Lecture-25: Markov Chains: Invariant Distribution

1 Invariant Distribution

Let $X = (X_n : n \in \mathbb{N}_0)$ be a time-homogeneous Markov chain on state space S with transition probability matrix P. A probability distribution $\pi = (\pi_i \geqslant 0 : i \in S)$ such that $\sum_{i \in S} \pi_i = 1$ is said to be **stationary distribution** or invariant distribution for the Markov chain X if $\pi = \pi P$, that is $\pi_j = \sum_{i \in S} \pi_i P_{ij}$ for all $j \in S$.

Remark 1. Facts about the invariant distribution π.

i. The global balance equation $\pi = \pi P$ is a matrix equation, that is we have a collection of $|S|$ equations $\pi_j = \sum_{i \in S} \pi_i P_{ij}$ for each $j \in S$.

ii. Balance equation across cuts is $\pi_j(1 - P_{jj}) = \pi_j \sum_{i \neq j} P_{ji} = \sum_{i \neq j} \pi_i P_{ij}$.

iii. The invariant distribution π is left eigenvector of stochastic matrix P with the largest eigenvalue 1. The all ones vector is the right eigenvector of this stochastic matrix P for the eigenvalue 1.

iv. From the Chapman-Kolmogorov equation for initial probability vector π, we have $\pi = \pi P^n$ for $n \in \mathbb{N}$. That is, if $P(X_0 = i) = \pi_i$ for each $i \in S$, then $P(X_n = j) = \pi_j$ for each $j \in S$ and all $n \in \mathbb{N}_0$, since $P(X_n = j) = \sum_{i \in S} P(X_0 = i) p_{ij}^{(n)}$.

v. Resulting process with initial distribution π is stationary, and hence have shift-invariant finite dimensional distributions. For example, for any $k, n \in \mathbb{N}$ and $i_1, \dots, i_n \in S$, we have

$$P(X_0 = i_0, \dots, X_n = i_n) = P(X_k = i_0, \dots, X_{k+n} = i_n) = \pi_{i_0} p_{i_0 i_1} \dots p_{i_{n-1} i_n}.$$

vi. If the Markov chain is irreducible, with $\pi_i > 0$ for some $i \in S$. Then for any $j \in S$, we have $p_{ij}^{(m)} > 0$ for some $m \in \mathbb{N}$. Hence, $\pi_j \geqslant \pi_i p_{ij}^{(m)} > 0$. That is, the entire invariant vector is positive.

vii. Any scaled version of π satisfies the global balance equation. Therefore, $\sum_{i \in S} \pi_i$ must be finite for positive recurrent Markov chains, to normalise such vectors and get a unique invariant measure.

Theorem 1.1. *An irreducible Markov chain with transition probability matrix P is positive recurrent iff there exists a unique invariant probability measure π on state space S that satisfies global balance equation $\pi = \pi P$ and $\pi_i = \frac{1}{\mu_{ii}} > 0$ for all $i \in S$.*

Proof. Let X be a positive recurrent Markov chain on state space S, with $X_0 = i$. Let H_i be the first recurrence time to state i, and let $N_j(n) = \sum_{k=1}^{n} 1\{X_k = j\}$ be the number of visits to state $j \in S$ in the first n steps of the Markov chain. It follows that $N_i(H_i) = 1$ and $\sum_{j \in S} N_j(n) = n$ for each $n \in \mathbb{N}$. Taking expectation, we denote $v_j \triangleq \mathbb{E}_i[N_j(H_i)]$ for each $j \in S$. We observe that $v_j \geqslant 0$ for each state $j \in S$, in particular $v_i = 1$, and $\sum_{j \in S} v_j = \mathbb{E}_i H_i = \mu_{ii} < \infty$ since X is positive recurrent.

We will show that the vector $v = (v_i : i \in S)$ satisfies the global balance equations $v = vP$, and since v is summable, $\pi = \frac{v}{\sum_{i \in S} v_i}$ is an invariant distribution for the Markov chain X. To see that the vector v satisfies the global balance equations, we observe from the monotone convergence theorem

$$v_j = \mathbb{E}_i N_j(H_i) = \mathbb{E}_i \sum_{n \in \mathbb{N}} 1\{X_n = j, n \leqslant H_i\} = \sum_{n \in \mathbb{N}} P_i(X_n = j, n \leqslant H_i).$$

Let $\lambda_{ij}^{(n)} \triangleq P_i(X_n = j, n \leqslant H_i)$. Observe that $\lambda_{ij}^{(1)} = p_{ij}$ for each $j \in S$. For $n \geqslant 2$, we have $\lambda_{ij}^{(n)} = \sum_{\ell \neq i} \lambda_{i\ell}^{(n-1)} p_{\ell j}$, and hence we have for each $j \in S$,

$$v_j = p_{ij} + \sum_{n \geqslant 2} \sum_{\ell \neq i} \lambda_{i\ell}^{(n-1)} p_{\ell j} = p_{ij} + \sum_{\ell \neq i} p_{\ell j} \sum_{n \in \mathbb{N}} P_i(X_n = \ell, n \leqslant H_i) = v_i p_{ij} + \sum_{\ell \neq i} v_\ell p_{\ell j} = \sum_{i \in S} v_i p_{ij}.$$

Hence, $\pi = \frac{\nu}{\sum_{i\in S}\nu_i}$ is an invariant measure of the transition matrix P, and $\pi_i = \frac{\nu_i}{\sum_{i\in S}\nu_i} = \frac{1}{\mu_{ii}} > 0$. Next, we show that this is a unique invariant measure independent of the initial state i, and hence $\pi_j = \frac{1}{\mu_{jj}} > 0$ for all $j \in S$. For uniqueness, we observe from the Chapman-Kolmogorov equations and invariance of π that for any $j \in S$

$$\pi_j = \sum_{i\in S}\pi_i \frac{1}{n}\sum_{k=1}^{n} p_{ij}^{(k)}.$$

Taking limit $n \to \infty$ on both sides, and exchanging limit and summation on right hand side using bounded convergence theorem for summable series π, we get for all $j \in S$

$$\pi_j = \frac{1}{\mu_{jj}}\sum_{i\in S}\pi_i = \frac{1}{\mu_{jj}} > 0.$$

Conversely, let π be the positive invariant distribution of Markov chain X. Then, if the Markov chain was transient or null recurrent, we would have $\lim_{n\in\mathbb{N}} \frac{1}{n}\sum_{k=1}^{n} p_{ij}^{(k)} = 0$. Since π is an invariant vector, we get $\pi = \pi P^k$ for each $k \in \mathbb{N}$ and hence $\pi = \pi \frac{1}{n}\sum_{k=1}^{n} P^k$. Taking limit on both sides, we have $\pi = 0$, yielding a contradiction for its positivity. □

Corollary 1.2. *An irreducible Markov chain on a finite state space has a unique and positive stationary distribution π.*

Remark 2. Additional remarks about the stationary distribution π.

i. For a Markov chain with multiple positive recurrent communicating classes $\mathscr{C}_1, \dots, \mathscr{C}_m$, one can find the positive equilibrium distribution for each class, and extend it to the entire state space S denoting it by π_k for class $k \in [m]$. It is easy to check that any convex combination $\pi = \sum_{k=1}^{m} \alpha_m \pi_m$ satisfies the global balance equation $\pi = \pi P$, where $\alpha_k \geqslant 0$ for each $k \in [m]$ and $\sum_{k=1}^{m} \alpha_m = 1$. Hence, a Markov chain with multiple positive recurrent classes have a convex set of invariant probability measures, with the individual invariant distribution π_k for each positive recurrent class $k \in [m]$ are the extreme points.

ii. Let $\nu(0) = e_i$, that is let the initial state of the positive recurrent Markov chain be $X_0 = i$. Then, we know that
$$\pi_j = \frac{1}{\mu_{jj}} = \lim_{n\in\mathbb{N}} \frac{1}{n}\sum_{k=1}^{n} p_{ij}^{(k)} = \lim_{n\in\mathbb{N}} \frac{1}{n}\mathbb{E}_i N_j(n).$$
That is, π_j is limiting average of number of visits to state $j \in S$.

iii. If the positive recurrent Markov chain is aperiodic, then limiting probability of being in a state j is its invariant probability, that is $\pi_j = \lim_{n\in\mathbb{N}} p_{ij}^{(n)}$.

An irreducible, aperiodic, positive recurrent Markov chain is called **ergodic**.

Theorem 1.3. *For an irreducible, aperiodic, positive recurrent Markov chain X with invariant distribution π, and nth step distribution $\nu(n)$, we have $\lim_{n\in\mathbb{N}} \nu(n) = \pi$.*

Proof. Consider independent time homogeneous Markov chains $X = (X_n : n \in \mathbb{N}_0)$ and $Y = (Y_n : n \in \mathbb{N}_0)$ each with transition matrix P. The initial state of Markov chain X is assumed to be $X_0 = i$, whereas the Markov chain Y is assumed to have an initial distribution π. It follows that Y is a stationary process, while X is not. In particular, $\nu_j(n) = P(X_n = j) = p_{ij}^{(n)}$ and $P(Y_n = j) = \pi_j$. Let $\tau = \inf\{n \in \mathbb{N}_0 : X_n = Y_n\}$ be the first time that two Markov chains meet, called the **coupling time**.

First, we show that the coupling time is almost surely finite. To this end, we define a a new Markov chain on state space $S \times S$ with transition probability matrix Q such that $q((i,i'),(j,j')) = p_{ij}p_{i'j'}$ for each $(i,i'),(j,j') \in S \times S$. The n-step transition probabilities are given by $q^{(n)}((i,i'),(j,j')) = p_{ij}^{(n)} p_{i'j'}^{(n)}$. Since the Markov chain X with transition probability matrix P is irreducible and aperiodic, for each $i, j, i', j' \in S$ there exists an $n \in \mathbb{N}_0$ such that $q^{(n)}((i,i'),(j,j')) = p_{ij}^{(n)} p_{i'j'}^{(n)} > 0$ from the previous Lemma. Hence, the irreducibility of this new **product** Markov chain follows. It is easy to check that $\theta(i,i') = \pi_i pi_{i'}$ is the invariant distribution for this product Markov chain, since $\theta(i,i') > 0$ for each $(i,i') \in S \times S$, $\sum_{i,i'\in S}\theta(i,i') = 1$, and for each $(j,j') \in S \times S$, we have

$$\sum_{i,i'\in S}\theta(i,i')q((i,i'),(j,j')) = \sum_{i\in S}\pi_i p_{ij}\sum_{i'\in S}\pi_{i'}p_{i'j'} = \pi_j\pi_{j'} = \theta(j,j').$$

This implies that the product Markov chain is positive recurrent, and each state $(i,i) \in S \times S$ is reachable with unit probability from any initial state $(j,k) \in S \times S$. In particular, the coupling time is almost surely finite.

Second, we show that from the coupling time onwards, the evolution of two Markov chains is identical in distribution. That is, $P(X_n = j, n \geqslant \tau) = P(Y_n = j, n \geqslant \tau)$ for each $j \in S$ and $n \in \mathbb{N}_0$. This follows from the fact that τ is stopping time for the joint process $((X_n, Y_n) : n \in \mathbb{N}_0)$, have identical transition matrix, and that $X_\tau = Y_\tau$.

We can write the difference for any $j \in S$, as

$$|p_{ij}^{(n)} - \pi_j| = |P(X_n = j, n < \tau) - P(Y_n = j, n < \tau)| \leqslant 2P(\tau > n).$$

Since the coupling time is almost surely finite, $\sum_{n \in \mathbb{N}} P(\tau = n) = 1$ and the tail sum $P(\tau > n)$ goes to zero as n grows large, and the result follows. □

Example 1.4 (Single Server Queue).