[AI2613 Lecture 5] Positive Recurrence, SLLN for MC, FT of Countably Infinite Markov Chains

March 28, 2022

1 Recurrence and Positive Recurrence

Recall that we say a state i is *recurrent* if $\mathbf{P}_i[T_i < \infty] = 1$ or equivalently $\mathbf{E}_i[N_i] = \infty$. Otherwise, we say the state is *transient*. A transient state j will be visited for finite times with probability 1. From Proposition 3 of last lecture, we know that *recurrence* is a class property, that is, given a recurrent state i, all the other states that i can reach in finite steps are also recurrent. We are only concerned with irreducible Markov chains in this lecture. So we may say a Markov chain is recurrent or transient in the future.

Here we follow the notations of the last lecture, that is: $X_0, X_1, \dots, X_t, \dots$ is a sequence of variables that follows the Markov chain P. $T_i \triangleq \inf\{t > 0 : X_t = i\}$, $N_i \triangleq \sum_{t=0}^{\infty} \mathbb{1}[X_t = i]$, $\mathbf{P}_i[\cdot] = \mathbf{Pr}[\cdot | X_0 = i]$ and $\mathbf{E}_i[\cdot] = \mathbf{E}[\cdot | X_0 = i]$.

Example 1 (Drunk person and drunk bird) *Imagine a random walk on a grid that we pick a direction uniformly at random at each time step. Can we go back to the original point with probability* 1*? Or equivalently, is this Markov chain recurrent or transient?*

First we consider the one-dimensional grid. Let $X_0 = 0$ *and* $X_{t+1} = X_t + \Delta$ *where* Δ *is uniformly at random picked from* $\{-1, 1\}$*. Then,*

$$\mathbf{E}_0[N_0] = \mathbf{E}_0\left[\sum_{t=0}^{\infty} \mathbb{1}[X_t = 0]\right] = \sum_{t=0}^{\infty} \mathbf{P}_0[X_t = 0] = \sum_{m=0}^{\infty} \mathbf{P}_0[X_{2m} = 0].$$

where the last equality follows from the fact that we can not go back within exactly odd steps. Then let's compute $\mathbf{P}_0[X_{2m} = 0]$ *using the Stirling's formula. For* $m \geq 1$,

Stirling's formula: $n! = \sqrt{2\pi n}\left(\frac{n}{e}\right)^n (1 + o(1))$.

$$\mathbf{P}_0[X_{2m} = 0] = \frac{\binom{2m}{m}}{2^{2m}} \approx \frac{\sqrt{4\pi m}\left(\frac{2m}{e}\right)^{2m}}{2\pi m\left(\frac{m}{e}\right)^{2m}} \cdot 2^{-2m} = \frac{1}{\sqrt{\pi m}}.$$

Thus, $\mathbf{E}_0[N_0] = \sum_{m=0}^{\infty} \mathbf{P}_0[X_{2m} = 0] \approx 1 + \sum_{m=1}^{\infty} \frac{1}{\sqrt{\pi m}}$ *which is divergent. This indicates that the Markov chain for random walk on one-dimensional grid is recurrent.*

For d*-dimensional grid, we regard the game as independently pick* Δ_i *u.a.r. from* $\{-1, 1\}$ *for* $i \in [d]$ *at each time step and walk to* $X_{t+1} = X_t + (\Delta_1, \Delta_2, \dots, \Delta_d)$*. So we have that* $\mathbf{P}_i[X_{2m} = \mathbf{0}] = (\mathbf{P}_i[X_{2m}(1) = 0])^d \approx \left(\frac{1}{\sqrt{\pi m}}\right)^d$*. We know that* $1 + \sum_{m=1}^{\infty}\left(\frac{1}{\sqrt{\pi m}}\right)^d$ *is divergent if and only if* $d \leq 2$*. Thus, only if the dimension of the grid is* 1 *or* 2*, the random walk is recurrent.*

Definition 1 (Positive recurrence) *If a state* i *is recurrent and* $\mathbf{E}_i[T_i] < \infty$*, we say it is* positive recurrent*. If the state is recurrent but with* $\mathbf{E}_i[T_i] = \infty$*, then we say it is* null recurrent.

Now we give some examples to distinguish the concept of null recurrence and positive recurrence.

Example 2 (Drunk person) *We have proved that the Markov chain of drunk person is recurrent. We can further verify that it is null recurrent by noticing that*

$$\mathbf{E}_i[T_i] = \sum_{t=1}^{\infty} \mathbf{P}_i[\mathbb{1}[X_t = i]] \cdot t \approx \sum_{m=1}^{\infty} \left(\frac{1}{\sqrt{\pi m}}\right)^2 \cdot 2m = \infty.$$

Example 3 *Recall the random walk on $\mathbb{N}$ we talked about in the last lecture: at each state i, go to $i+1$ w.p. p and go to $i-1$ w.p $1-p$ (if $i=0$, w.p. $1-p$ stay at 0). Then the following statements hold:*

- *When $p > \frac{1}{2}$, the Markov chain is transient.*
- *When $p = \frac{1}{2}$, it is null recurrent.*
- *When $p < \frac{1}{2}$, it is positive recurrent.*

2 Laws of Large Numbers

$X_1, X_2, \dots$ is an infinite sequence of independent and identically distributed Lebesgue integrable random variables with expected value $\mathbf{E}[X_1] = \mathbf{E}[X_2] = \cdots = \mu < \infty$. Let $\bar{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ be the sample average. Then we have the following two laws of large numbers.

Theorem 2 (Weak law of large numbers or Khinchin's law) *The sample average converge in probability towards the expected value:*

$$\bar{X}_n \xrightarrow{p} \mu \quad \text{when } n \to \infty.$$

That is, for any positive value ε,

$$\lim_{n\to\infty} \mathbf{Pr}\left[\left|\bar{X}_n - \mu\right| < \epsilon\right] = 1.$$

Theorem 3 (Strong law of large numbers or Kolmogorov's law) *The sample average converges almost surely or with probability 1 to the expected value:*

$$\bar{X}_n \xrightarrow{\text{a.s.}} \mu \quad \text{when } n \to \infty.$$

That is,

$$\mathbf{Pr}\left[\lim_{n\to\infty} \bar{X}_n \to \mu\right] = 1.$$

Let $(\Omega, \mathcal{F}, P)$ be the probability space. Here $\bar{X}_n \to \mu$ means $\exists M \in \mathcal{F}$ satisifying

- P(M)=1;
- $\forall \omega \in M, \bar{X}_n(\omega) \xrightarrow{n\to\infty} \mu$.

As the name of the laws shows, *convergence in probability* is weaker than *convergence with probability* 1. Consider a sequence of independent random variables $X_1, X_2, \dots$ that X_n is 1 with probability $\frac{1}{n}$ and X_n is 0 with probability $1 - \frac{1}{n}$. Then the sequence converges to 0 in probability but not with probability 1 since we cannot find an $M \in \mathcal{F}$ with measure 1 such that $\bar{X}_n(\omega) \xrightarrow{n\to\infty} 0$ for every $\omega \in M$.

Theorem 4 (Strong law of large numbers for Markov chains) *If there is a finite path from state i to j, then*

$$\mathbf{P}_i\left[\lim_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n}\mathbb{1}[X_t=j]=\frac{1}{\mathbf{E}_j[T_j]}\right]=1.$$

Proof. If j is transient, then the random process will visit j for finite times with probability 1. Thus $\mathbf{P}_i\left[\lim_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n}\mathbb{1}[X_t=j]=\frac{1}{\mathbf{E}_j[T_j]}\right]=\mathbf{P}_i\left[\lim_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n}\mathbb{1}[X_t=j]=0\right]=1$.

If j is recurrent, we first prove the theorem for $i=j$. We call a loop from j to j a cycle (we visit j only at the beginning and end of the loop). Denote C_r as the length of the r^{th} cycle during the process. Let $S_k=\sum_{r=1}^{k}C_r$. Let k_n be the number of cycles before the $n+1$ step, that is, $k_n=\max\{k|S_k\le n\}$. Then we have $S_{k_n}\le n<S_{k_n+1}$ and consequently $\frac{S_{k_n}}{k_n}\le\frac{n}{k_n}<\frac{S_{k_n+1}}{k_n}$. Note that with probability 1, $k_n\to\infty$ when $n\to\infty$. We have with probability 1 that

$$\lim_{k\to\infty}\frac{S_k}{k}\le\lim_{n\to\infty}\frac{n}{k_n}<\lim_{k\to\infty}\frac{S_{k+1}}{k}.$$

Note that $S_k=\sum_{r=1}^{k}C_r$ where each C_r is an i.i.d random variable with mean $\mathbf{E}_j\left[T_j\right]$. So by SLLN (Theorem 3), we have $\lim_{k\to\infty}\frac{S_k}{k}=\mathbf{E}_j\left[T_j\right]$ and $\lim_{k\to\infty}\frac{S_{k+1}}{k}=\lim_{k\to\infty}\frac{S_{k+1}}{k+1}\cdot\frac{k+1}{k}=\mathbf{E}_j\left[T_j\right]$. As a result, with probability 1,

$$\mathbf{E}_j\left[T_j\right]=\lim_{n\to\infty}\frac{n}{k_n}=\lim_{n\to\infty}\frac{n}{\sum_{t=1}^{n}\mathbb{1}[X_t=j]}.$$

If j is recurrent and $i\ne j$, let $T_{i\to j}$ be the first time visiting j. Then we have $\frac{S_{k_n}+T_{i\to j}}{k_n}\le\frac{n}{k_n}<\frac{S_{k_n+1}+T_{i\to j}}{k_n}$. Since $\mathbf{P}_i[T_j<\infty]=1$, $\mathbf{P}_i[\lim_{k\to\infty}\frac{T_{i\to j}}{k}=0]=1$. The remaining proof is the same with the situation that $i=j$. □

Corollary 5 *Let P be the transition function of an irreducible Markov chain where $P^t(i,j)=\mathbf{Pr}[X_t=j|X_0=i]$. Then for any states i,j,*

$$\lim_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n}P^t(i,j)=\frac{1}{\mathbf{E}_j[T_j]}.$$

Proof. By the strong law of large numbers for Markov chains, there exists a set $M\in\mathcal{F}$ such that $P(M)=1$ and $\lim_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n}\mathbb{1}[X_t(\omega)=j]=\frac{1}{\mathbf{E}_j[T_j]}$ for any $\omega\in M$. Then,

$$\begin{aligned}\lim_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n}P^t(i,j)&=\lim_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n}\mathbf{E}_i\left[\mathbb{1}[X_t=j]\right]\\&=\lim_{n\to\infty}\mathbf{E}_i\left[\frac{1}{n}\sum_{t=1}^{n}\mathbb{1}[X_t=j]\right]\\&=\mathbf{E}_i\left[\lim_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n}\mathbb{1}[X_t=j]\right]\\&=\frac{1}{\mathbf{E}_j[T_j]},\end{aligned}$$

where the third equation follows from the bounded convergence theorem. □

Bounded Convergence Theorem: If $X_n\xrightarrow{a.s.}X$ and $\mathbf{E}[X]<\infty$, then $\mathbf{E}[X_n]\to\mathbf{E}[X]$.

3 *Fundamental Theorem*

First we introduce some abbreviations to simplify the expression:

- Aperiodicity:[A],
- Irreducibility:[I],
- Recurrence:[R],
- Positive Recurrence: [PR],
- Has a stationary distribution:[S],
- Has a unique stationary distribution:[U],
- Convergence:[C],
- Finiteness:[F].

The finite FTMC can be written as: [F]+[A]+[I]⇒[S]+[U]+[C]. For infinite Markov chains, the theorem need to be modified as: [PR]+[A]+[I]⇒[S]+[U]+[C]. We will first prove the existence and uniqueness of the stationary distribution in this lecture.(i.e. [S] and [U])

Theorem 6 *[I]+[PR]⇒[S]+[U].*

Proof. [Proof of [U]] Let $\mathcal{S}$ be the set of states. Assume π is a stationary distribution of the Markov chain, i.e.,

$$\forall j \in \mathcal{S},\ \forall t \geq 0,\ \sum_{i \in \mathcal{S}} \pi(i) P^t(i,j) = \pi(j).$$

This yields that for $n \geq 1$,

$$\frac{1}{n} \sum_{i \in \mathcal{S}} \pi(i) \sum_{t=1}^{n} P^t(i,j) = \pi(j).$$

Taking $n \to \infty$ and applying the dominated convergence theorem, we have

Dominated Convergence Theorem: If $\int_S |f_n| < \infty$, then $\lim_{n\to\infty} \int_S f_n = \int_S \lim_{n\to\infty} f_n$.

$$\pi(j) = \sum_{i \in \mathcal{S}} \pi(i) \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} P^t(i,j) = \sum_{i \in \mathcal{S}} \pi(i) \cdot \frac{1}{\mathbf{E}_j[T_j]} = \frac{1}{\mathbf{E}_j[T_j]}.$$

□

Proof. [Proof of [S]] Then we prove the above π is a stationary distribution.

$\mathcal{S}$ is finite. We first assume $\mathcal{S}$ is finite, so that we can safely exchange the order of taking limitation and summation in the calculations below.

$$\sum_{j \in \mathcal{S}} \pi(j) = \sum_{j \in \mathcal{S}} \frac{1}{\mathbf{E}_j[T_j]} = \sum_{j \in \mathcal{S}} \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} P^t(i,j)$$
$$= \lim_{n\to\infty} \sum_{j \in \mathcal{S}} \frac{1}{n} \sum_{t=1}^{n} P^t(i,j) = \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} \sum_{j \in \mathcal{S}} P^t(i,j) = 1.$$

This indicates that π is a legal distribution. We then verify that π is indeed the stationary distribution.

Note that $P^{t+1}(i,j) = \sum_{k\in\mathcal{S}} P^t(i,k)P(k,j)$. Then

$$\begin{aligned}\frac{1}{\mathbf{E}_j[T_j]} &= \lim_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n} P^t(i,j) = \lim_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n} P^{t+1}(i,j)\\ &= \lim_{n\to\infty}\sum_{k\in\mathcal{S}} P(k,j)\frac{1}{n}\sum_{t=1}^{n} P^t(i,k) = \sum_{k\in\mathcal{S}} P(k,j)\lim_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n} P^t(i,k)\\ &= \sum_{k\in\mathcal{S}} P(k,j)\cdot\frac{1}{\mathbf{E}_k[T_k]}.\end{aligned}$$

That is,

$$\pi(j) = \sum_{k\in\mathcal{S}} P(k,j)\pi(k).$$

It is worth noting that [PR] is equivalent to [I] when $\mathcal{S}$ is finite.

$\mathcal{S}$ is infinite. When $\mathcal{S}$ is (countably) infinite, we consider every finite subset A of $\mathcal{S}$. Then

$$\begin{aligned}\sum_{j\in A}\pi(j) = \sum_{j\in A}\frac{1}{\mathbf{E}_j[T_j]} &= \sum_{j\in A}\lim_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n} P^t(i,j)\\ &= \lim_{n\to\infty}\sum_{j\in A}\frac{1}{n}\sum_{t=1}^{n} P^t(i,j) = \lim_{n\to\infty}\frac{1}{n}\sum_{t=1}^{n}\sum_{j\in A} P^t(i,j) < 1.\end{aligned}$$

Therefore

$$\sum_{j\in\mathcal{S}}\pi(j) = \sup_{\text{finite } A\subseteq\mathcal{S}}\sum_{j\in A}\pi(j) =: C \le 1.$$

Since $[PR]$, we know that $C \neq 0$. In the following, we will prove that π/C is a stationary distribution. Then $C = 1$ follows from the uniqueness of the stationary distribution we just proved.

For every finite $A \subseteq \mathcal{S}$, we have

$$\sum_{k\in A} P(k,j)\cdot\frac{1}{\mathbf{E}_k\left[T_k\right]} \le \frac{1}{\mathbf{E}_j\left[T_j\right]}.$$

Therefore,

$$\sum_{k\in\mathcal{S}} P(k,j)\cdot\frac{1}{\mathbf{E}_k\left[T_k\right]} = \sup_{\text{finite } A\subseteq\mathcal{S}}\sum_{k\in A} P(k,j)\cdot\frac{1}{\mathbf{E}_k\left[T_k\right]} \le \frac{1}{\mathbf{E}_j\left[T_j\right]}.$$

We show that indeed the equality holds. Assume for a contradiction that

$$\sum_{k\in\mathcal{S}} P(k,j)\cdot\frac{1}{\mathbf{E}_k\left[T_k\right]} < \frac{1}{\mathbf{E}_j\left[T_j\right]}.$$

Summing the both sides over all $j \in \mathcal{S}$, we obtain

$$\sum_{k\in\mathcal{S}}\frac{1}{\mathbf{E}_k\left[T_k\right]} < \sum_{j\in\mathcal{S}}\frac{1}{\mathbf{E}_j\left[T_j\right]},$$

which is a contradiction. As a result, we know

$$\sum_{k \in \mathcal{S}} P(k, j) \cdot \frac{1}{\mathbf{E}_k\left[T_k\right]} = \frac{1}{\mathbf{E}_j\left[T_j\right]},$$

and $\hat{\pi}(j) = \frac{1}{C \cdot \mathbf{E}_j\left[T_j\right]}$ is a stationary distribution. By the uniqueness of the distribution, we have $C = 1$.

□