Markov Chains II

Contents

1. Markov chains with a single recurrence class
2. Uniqueness of Stationary Distributions

1 Markov chains with a single recurrence class

Recall the relations $\to, \leftrightarrow$ introduced in the previous lecture for the class of finite state Markov chains. Recall that we defined a state $i$ to be recurrent if whenever $i \to j$ we also have $j \to i$, namely $i \leftrightarrow j$. We have observed that $\leftrightarrow$ is an equivalency relation, so that set of recurrent states is partitioned into equivalency classes $R_1, \ldots, R_r$. The remaining states $T$ are transient.

**Lemma 1.** For every $l = 1, \ldots, r$ and every $i \in R_l, j \notin R_l$ we must have $p_{i,j} = 0$.

This means that once the chain is in some recurrent class $R$ it stays there forever.

**Proof.** The proof is simple: $p_{i,j} > 0$ implies $i \to j$. Since $i$ is recurrent then also $j \to i$ implying $j \in R$ - contradiction. 

Introduce the following basic random quantities. Given states $i, j$ let

$$T_i = \min\{n \geq 1 : X_n = i | X_0 = i\}.$$ 

In case no such $n$ exists, we set $T_i = \infty$. Thus the range of $T_i$ is $\mathcal{N} \cup \{\infty\}$. The quantity is called the *first passage time*. Let $\mu_i = \mathbb{E}[T_i]$, possibly with $\mu_i = \infty$. This is called mean recurrence time of the state $i$. 


Lemma 2. For every state \( i \in \mathcal{T} \), \( \mathbb{P}(X_n = i, \text{ i.o.}) = 0 \). Namely, almost surely, after some finite time \( n_0 \), the chain will never return to \( i \). In addition \( \mathbb{E}[T_i] = \infty \).

Proof. By definition there exists a state \( j \) such that \( i \to j \), but \( j \not\to i \). It then follows that \( \mathbb{P}(T_i = \infty) > 0 \) implying \( \mathbb{E}[T_i] = \infty \). Now, let us establish the first part.

Let \( I_{i,m} \) be the indicator of the event that the M.c. returned to state \( i \) at least \( m \) times. Notice that \( \mathbb{P}(I_{i,1}) = \mathbb{P}(T_i < \infty) < 1 \). Also by M.c. property we have \( \mathbb{P}(I_{i,m} \mid I_{i,m-1}) = \mathbb{P}(T_i < \infty) \), as conditioning that at some point the M.c. returned to state \( i \) \( m-1 \) times does not impact its likelihood to return to this state again. Also notice \( I_{i,m} \subset I_{i,m-1} \). Thus \( \mathbb{P}(I_{i,m}) = \mathbb{P}(I_{i,m} \mid I_{i,m-1}) \mathbb{P}(I_{i,m-1}) = \mathbb{P}(T_i < \infty) \mathbb{P}(I_{i,m-1}) = \cdots = \mathbb{P}^m(T_i < \infty) \). Since \( \mathbb{P}(T_i < \infty) < 1 \), then by continuity of probability property we obtain \( \mathbb{P}(\cap_m I_{i,m}) = \lim_{m \to \infty} \mathbb{P}(I_{i,m}) = \lim_{m \to \infty} \mathbb{P}^m(T_i < \infty) = 0 \). Notice that the event \( \bigcap_m I_{i,m} \) is precisely the event \( X_n = i, \text{ i.o.} \).

Exercise 1. Show that \( \mathcal{T} \neq \mathcal{X} \). Namely, in every finite state M.c. there exists at least one recurrent state.

Exercise 2. Let \( i \in \mathcal{T} \) and let \( \pi \) be an arbitrary stationary distribution. Establish that \( \pi_i = 0 \).

Exercise 3. Suppose M.c. has one recurrent class \( \mathcal{R} \). Show that for every \( i \in \mathcal{R} \) \( \mathbb{P}(X_n = i, \text{ i.o.}) = 1 \). Moreover, show that there exists \( 0 < q < 1 \) and \( C > 0 \) such that \( \mathbb{P}(T_i > t) \leq Cq^t \) for all \( t \geq 0 \). As a result, show that \( \mathbb{E}[T_i] < \infty \).

We now focus on the family of Markov chains with only one recurrent class. Namely \( \mathcal{X} = \mathcal{T} \cup \mathcal{R} \). If in addition \( \mathcal{T} = \emptyset \), then such a M.c. is called irreducible.

2 Uniqueness of the stationary distribution

We now establish a fundamental result on M.c. with a single recurrence class.

Theorem 1. A finite state M.c. with a single recurrence class has a unique stationary distribution \( \pi \), which is given as \( \pi_i = \frac{1}{\mu_i} \) for all states \( i \). Specifically, \( \pi_i > 0 \) iff the state \( i \) is recurrent.
Proof. Let $P$ be the transition matrix of the chain. We let the state space be $\mathcal{X} = \{1, \ldots, N\}$. We fix an arbitrary recurrent state $k$. We know that one exists by Exercise 1. Assume $X_0 = k$. Let $N_i$ be the number of visits to state $i$ between two successive visits to state $k$. In case $i = k$, the last visit is counted but the initial is not. Namely, in the special case $i = k$ the number of visits is $1$ with probability one. Let $\hat{\rho}_i(k) = \mathbb{E}[N_i]$. Consider the event $\{X_n = i, T_k \geq n\}$ and consider the indicator function $I_{X_n = i, T_k \geq n}$ = $\sum_{1 \leq n \leq T_k} I_{X_n = i}$. Notice that this sum is precisely $N_i$. Namely, $\hat{\rho}_i(k) = \sum_{n \geq 1} \mathbb{P}(X_n = i, T_k \geq n | X_0 = k).$ (1)

Then using the formula $\mathbb{E}[Z] = \sum_{n \geq 1} \mathbb{P}(Z \geq n)$ for integer valued r.v., we obtain

$$\sum_i \rho_i(k) = \sum_{n \geq 1} \mathbb{P}(T_k \geq n | X_0 = k) = \mathbb{E}[T_k] = \mu_k. \quad (2)$$

Since $k$ is recurrent, then by Exercise 3, $\mu_k < \infty$ implying $\rho_i(k) < \infty$. We let $\rho(k)$ denote the vector with components $\rho_i(k)$.

**Lemma 3.** $\rho(k)$ satisfies $\rho^T(k) = \rho^T(k) P$. In particular, for every recurrent state $k$, $\pi_i = \frac{\rho_i(k)}{\mu_k}, 1 \leq i \leq N$ defines a stationary distribution.

Proof. The second part follows from (2) and the fact that $\mu_k < \infty$. Now we prove the first part. We have for every $n \geq 2$

$$\mathbb{P}(X_n = i, T_k \geq n | X_0 = k) = \sum_{j \neq k} \mathbb{P}(X_n = i, X_{n-1} = j, T_k \geq n | X_0 = k)$$

$$= \sum_{j \neq k} \mathbb{P}(X_{n-1} = j, T_k \geq n - 1 | X_0 = k)p_{j,i} \quad (3)$$

Observe that $\mathbb{P}(X_1 = i, T_k \geq 1 | X_0 = k) = p_{k,i}$. We now sum the (3) over $n$ and apply it to (1) to obtain

$$\rho_i(k) = p_{k,i} + \sum_{j \neq k} \sum_{n \geq 2} \mathbb{P}(X_{n-1} = j, T_k \geq n - 1 | X_0 = k)p_{j,i}$$

3
We recognize \( \sum_{n \geq 2} \mathbb{P}(X_{n-1} = j, T_k \geq n) = 0 \) as \( \rho_j(k) \). Using \( \rho_k(k) = 1 \) we obtain

\[
\rho_i(k) = \rho_k(k)p_{k,i} + \sum_{j \neq k} \rho_j(k)p_{j,i} = \sum_j \rho_j(k)p_{j,i}
\]

which is in vector form precisely \( \hat{\rho}^T(k) = \rho^T(k)P \). \( \square \)

We now return to the proof of the theorem. Let \( \pi \) denote an arbitrary stationary distribution of our M.c. We know one exists by Lemma 3 and, independently by our linear programming based proof. By Exercise 2 we already know that \( \pi_i = 1/\mu_i = 0 \) for every transient state \( i \).

We now show that in must be that \( \pi_k = 1/\mu_k \) for every recurrent state \( k \). In particular, the stationary distribution is unique. Assume that at time zero we start with distribution \( \hat{\pi} \). Namely \( \mathbb{P}(X_0 = i) = \pi_i \) for all \( i \). Of course this implies that \( \mathbb{P}(X_n = i) \) is also \( \pi_i \) for all \( n \). On the other hand, fix any recurrent state \( k \) and consider

\[
\mu_k \pi_k = \mathbb{E}[T_k | X_0 = k] \mathbb{P}(X_0 = k)
\]

\[
= \sum_{n \geq 1} \mathbb{P}(T_k \geq n | X_0 = k) \mathbb{P}(X_0 = k)
\]

\[
= \sum_{n \geq 1} \mathbb{P}(T_k \geq n, X_0 = k).
\]

On the other hand \( \mathbb{P}(T_k \geq 1, X_0 = k) = \mathbb{P}(X_0 = k) \) and for \( n \geq 2 \)

\[
\mathbb{P}(T_k \geq n, X_0 = k) = \mathbb{P}(X_0 = k, X_j \neq k, 1 \leq j \leq n - 1)
\]

\[
= \mathbb{P}(X_j \neq k, 1 \leq j \leq n - 1) - \mathbb{P}(X_j \neq k, 0 \leq j \leq n - 1)
\]

\[
= \mathbb{P}(X_j \neq k, 0 \leq j \leq n - 2) - \mathbb{P}(X_j \neq k, 0 \leq j \leq n - 1)
\]

\[
= a_{n-2} - a_{n-1},
\]

where \( a_n = \mathbb{P}(X_j \neq k, 0 \leq j \leq n) \) and (*) follows from stationarity of \( \pi \). Now \( a_0 = \mathbb{P}(X_0 \neq k) \). Putting together, we obtain

\[
\mu_k \pi_k = \mathbb{P}(X_0 = k) + \sum_{n \geq 2} (a_{n-2} - a_{n-1})
\]

\[
= \mathbb{P}(X_0 = k) + \mathbb{P}(X_0 \neq k) - \lim_n a_n
\]

\[
= 1 - \lim_n a_n
\]
But by continuity of probabilities \( \lim_{n} a_n = \mathbb{P}(X_n \neq k, \forall n) \). By Exercise 3, the state \( k \), being recurrent is visited infinitely often with probability one. We conclude that \( \lim_{n} a_n = 0 \), which gives \( \mu_k \pi_k = 1 \), implying that \( \pi_k \) is uniquely defined as \( 1/\mu_k \).

3 Ergodic theorem

Let \( N_i(t) \) denote the number of times the state \( i \) is visited during the times \( 0, 1, \ldots, t \). What can be said about the behavior of \( N_i(t)/t \) when \( t \) is large? The answer turns out to be very simple: it is \( \pi_i \). These type of results are called ergodic properties, as they show how the time average of the system, namely \( \bar{N}_i(t)/t \) relates to the spatial average, namely \( \pi_i \).

**Theorem 2.** For arbitrary starting state \( X_0 = k \) and for every state \( i \),
\[
\lim_{t \to \infty} \frac{N_i(t)}{t} = \pi_i
\]
almost surely. Also
\[
\lim_{t \to \infty} \frac{\mathbb{E}[N_i(t)]}{t} = \pi_i.
\]

**Proof.** Suppose \( X_0 = k \). If \( i \) is a transient state, then, as we have established, almost surely after some finite time, the chain will never enter \( i \), meaning \( \lim_{t} N_i(t)/t = 0 \) almost surely. Since also \( \pi_i = 0 \), then we have established the required equality for the case when \( i \) is a transient state.

Suppose now \( i \) is a recurrent state. Let \( T_1, T_2, T_3, \ldots \) denote the time of successive visits to \( i \). Then the sequence \( T_n, n \geq 2 \) is i.i.d. Also \( T_1 \) is independent from the rest of the sequence, although it distribution is different from the one of \( T_m, m \geq 2 \) since we have started the chain from \( k \) which is in general different from \( i \). By the definition of \( N_i(t) \) we have
\[
\sum_{1 \leq m \leq N_i(t)} T_m \leq t < \sum_{1 \leq m \leq N_i(t)+1} T_m
\]
from which we obtain
\[
\frac{\sum_{1 \leq m \leq N_i(t)} T_m}{N_i(t)} \leq \frac{t}{N_i(t)} < \frac{\sum_{1 \leq m \leq N_i(t)+1} T_m}{N_i(t)+1} \frac{N_i(t)+1}{N_i(t)}. \quad (5)
\]
We know from Exercise 3 that $\mathbb{E}[T_m] < \infty, m \geq 2$. Using a similar approach it can be shown that $\mathbb{E}[T_1] < \infty$, in particular $T_1 < \infty$ a.s. Applying SLLN we have that almost surely

$$\lim_{n \to \infty} \frac{\sum_{2 \leq m \leq n} T_m}{n} = \lim_{n \to \infty} \frac{\sum_{2 \leq m \leq n} T_m}{n - 1} = \mathbb{E}[T_2]$$

which further implies

$$\lim_{n \to \infty} \frac{\sum_{1 \leq m \leq n} T_m}{n} = \lim_{n \to \infty} \frac{\sum_{2 \leq m \leq n} T_m}{n} + \lim_{n \to \infty} \frac{T_1}{n} = \mathbb{E}[T_2]$$

almost surely.

Since $i$ is a recurrent state then by Exercise 3, $N_i(t) \to \infty$ almost surely as $t \to \infty$. Combining the preceding identity with (5) we obtain

$$\lim_{t \to \infty} \frac{t}{N_i(t)} = \mathbb{E}[T_2] = \mu_i,$$

from which we obtain $\lim_t N_i(t)/t = \mu_i^{-1} = \pi_i$ almost surely.

To establish the convergence in expectation, notice that $N_i(t) \leq t$ almost surely, implying $N_i(t)/t \leq 1$. Applying bounded convergence theorem, we obtain that $\lim_t \mathbb{E}[N_i(t)]/t = \pi_i$, and the proof is complete.

4 Markov chains with multiple recurrence classes

How does the theory extend to the case when the M.c. has several recurrence classes $R_1, \ldots, R_r$? The summary of the theory is as follows (the proofs are very similar to the case of single recurrent class case and is omitted). It turns out that such a M.c. chain possesses $r$ stationary distributions $\pi^i = (\pi^i_1, \ldots, \pi^i_N), 1 \leq i \leq r$, each “concentrating” on the class $R_i$. Namely for each $i$ and each state $k \notin R_i$ we have $\pi^i_k = 0$. The $i$-th stationary distribution is described by $\pi^i_k = 1/\mu_k$ for all $k \in R_i$ and where $\mu_k$ is the mean return time from state $k \in R_j$ into itself. Intuitively, the stationary distribution $\pi^i$ corresponds to the case when the M.c. “lives” entirely in the class $R_i$. One can prove that the family of all of the stationary distributions of such a M.c. can be obtained by taking all possible convex combinations of $\pi^i, 1 \leq i \leq r$, but we omit the proof. (Exercise: show that a convex combination of stationary distributions is a stationary distribution).
References

