). Let $G = (\mathcal{X}, E)$ be a finite graph where $\mathcal{X} \triangleq \{0, \dots, n-1\}$ and $E = \{(x, x+1) : x \in \mathcal{X}\}$ where addition is modulo n. Let $Z : \Omega \to \{-1, 1\}^{\mathbb{N}}$ be a random *i.i.d.* sequence of step-sizes with $\mathbb{E}Z_n = 2p - 1$. We denote the location of particle after n random steps by $X_n \triangleq X_0 + \sum_{i=1}^{n} Z_i$. It follows that the random walk $X : \Omega \to \mathcal{X}^{\mathbb{N}}$ is an irreducible homogeneous Markov chain with period 2 if n is even. The Markov chain X is aperiodic if n is odd.

Proposition 1.11. *If a Markov chain $X : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$ on a finite state space $\mathcal{X}$ is irreducible and aperiodic, then there exists an integer n_0 such that $p_{xy}^{(n)} > 0$ for all $x, y \in \mathcal{X}$ and $n \geqslant n_0$.*

Proof. Since periodicity is a class property, it follows that $\gcd(A_x) = 1$ for all states $x \in \mathcal{X}$. Further, we have $m_x \in A_x$ such that $n \in A_x$ for all $n \geqslant m_x$. Further for any pair of states $x, y \in \mathcal{X}$, we can find $n_{xy} \in \mathbb{N}$ such that $p_{xy}^{(n_{xy})} > 0$ from the irreducibility of the Markov chain. It follows that $p_{xy}^{(n)} > 0$ for all $n \geqslant n_{xy} + m_y \in \mathbb{N}$. Since the state space $\mathcal{X}$ is finite, we have a finite $n_0 \triangleq \sup_{x \in \mathcal{X}} m_x + \sup_{x,y \in \mathcal{X}} n_{xy} \in \mathbb{N}$ such that $p_{xy}^{(n)} > 0$ for any state $x, y \in \mathcal{X}$ for all $n \geqslant a$. $\square$

2 Transient and recurrent states

2.1 Hitting and return times

Definition 2.1. For a homogeneous Markov chain $X : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$, we can define the **first hitting time** to state $x \in \mathcal{X}$, as $\tau_x^+ \triangleq \inf\{n \in \mathbb{N} : X_n = x\}$. If $X_0 = x$, then τ_x^+ is called the **first return time** to state x.

Lemma 2.2. *For an irreducible Markov chain $X : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$ on finite state space $\mathcal{X}$, we have $\mathbb{E}_x \tau_y^+ < \infty$ for all states $x, y \in \mathcal{X}$.*

Proof. From the definition of irreducibility, for each pair of states $z, w \in \mathcal{X}$, we have a positive integer $n_{zw} \in \mathbb{N}$ such that $p_{zw}^{n_{zw}} > \epsilon_{zw} > 0$. Since the state space $\mathcal{X}$ is finite, we define $\epsilon \triangleq \inf_{z,w \in \mathcal{X}} \epsilon_{zw} > 0$ and $r \triangleq \sup_{z,w \in \mathcal{X}} n_{zw} \in \mathbb{N}$. Hence, there exists a positive integer $r \in \mathbb{N}$ and a real $\epsilon > 0$ such that $p_{zw}^{(n)} > \epsilon$ for some $n \leqslant r$ and all states $z, w \in \mathcal{X}$. It follows that $P_z(\cup_{n \in [r]} \{X_n = y\}) > \epsilon$ or $P_z\left\{\tau_y^+ > r\right\} < 1 - \epsilon$ for any initial condition $X_0 = z \in \mathcal{X}$ and state $y \in \mathcal{X}$.

Fix $k \in \mathbb{N}$. We observe that $\left\{\tau_y^+ > kr\right\} = \cup_{z \neq y}\left\{\tau_y^+ > kr, \tau_y^+ > (k-1)r, X_{(k-1)r} = z\right\}$. Therefore,

$$P_x\left\{\tau_y^+ > kr\right\} = \sum_{z \neq y} P_x\left\{\tau_y^+ > (k-1)r, X_{(k-1)r} = z\right\} P\Big(\left\{\tau_y^+ > kr\right\} \mid \left\{X_{(k-1)r} = z, \tau_y^+ > (k-1)r, X_0 = x\right\}\Big).$$

We observe that $\left\{X_{(k-1)r} = z, \tau_y^+ > (k-1)r, X_0 = x\right\} \in \mathcal{F}_{(k-1)r}$ for all $z \neq y$. From the Markov property and the time homogeneity of X, we can write

$$P(\left\{\tau_y^+ > kr\right\} \mid \left\{X_{(k-1)r} = z, \tau_y^+ > (k-1)r, X_0 = x\right\}) = P(\left\{\tau_y^+ > kr\right\} \mid \left\{X_{(k-1)r} = z\right\}) = P_z\left\{\tau_y^+ > r\right\} < (1-\epsilon).$$

It follows that $P_x\left\{\tau_y^+ > kr\right\} < P_x\left\{\tau_y^+ > (k-1)r\right\}(1-\epsilon)$. By induction, we have $P_x\left\{\tau_y^+ > kr\right\} < (1-\epsilon)^k$. Since $P_x\left\{\tau_y^+ > n\right\}$ is decreasing in n, we can write

$$\mathbb{E}_x \tau_y^+ = \sum_{k \in \mathbb{Z}_+} \sum_{i=0}^{r-1} P_x\left\{\tau_y^+ > kr + i\right\} \leqslant \sum_{k \in \mathbb{Z}_+} r P_x\left\{\tau_y^+ > kr\right\} < \frac{r}{\epsilon} < \infty.$$

$\square$

Corollary 2.3. *For an irreducible Markov chain $X : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$ on finite state space $\mathcal{X}$, we have $P_x\left\{\tau_y^+ < \infty\right\} = 1$ for all states $x, y \in \mathcal{X}$.*

Proof. This follows from the fact that τ_y^+ is a positive random variable with finite mean for all states $y \in \mathcal{X}$ and any initial state $x \in \mathcal{X}$. $\square$

2.2 Recurrence and transience

Definition 2.4. We denote the probability of the first transition into state y at time n from the initial state x by $f_{xy}^{(n)} \triangleq P_x\left\{\tau_y^+ = n\right\}$. The probability of eventually entering state y from the initial state x is denoted by $f_{xy} \triangleq P_x\left\{\tau_y^+ < \infty\right\} = \sum_{n=1}^{\infty} f_{xy}^{(n)}$.

Definition 2.5. A state y is said to be **transient** if $f_{yy} < 1$, **recurrent** if $f_{yy} = 1$, and **positive recurrent** if $\mathbb{E}_y \tau_y^+ < \infty$.

Definition 2.6. For a discrete time process $X : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$, the total number of visits to a state $y \in \mathcal{X}$ in first n steps is denoted by $N_y(n) \triangleq \sum_{i=1}^{n} \mathbb{1}_{\{X_i = y\}}$. The total number of visits to state $y \in \mathcal{X}$ is denoted by $N_y \triangleq N_y(\infty)$.

Remark 4. From the linearity of expectations and monotone convergence theorem, we get $\mathbb{E}_y N_y = \sum_{n\in\mathbb{N}} p_{yy}^{(n)}$.

Lemma 2.7. *Consider a time homogeneous Markov chain $X : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$. For each $m \in \mathbb{Z}_+$ and state $x, y \in \mathcal{X}$, we have*

$$P_x\{N_y = m\} = \begin{cases} 1 - f_{xy} & m = 0, \\ f_{xy} f_{yy}^{m-1}(1 - f_{yy}) & m \in \mathbb{N}. \end{cases}$$

Proof. We can write the following equality

$$\{N_y = m\} = \left\{\tau_y^+(m) < \infty, \tau_y^+(m+1) = \infty\right\} = \cap_{k=1}^{m}\left\{\tau_y^+(k) - \tau_y^+(k-1) < \infty\right\} \cap \left\{\tau_y^+(m+1) - \tau_y^+(m) = \infty\right\}.$$

For each $k \in \mathbb{N}$, the kth hitting time $\tau_y^+(k)$ to the state y is adapted to the natural filtration $\mathcal{F}_\bullet$ of the process X. From strong Markov property, the next return to state y is independent of the past. That is, $(\tau_y^+(k) : k \in \mathbb{N})$ is a delayed renewal sequence for initial state $X_0 = x \neq y$. It follows that

$$P_x\{N_y = m\} = P_x\left\{\tau_y^+(1) < \infty\right\} \prod_{k=2}^{m} P_y\left\{\tau_y^+(k) - \tau_y^+(k-1) < \infty\right\} P_y\left\{\tau_y^+(m+1) = \infty\right\}.$$

□

Corollary 2.8. *For a homogeneous Markov chain $X : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$, we have*

$$P_y\{N_y < \infty\} = \mathbb{1}_{\{f_{yy}<1\}}, \qquad \mathbb{E}_y N_y = \frac{f_{yy}}{1 - f_{yy}}.$$

Proof. From the additivity of probability of disjoint events and the expression for the conditional probability mass function $P_y\{N_y = m\}$ in Lemma 2.7, we obtain

$$P_y\{N_y < \infty\} = \sum_{m\in\mathbb{Z}_+} P_y\{N_y = m\} = (1 - f_{yy}) \sum_{n\in\mathbb{Z}_+} f_{yy}^{m} = \mathbb{1}_{\{f_{yy}<1\}}.$$

Similarly, can compute the mean $\mathbb{E}_y N_y$ using the conditional distribution of N_y given initial state y. □

Remark 5. In particular, this corollary implies the following.

1. A transient state is visited a finite amount of times almost surely.
2. A recurrent state is visited infinitely often almost surely.
3. Since $\sum_{y\in\mathcal{X}} N_y = \infty$, it follows that not all states can be transient in a finite state Markov chain.

Proposition 2.9. *A state $y \in \mathcal{X}$ is recurrent iff $\sum_{k\in\mathbb{N}} p_{yy}^{(k)} = \infty$.*

Proof. For any state $y \in \mathcal{X}$, we can write $p_{yy}^{(k)} = P_y\{X_k = y\} = \mathbb{E}_y \mathbb{1}_{\{X_k = y\}}$. Using monotone convergence theorem to exchange expectation and summation, we obtain $\sum_{k\in\mathbb{N}} p_{yy}^{(k)} = \mathbb{E}_y \sum_{k\in\mathbb{N}} \mathbb{1}_{\{X_k=y\}} = \mathbb{E}_y N_y$. Thus, $\sum_{k\in\mathbb{N}} p_{yy}^{(k)}$ represents the expected number of returns $\mathbb{E}_y N_y$ to a state y starting from state y, which we know to be finite if the state is transient and infinite if the state is recurrent. □

Proposition 2.10. *Transience and recurrence are class properties.*

Proof. Let $x \leftrightarrow y$. Then from the reachability, there exist some $m, n > 0$, such that $p_{xy}^{(m)} > 0$ and $p_{yx}^{(n)} > 0$. Let x be a recurrent state, then $\sum_{s\in\mathbb{Z}_+} p_{xx}^{(s)} = \infty$. We show that $\sum_{k\in\mathbb{Z}_+} p_{yy}^{(k)} = \infty$, by observing

$$\sum_{k\in\mathbb{Z}_+} p_{yy}^{(k)} \geqslant \sum_{s\in\mathbb{Z}_+} p_{yy}^{(m+n+s)} \geqslant \sum_{s\in\mathbb{Z}_+} p_{yx}^{(n)} p_{xx}^{(s)} p_{xy}^{(m)} = \infty.$$

Let x be a transient state, then $\sum_{s\in\mathbb{Z}_+} p_{xx}^{(s)} < \infty$. We show that $\sum_{k\in\mathbb{Z}_+} p_{yy}^{(k)} < \infty$, by observing

$$\sum_{k\in\mathbb{Z}_+} p_{yy}^{(k)} \leqslant \frac{\sum_{k\in\mathbb{Z}_+} p_{xx}^{(m+n+k)}}{p_{yx}^{(n)} p_{xy}^{(m)}} \leqslant \frac{\sum_{s\in\mathbb{Z}_+} p_{xx}^{(s)}}{p_{yx}^{(n)} p_{xy}^{(m)}} < \infty.$$

□

Exercise 2.11. If y is recurrent and x such that $x \leftrightarrow y$, then show that $f_{xy} = 1$.

A Bézout's identity

Lemma A.1 (Bézout). *Consider a vector $a \in (\mathbb{Z} \setminus \{0\})^n$ with gcd d, and define the set $S \triangleq \{\langle a, x\rangle : x \in \mathbb{Z}^n, \langle a, x\rangle > 0\}$. Then d is the smallest element of set S and $d|s$ for all $s \in S$.*

Proof. Let $I \subseteq \{i \in [n] : a_i > 0\}$, then $x \in \{-1, 1\}^n$ defined as $x_i = 1$ for $i \in I$ and $x_i = -1$ for $i \notin I$ ensures that $\langle a, x\rangle = \sum_{i=1}^n |a|_i > 0$. It follows that S is non-empty. Let $g = \langle a, x\rangle$ be the minimum element of S.

First, we show that $g|a_i$ for all $i \in [n]$. Let $0 < r_i < g$ be the reminder when g divides a_i. Then, we can write $r_i = a_i - gv_i$ for some $v_i \in \mathbb{Z}_+$. Therefore, $r_i = \sum_{j\neq i} -x_j v_i a_j + (1 - v_i x_i)a_i \in S$. However, this is a contraction since g is the smallest element of S, and the result follows.

Second, we show that if any $c \in \mathbb{Z}_+$ such that $c|a_i$ for all $i \in [n]$, then $c|g$ and hence $g = d$. Since $g = \langle a, x\rangle$, this implies that $c|g$ and the result follows.

Further, any $s \in S$ can be written as $s = \langle a, z\rangle$ and $d|a_i$ for all $i \in [n]$ and hence $d|s$. □

Lemma A.2. *If A is a set closed under addition and* $\gcd(A) = 1$, *then there exists $m_0 \in A$ such that $m \in A$ for all $m \geqslant m_0$.*

Proof. Let A be a set generated by positive integers $a \triangleq (a_1, a_2, \dots, a_n)$ such that $A = \{\langle a, x\rangle : x \in \mathbb{Z}_+^{\mathbb{N}}, x \neq 0\}$. Let g be the smallest element in A. If $g = 1$, then $A = \mathbb{N}$ and there is nothing to prove. We consider the case when $g > 1$. From Bézout's Lemma, there exists $v \in \mathbb{Z}^n$ such that $1 = \langle a, v\rangle$. Since $1 \notin A$, it implies that there exists a non-empty subset $I \triangleq \{i \in [n] : v_i < 0\}$. We define

$$c_1 \triangleq \sum_{i\notin I} v_i a_i \in A, \qquad c_2 \triangleq \sum_{i\in I} -v_i a_i \in A.$$

We observe that $c_1 - c_2 = 1$. Since $1 \notin A$, both $c_1, c_2 > 1$. Let $m_0 = c_2^2$, then for any $m \geqslant m_0$, we can write $m = kc_2 + \ell$ where the remainder $0 \leqslant \ell < c_2$. We can write $c_2^2 \leqslant m = kc_2 + \ell < (k+1)c_2$. Since $c_2 > 0$, this implies that $k > c_2 - 1 \geqslant \ell$. Thus, we can write

$$m = kc_2 + \ell = kc_2 + \ell(c_1 - c_2) = (k - \ell)c_2 + \ell c_1 \in A.$$

□