Lecture-17: Invariant Distribution

1 Invariant distribution

Definition 1.1. For a time homogeneous Markov chain $X : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$ with transition matrix $p \in \mathcal{M}(\mathcal{X})^{\mathcal{X}}$, a distribution $\pi \in \mathcal{M}(\mathcal{X})$ is called *invariant* if it is a left eigenvector of the probability transition matrix p with eigenvalue unity, or

$$\pi = \pi p.$$

Remark 1. Fix $n \in \mathbb{N}$. Recall that $\nu(n) \in \mathcal{M}(\mathcal{X})$ denotes the probability distribution of the Markov chain X at time step n, i.e. $\nu_x(n) = P\{X_n = x\}$ for all states $x \in \mathcal{X}$. We observe that $\nu(n) = \nu(0)p^n$ and hence if $\nu(0) = \pi$, then $\nu(n) = \pi$ for all time-steps $n \in \mathbb{N}$.

Definition 1.2. For a time homogeneous Markov chain $X : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$ with transition matrix $p \in \mathcal{M}(\mathcal{X})^{\mathcal{X}}$, the *stationary distribution* is defined as $\nu(\infty) \triangleq \lim_{n\to\infty} \nu(n)$.

Remark 2. For a Markov chain with initial distribution being invariant, the stationary distribution is invariant distribution.

Example 1.3 (Simple random walk on a directed graph). Let $G = (\mathcal{X}, E)$ be a finite directed graph. We define a simple random walk $X : \Omega \to \mathcal{X}^{\mathbb{N}}$ on this graph as a Markov chain with state space $\mathcal{X}$ and transition matrix $P \in \mathcal{M}(\mathcal{X})^{\mathcal{X}}$ where $p_{xy} \triangleq \frac{1}{\deg_{\text{out}}(x)} \mathbb{1}_{\{(x,y)\in E\}}$. We observe that vector $(\deg_{\text{out}}(x) : x \in \mathcal{X})$ is a left eigenvector of the transition matrix P with unit eigenvalue. Indeed we can very that

$$\sum_{x\in\mathcal{X}} \deg_{\text{out}}(x) p_{xy} = \sum_{x\in\mathcal{X}} \mathbb{1}_{\{(x,y)\in E\}} = \deg_{\text{out}}(y).$$

Since $\sum_{x\in\mathcal{X}} \deg_{\text{out}}(x) = 2|E|$, it follows that $\pi \in \mathcal{M}(\mathcal{X})$ defined by $\pi_x \triangleq \frac{\deg_{\text{out}}(x)}{2|E|}$ for each $x \in \mathcal{X}$, is the equilibrium distribution of this simple random walk.

1.1 Existence

Proposition 1.4. *Consider an irreducible and aperiodic homogeneous discrete time Markov chain $X : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$ with transition matrix $p \in \mathcal{M}(\mathcal{X})^{\mathcal{X}}$ and starting from initial state $X_0 = x$. We define a positive vector $\tilde{\pi}^x \in \mathbb{R}_+^{\mathcal{X}}$ as $\tilde{\pi}^x_y \triangleq \mathbb{E}_x N_y(\tau_x^+(1))$ for all $y \in \mathcal{X}$. Then the following statements hold true.*
(a) If x is recurrent, then $\tilde{\pi}^x$ is a left eigenvector of p with eigenvalue unity. That is, $\tilde{\pi}^x = \tilde{\pi}^x P$.
(b) If x is positive recurrent, then $\pi \triangleq \frac{\tilde{\pi}^x}{\mathbb{E}_x \tau_x^+(1)}$ is an invariant distribution of p.

Proof. (a) We first show that $\tilde{\pi}^x$ is a left eigenvector for the transition probability matrix p for time homogeneous discrete time Markov chain X. Recall that $N_y(\tau_x^+(1)) = \sum_{n\in\mathbb{N}} \mathbb{1}_{\{n\leqslant\tau_x^+(1)\}} \mathbb{1}_{\{X_n=y\}}$ and $p_{wz} = P(\{X_{n+1} = z\} \mid \{X_n = w\})$. Using the monotone convergence theorem, we can write

$$\sum_{w\in\mathcal{X}} \tilde{\pi}^x_w p_{wz} = \sum_{w\in\mathcal{X}} \sum_{n\in\mathbb{N}} P_x\left\{\tau_x^+(1) \geqslant n, X_n = w\right\} P(\{X_{n+1} = z\} \mid \{X_n = w\}).$$

We first focus on the term $w = x$. We see that $\{X_n = x, \tau_x^+(1) \geqslant n\} = \{\tau_x^+(1) = n\}$ and hence for a recurrent state x, we have

$$\tilde{\pi}^x_x p_{xz} = p_{xz} \sum_{n\in\mathbb{N}} P_x\left\{\tau_x^+ = n\right\} = p_{xz} P_x\left\{\tau_x^+ < \infty\right\} = p_{xz}.$$

We next focus on the terms $w \neq x$. We observe that $\{X_n = w, \tau_x^+ \geqslant n\} = \{X_n = w, \tau_x^+ \geqslant n+1\} \in \mathcal{F}_n$. From the Markov property for X, we have

$$p_{wz} = P(\{X_{n+1} = z\} \mid \{X_n = w\}) = P(\{X_{n+1} = z\} \mid \{X_n = w, \tau_x^+ \geqslant n+1, X_0 = x\}).$$

Therefore, from the definition of conditional probability, we obtain $p_{wz} P_x \{X_n = w, \tau_x^+ \geqslant n+1, X_0 = x\} = P_x \{X_{n+1} = z, X_n = w, \tau_x^+ \geqslant n+1\}$, and hence

$$\sum_{w \neq x} \tilde{\pi}_w^x p_{wz} = \sum_{n \in \mathbb{N}} \sum_{w \neq x} P_x \{X_n = w, X_{n+1} = z, \tau_x^+ \geqslant n+1\} = \sum_{n \in \mathbb{N}} P_x \{X_{n+1} = z, \tau_x^+ \geqslant n+1\}$$
$$= \tilde{\pi}_z^x - P_x \{X_1 = z, \tau_x^+ \geqslant 1\} = \tilde{\pi}_z^x - p_{xz}.$$

The result follows from summing both the cases.

(b) For a positive recurrent state x, it suffices to show that π is a distribution on state space $\mathcal{X}$. Recall that $\tilde{\pi}_y^x = \mathbb{E}_x N_y(\tau_x^+(1))$ and $\sum_{y \in \mathcal{X}} N_y(\tau_x^+(1)) = \tau_x^+(1)$, it follows from the monotone convergence theorem that

$$\sum_{y \in \mathcal{X}} \tilde{\pi}_y^x = \mathbb{E}_x \sum_{y \in \mathcal{X}} N_y(\tau_x^+(1)) = \mathbb{E}_x \tau_x^+(1) < \infty.$$

□

Remark 3. For the positive vector $\tilde{\pi}^x$ defined in Proposition 1.4, we have $0 \leqslant \tilde{\pi}_y^x \leqslant \mathbb{E}_x \tau_x^+(1)$ for all states $y \in \mathcal{X}$. In addition, $N_x(\tau_x^+(1)) = 1$ and hence $\tilde{\pi}_x^x = 1$.

1.2 Uniqueness

Definition 1.5. Consider a time homogeneous Markov chain $X : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$ with transition probability matrix $p \in \mathcal{M}(\mathcal{X})^{\mathcal{X}}$. A function $h \in \mathbb{R}^{\mathcal{X}}$ is called *harmonic at x* if $h(x) = \mathbb{E}_x h(X_1) = (ph)_x = \sum_{y \in \mathcal{X}} p_{xy} h(y)$. A function is *harmonic on a subset $D \subset \mathcal{X}$* if it is harmonic at every state $x \in D$.

Lemma 1.6. *For a finite state irreducible Markov chain, a function f harmonic on all states in $\mathcal{X}$ is a constant.*

Proof. Suppose h is not a constant, then there exists a state $x \in \mathcal{X}$, such that $h(x) \geqslant h(y)$ for all states $y \in \mathcal{X}$. Since the Markov chain is irreducible, there exists a state $z \in \mathcal{X}$ such that $p_{xz} > 0$. Let's assume $h(z) < h(x)$, then

$$h(x) = p_{xz} h(z) + \sum_{y \neq z} p_{xy} h(y) < h(x).$$

This implies that $h(x) = h(z)$ for all states z such that $p_{xz} > 0$. By induction, this implies that $h(x) = h(y)$ for any state y reachable from state x. Since all states are reachable from state x by irreducibility, this implies h is a constant on the state space $\mathcal{X}$. □

Corollary 1.7. *For any irreducible and aperiodic finite state Markov chain, there exists a unique invariant distribution π.*

Proof. For an aperiodic and irreducible discrete time Markov chain $X : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$ with finite state space $\mathcal{X}$, we have $\mathbb{E}_x \tau_y^+(1) < \infty$ for all states $x, y \in \mathcal{X}$. In particular, X is positive recurrent and hence there exists a positive invariant distribution $\pi \in \mathcal{M}(\mathcal{X})$. Further, from previous Lemma we have that the dimension of null-space of $(p - I)$ is unity. Hence, the rank of $p - I$ is $|\mathcal{X}| - 1$. Therefore, all vectors satisfying $\nu = \nu p$ are scalar multiples of π. □

1.3 Stationary distribution

Theorem 1.8. *For a finite state irreducible and aperiodic Markov chain $X : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$, its invariant distribution is same as its stationary distribution.*

Proof. For a finite state irreducible and aperiodic discrete time Markov chain $X : \Omega \to \mathcal{X}^{\mathbb{Z}_+}$, we have $\mathbb{E}_x \tau_y^+(1) < \infty$ and hence $P_x \left\{\tau_y^+(1) < \infty\right\} = 1$ for all states $x, y \in \mathcal{X}$. That is, the return times are finite almost surely, and hence we can apply strong Markov property at these stopping times to obtain that discrete time Markov chain X is a regenerative process with delayed renewal sequence $\tau_y^+ : \Omega \to \mathbb{N}^{\mathbb{N}}$, where $\tau_y^+(0) \triangleq 0$, and $\tau_y^+(k) \triangleq \inf\left\{n > \tau_y^+(k-1) : X_n = y\right\}$. We can create an on-off alternating renewal function on this discrete time Markov chain X, which is on when in state y. Then, from the limiting on probability of alternating renewal function, we know that

$$\pi_y \triangleq \lim_{n\to\infty} P_x\{X_n = y\} = \lim_{n\to\infty} \frac{1}{n}\sum_{k=1}^{n} \mathbb{1}_{\{X_k=y\}} = \frac{1}{\mathbb{E}_y \tau_y^+(1)}.$$

We observe that $\pi_y = \frac{\tilde{\pi}_y^y}{\mathbb{E}_y \tau_y^+(1)}$ for each state $y \in \mathcal{X}$. From the uniqueness of invariant distribution, it follows that π is the unique invariant distribution of the discrete time Markov chain X. We observe that π_x is the long-term average of the amount of time spent in state x and from renewal reward theorem $\pi_x = \frac{1}{\mathbb{E}_x \tau_x^+(1)}$. □