Fundamentals of probability.
6.436/15.085

LECTURE 25
Markov chains III. Periodicity, Mixing, Absorption

25.1. Periodicity

Previously we showed that when a finite state M.c. has only one recurrent class and π is the corresponding stationary distribution, then $\mathbb{E}[N_i(t)|X_0 = k]/t \to \pi_i$ as $t \to \infty$, irrespective of the starting state k. Since $N_i(t) = \sum_{n=1}^{t} \mathbf{1}_{\{X_n=i\}}$ is the number of times state i is visited up till time t, we have shown that $\frac{1}{t}\sum_{n=1}^{t} \mathbb{P}(X_n = i|X_0 = k) \to \pi_i$ for every state k, i.e., $p_{ki}^{(n)}$ converges to π_i in the Cesaro sense. However, $p_{ki}^{(n)}$ need not converge, as the following example shows. Consider a 2 state Markov Chain with states $\{1, 2\}$ and $p_{12} = 1 = p_{21}$. Then $p_{12}^{(n)} = 1$ when n is odd and 0 when n is even.

Let x be a recurrent state and consider all the times when x is accessible from itself, i.e., the times in the set $I_x = \{n \geq 1 : p_{xx}^{(n)} > 0\}$ (note that this set is non-empty since x is a recurrent state). One property of I_x we will make use of is that it is closed under addition, i.e., if $m, n \in I_x$, then $m + n \in I_x$. This is easily seen by observing that $p_{xx}^{(m+n)} \geq p_{xx}^{(m)} p_{xx}^{(n)} > 0$. Let d_x be the greatest common divisor of the numbers in I_x. We call d_x the *period* of x. We now show that all states in the same recurrent class has the same period.

Lemma 25.1. *If x and y are in the same recurrent class, then $d_x = d_y$.*

Proof. Let m and n be such that $p_{xy}^{(m)}, p_{yx}^{(n)} > 0$. Then $p_{yy}^{(m+n)} \geq p_{xy}^{(m)} p_{yx}^{(n)} > 0$. So d_y divides $m+n$. Let l be such that $p_{xx}^{(l)} > 0$, then $p_{yy}^{(m+n+l)} \geq p_{yx}^{(n)} p_{xx}^{(l)} p_{xy}^{(m)} > 0$. Therefore d_y divides $m+n+l$, hence it divides l. This implies that d_y divides d_x. A similar argument shows that d_x divides d_y, so $d_x = d_y$. □

A recurrent class is said to be *periodic* if the period d is greater than 1 and *aperiodic* if $d = 1$. The 2 state Markov Chain in the example above has a period of 2 since $p_{11}^{(n)} > 0$ iff n is even. A recurrent class with period d can be divided into d subsets, so that all transitions from one subset lead to the next subset.

Why is periodicity of interest to us? It is because periodicity is exactly what prevents the convergence of $p_{xy}^{(n)}$ to π_y. Suppose y is a recurrent state with period $d > 1$. Then $p_{yy}^{(n)} = 0$ unless n is a multiple of d, but $\pi_y > 0$. However, if $d = 1$, we have positive probability of returning to y for all time steps n sufficiently large.

Lemma 25.2. *If $d_y = 1$, then there exists some $N \geq 1$ such that $p_{yy}^{(n)} > 0$ for all $n \geq N$.*

Proof. We first show that $I_y = \{n \geq 1 : p_{yy}^{(n)} > 0\}$ contains two consecutive integers. Let n and $n + k$ be elements of I_y. If $k = 1$, then we are done. If not, then since $d_y = 1$, we can find a $n_1 \in I_y$ such that k is not a divisor of n_1. Let $n_1 = mk + r$ where $0 < r < k$. Consider $(m + 1)(n + k)$ and $(m + 1)n + n_1$, which are both in I_y since I_y is closed under addition. We have

$$(m + 1)(n + k) - (m + 1)n + n_1 = k - r < k.$$

So by repeating the above argument at most k times, we eventually obtain a pair of consecutive integers $m, m + 1 \in I_y$. If $N = m^2$, then for all $n \geq N$, we have $n - N = km + r$, where $0 \leq r < m$. Then $n = m^2 + km + r = r(1 + m) + (m - r + k)m \in I_y$. □

When a Markov chain has one recurrent class (irreducible) and aperiodic, we have that the steady state behavior is given by the stationary distribution. This is also known as *mixing*.

Theorem 25.3. *Consider an irreducible, aperiodic Markov chain. Then for all states x, y, $\lim_{n\to\infty} p_{xy}^{(n)} = \pi_y$.*

For the case of periodic chains, there is a similar statement regarding convergence of $p_{xy}^{(n)}$, but now the convergence holds only for certain subsequences of the time index n. See [**1**] for further details.

There are at least two generic ways to prove this theorem. One is based on the Perron-Frobenius Theorem which characterizes eigenvalues and eigenvectors of non-negative matrices. Specifically the largest eigenvalue of P is equal to unity and all other eigenvalues are strictly smaller than unity in absolute value. The P-F Theorem is especially useful in the special case of so-called *reversible* M.c.. These are irreducible M.c. for which the unique stationary distribution satisfies $\pi_x p_{xy} = \pi_y p_{yx}$ for all states x, y. Then the following important refinement of Theorem 25.4 is known.

Theorem 25.4. *Consider an irreducible aperiodic Markov chain which is reversible. Then there exists a constant C such that for all states x, y, $|p_{xy}^{(n)} - \pi_y| \leq C|\lambda_2|^n$, where λ_2 is the second largest (in absolute value) eigenvalue of P.*

Since by P-F Theorem $|\lambda_2| < 1$, this theorem is indeed a refinement of Theorem 25.4 as it gives a concrete rate of convergence to the steady-state.

25.2. Absorption Probabilities and Expected Time to Absorption

We have considered the long-term behavior of Markov chains. Now, we study the short-term behavior. In such considerations, we are concerned with the behavior of the chain starting in a transient state, till it enters a recurrent state. For simplicity, we can therefore assume that every recurrent state i is *absorbing*, i.e., $p_{ii} = 1$. The Markov chain that we will work with in this section has only transient and absorbing states.

If there is only one absorbing state i, then $\pi_i = 1$, and i is reached with probability 1. If there are multiple absorbing states, the state that is entered is random, and we are interested in the absorbing probability

$$a_{ki} = \mathbb{P}(X_n \text{ eventually equals } i \mid X_0 = k),$$

i.e., the probability that state i is eventually reached, starting from state k. Note that $a_{ii} = 1$ and $a_{ji} = 0$ for all absorbing $j \neq i$. For k a transient state, we have

$$\begin{aligned} a_{ki} &= \mathbb{P}(\exists n : X_n = i \mid X_0 = k) \\ &= \sum_{j=1}^{N} \mathbb{P}(\exists n : X_n = i \mid X_1 = j) p_{kj} \\ &= \sum_{j=1}^{N} a_{ji} p_{kj}. \end{aligned}$$

So we can find the absorption probabilities by solving the above system of linear equations.

Example: Gambler's Ruin A gambler wins 1 dollar at each round, with probability p, and loses a dollar with probability $1-p$. Different rounds are independent. The gambler plays continuously until he either accumulates a target amount m or loses all his money. What is the probability of losing his fortune?

We construct a Markov chain with state space $\{0, 1, \ldots, m\}$, where the state i is the amount of money the gambler has. So state $i = 0$ corresponds to losing his entire fortune, and state m corresponds to accumulating the target amount. The states 0 and m are absorbing states. We have the transition probabilities $p_{i,i+1} = p$, $p_{i,i-1} = 1 - p$ for $i = 1, 2, \ldots, m-1$, and $p_{00} = p_{mm} = 1$. To find the absorbing probabilities for the state 0, we have

$$\begin{aligned} &a_{00} = 1, \\ &a_{m0} = 0, \\ &a_{i0} = (1-p) a_{i-1,0} + p a_{i+1,0}, \quad \text{for } i = 1, \ldots, m-1. \end{aligned}$$

Let $b_i = a_{i0} - a_{i+1,0}$, $\rho = (1-p)/p$, then the above equation gives us

$$\begin{aligned} (1-p)(a_{i-1,0} - a_{i0}) &= p(a_{i0} - a_{i+1,0}) \\ b_i &= \rho b_{i-1} \end{aligned}$$

so we obtain $b_i = \rho^i b_0$. Note that $b_0 + b_1 + \cdots + b_{m-1} = a_{00} - a_{m0} = 1$, hence $(1 + \rho + \ldots + \rho^{m-1}) b_0 = 1$, which gives us

$$b_i = \begin{cases} \frac{\rho^i (1-\rho)}{1-\rho^m}, & \text{if } \rho \neq 1, \\ \frac{1}{m}, & \text{otherwise.} \end{cases}$$

Finally, $a_{i,0}$ can be calculated. For $\rho \neq 1$, we have for $i = 1, \ldots, m-1$,

$$\begin{aligned} a_{i0} &= a_{00} - b_{i-1} - \ldots - b_0 \\ &= 1 - (\rho^{i-1} + \ldots + \rho + 1) b_0 \\ &= 1 - \frac{1-\rho^i}{1-\rho} \frac{1-\rho}{1-\rho^m} \\ &= \frac{\rho^i - \rho^m}{1-\rho^m} \end{aligned}$$

and for $\rho = 1$,

$$a_{i0} = \frac{m-i}{m}.$$

This shows that for any fixed i, if $\rho > 1$, i.e., $p < 1/2$, the probability of losing goes to 1 as $m \to \infty$. Hence, it suggests that if the gambler aims for a large target while under unfavorable odds, financial ruin is almost certain.

The expected time of absorption μ_k when starting in a transient state k can be defined as $\mu_k = \mathbb{E}[\min\{n \geq 1 : X_n \text{ is recurrent}\} \mid X_0 = k]$. A similar analysis by conditioning on the first step of the Markov chain shows that the expected time to absorption can be found by solving

$$\mu_k = 0 \quad \text{for all recurrent states } k,$$

$$\mu_k = 1 + \sum_{j=1}^{N} p_{kj}\mu_j \quad \text{for all transient states } k.$$

25.3. References

- Sections 6.4,6.6 [**2**].
- Section 5.5 [**1**].

BIBLIOGRAPHY

1. R. Durrett, *Probability: theory and examples*, Duxbury Press, second edition, 1996.
2. G. R. Grimmett and D. R. Stirzaker, *Probability and random processes*, Oxford University Press, 2005.

6.436J / 15.085J Fundamentals of Probability
Fall 2008