PROBABILITIES: MODELS AND APPLICATIONS.
2. LIMIT RESULTS FOR MARKOV CHAINS

O. GOUBET

1. Setting the problem

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and S a finite (or countable) set. In this chapter a random variable is a measurable map from $(\Omega, \mathcal{F}, P)$ to $(S, \mathcal{P}(S))$, where $\mathcal{P}(S)$ is the $\sigma-$ algebra of all subsets of S.

Consider here X_n a discrete in time homogeneous Markov chain. We are interested in the limit behavior of $\mathbb{P}(X_n = j)$ when n diverges towards ∞. This amounts to consider, for Φ_i the initial distribution at time 0, the limit of

$$\mathbb{P}(X_n = j) = \sum_{i \in S} \Phi_i P^n_{i,j}.$$

Let us recall that we can discuss the convergence in distribution (or law) of the Markov chain.

Definition 1.1. *X_n converges in law towards X ($X_n \to_{\mathcal{L}} X$) if for any function f defined in S*

$$\mathbb{E}[f(X_n)|X_0 = i] \to \mathbb{E}_i[f(X)], \text{ where}$$

$$\mathbb{E}_i[f(X_n)] = \sum_{j \in S} f(j)\mathbb{P}(X_n = j|X_0 = i) = \sum_{j \in S} f(j) P^n_{i,j}. \tag{1}$$

1.1. Period of a Markov chain.

Consider an example. Let $S = \{1, 2, 3\}$ and let $X_{2n} = 1$ and $X_{2n+1} \in \{2, 3\}$ almost surely (say with $\frac{1}{2}$ chance to reach either 2 or 3). This Markov chain has a periodic behavior. Let us write the transition matrix P. We have $P = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$ and then $P^3 = P$. This Markov chain is irreducible and has period 2.

Proposition 1.2 (Periodicity)**.** *For any irreducible Markov chain we can find a partition of S into $E_0, ...E_{d-1}$ such that for any i in E_k (with $E_d = E_0$)*

$$\sum_{j \in E_{k+1}} p_{i,j} = 1.$$

Here d is maximal under the integers that has this property ; d is the period of the chain. If $d = 1$ the chain is **aperiodic**.

Date: November 2020.

Proof. Admitted □

If period $d = 3$ the transition matrix has the following block diagonal form

$$P = \begin{pmatrix} 0 & A & 0 \\ 0 & 0 & B \\ C & 0 & 0 \end{pmatrix};$$

The dynamics of the Markov chain on this example can be understood looking at P^{nd+k} for $k = 0, 1, 2$.

1.2. **Invariant probability.** We are now interested in the limit when n goes to ∞ of $\mathbb{P}(X_n = j)$ or $\mathbb{P}_i(X_n = j) = \mathbb{P}(X_n = j|X_0 = i)$. Assume that this limit exists. Then, using Markov property

$$\mathbb{P}(X_{n+1} = j|X_0 = i) = \sum_{k\in S} \mathbb{P}(X_{n+1} = j, X_n = k|X_0 = i) =$$
$$\sum_{k\in S} \mathbb{P}(X_{n+1} = j|X_n = k, X_0 = i)\mathbb{P}(X_n = k|X_0 = i) =$$
$$\sum_{k\in S} \mathbb{P}(X_{n+1} = j|X_n = k)\mathbb{P}(X_n = k|X_0 = i) =$$
$$\sum_{k\in S} p_{k,j}\mathbb{P}(X_n = k|X_0 = i).$$

Setting the vector $\pi(j) = \lim_{n\to+\infty} \mathbb{P}(X_n = j|X_0 = i)$ and passing to the limit n goes to ∞ we then have

$$\pi(j) = \sum_{k\in S} \pi(k)p_{k,j}. \tag{2}$$

If S is finite then this means that $\pi = \pi P$ that is π is an eigenvector associated to 1 for P^* the adjoint of P. Therefore it makes sense to seek invariant probabilities to understand the long run dynamics of a Markov chain.

Definition 1.3 (Invariant probability)**.** *A nonnegative vector π solving the system of equations* (2) *is called an* **invariant measure** *for the transition probabilities P. If S is finite, then $\pi = \pi P$. If $\sum_j \pi_j = 1$ then π is called invariant probability.*

Definition 1.4 (Reversibility)**.** *A Markov chain is* **reversible** *if there exists a measure π such that*

$$\pi_i p_{i,j} = \pi_j p_{j,i}.$$

Let us observe that a measure π that satisfies this system of equations is an invariant measure; actually

$$\pi_i = \sum_j \pi_i p_{i,j} = \sum_j \pi_j p_{j,i}.$$

1.3. **Limit of a Markov chain.** Let us recall

Definition 1.5 (Recurrence). *For a discrete finite time Markov chain we say that a state i is* **recurrent** *if and only if*

$$\mathbb{P}(T^i < +\infty | X(0) = i) = 1.$$

If not we say that i is **transient**.

A recurrent state i is **positive recurrent** *if*

$$\mathbb{E}[T^i | X_0 = i] < \infty.$$

If not it is **null recurrent**.

One can prove that in the same recurrent class all sites are either positive of null recurrent. If S is finite the is no null recurrent state.

Theorem 1.6. *Let X_n be an irreducible recurrent Markov chain. There exists (up to a multiplication by a constant) a unique invariant measure solving* (2). *Moreover the Markov chain is positive recurrent if and only if this invariant measure can be normalized into an invariant probability.*

Proof. If S is finite this is the Perron Fröbenius Theorem that implies that 1 is an eigenvalue for P^* of order one and that there is a positive eigenvector. We postpone the proof of the theorem if S is countable. $\square$

1.3.1. *Ergodic theorems.* Assume that there exists a unique invariant probability measure for a Markov chain. An ergodic theorem asserts that the average in time of the Markov chain converges towards the spacial average.

Theorem 1.7 (Ergodic theorem). *Let X_n be an irreducible Markov chain. Let $\pi_i = \frac{1}{\mathbb{E}_i(T^i)}$. Then almost surely*

$$\frac{1}{n} \sum_{k=1}^{n} 1\!\!1_{X_k = i} \to \pi_i.$$

Therefore for any function f defined on S

$$\lim_{n \to +\infty} \frac{1}{n} \sum_{k=1}^{n} f(X_k) = \sum_{j \in S} \pi_j f(j).$$

Proof.

If the Markov chain is transient then $\sum_{k=1}^{n} 1\!\!1_{X_k = i} < +\infty$ almost surely. Then the conclusion follows promptly.

Assume that the Markov chain is recurrent. From Theorem 2.8 of the first chapter, we know that the chain visit any state an infinite number of times. Consider i that is recurrent. Let $T_1 = T^i$. Let define recursively

$$T_{n+1} = \inf\{k \geq n; X_{S_n + k} = i\},$$

where $S_n = \sum_{k \leq n} T_k$.

Due to the Markov property, we can prove that the T_n are independent random variables. By the Law of Large Numbers we then have almost surely

$$\lim_{n\to\infty} \frac{S_n}{n} = \mathbb{E}_i(T^i).$$

Let us introduce $n(m) = \sum_{k=1}^m \mathbb{1}_{X_k=i}$. Then we claim that $S_{n(m)} \leq m < S_{n(m+1)}$; on the one hand $S_{n(m)} \leq m$ since the number of times you reach a target is always less than the number of tries. On the other hand for $k \geq n(m)$, then $X_k \neq i$, therefore $X_{n(m+1)}$ is larger than m (int the other cas you have a contradiction with the definition of $S_{n(m)}$).

Therefore we have almost surely

$$\lim_{m\to\infty} \frac{n(m)}{m} = \frac{1}{\mathbb{E}_i(T^i)},$$

that completes the proof of the Theorem. □

1.3.2. *Miscellaneous results.* Let us state the main result of this section

Theorem 1.8. *Let X_n be an irreducible positive recurrent aperiodic Markov chain. Then for any states i and any initial distribution it holds true*

$$\lim_{n\to\infty} \mathbb{P}(X_n = i) = \pi_i = \frac{1}{\mathbb{E}[T^i|X_0 = i]},$$

where π is the unique invariant probability that solves the system of equations $\pi_j = \sum_{i\in S} \pi_i p_{i,j}$.

Let us recall that the following formula holds true in $\mathbb{N} \cup \{\infty\}$.

$$\mathbb{E}[T^i|X_0 = i] = \sum_{n=1}^{\infty} n\mathbb{P}(T^i = n|X_0 = i). \tag{3}$$

Besides, for null-recurrent states or transient states, we have

Theorem 1.9. *For any null-recurrent or transient state i, for any choice of initial distribution, we have $\lim_{n\to+\infty} \mathbb{P}(X_n = i) = 0$.*

Proof. We sketch the proof of Theorem 1.8 if S is finite. Actually we are interested in the limit when n diverges towards ∞ to ΦP^n where Φ is the initial distribution.

The assumption that X_n is an irreducible Markov chain means that the stochastic matrix P is irreducible. The fact that it is positive recurrent implies that there is only one eigenvalue of modulus 1. In fact we can chose π the corresponding eigenvector to be positive also (Perron Fröbenius Theorem). Then normalizing Φ such that $\sum_i \Phi_i = 1$, we have that the sequence ΦP^n converges towards π almost surely (the only case when this does not occur is when Φ belongs to the complementary of $\mathbb{R}\Phi$ that is an hyperplane of measure 0).

We know that if a sequence converges, it converges also in the Cesaro sense (in average) to the same limit. Then appealing the Ergodic Theorem we know what is the limit. This completes the proof of the Theorem if S is finite.

$\square$.

2. First step analysis

In this section we discuss the following issues. Consider i a transient state for a Markov chain with several communication classes. Assume S is finite. How many times X_n will visit i before leaving for another communication class ? What is the probability to be *absorbed* by a closed recurrent class (for instance a trapping site) ?

2.1. **The example of the gambler's ruin.** Example 5. (Gambler's ruin) Two players play *heads or tails* where head occurs with probability p. The successive outcomes is an i.i.d. sequence indexed by $n \in \mathbb{N}$. Let X_n be the fortune of player A at time n. Then

$$X_{n+1} = X_n + Z_{n+1},$$
$$Z_{n+1} = 1 \text{ if the result of the toss is head,}$$
$$Z_{n+1} = -1 \text{ if the result of the toss is tails.}$$

Assume that players A and B starts respectively with a euros and b euros. Then the game ends if one player is ruined, that is $X_n = 0$ or $X_n = a + b$. Let S be $\{0, 1......a + b\}$ the set of possible states for the fortune of player A at time n. Then 0 and $a + b$ are trapping (absorbing) states.

Introduce the stopping time

$$T = \inf\{n > 0; X_n = 0 \text{ or } X_n = a + b\},$$

and the probability of win for player A that is $u(a) = \mathbb{P}(X_T = a+b|X_0 = a)$.

The *first-step* analysis reads as follows. We plan to compute $u(i)$ for any state i. Let us assume that i is not a trapping state. We have the recurrence formula, due to translation invariance

$$\mathbb{P}(X_n = a + b|X_0 = i) =$$
$$p\mathbb{P}(X_{n-1} = a + b|X_0 = i + 1) + (1 - p)\mathbb{P}(X_{n-1} = a + b|X_0 = i - 1).$$

Therefore

$$u(i) = pu(i + 1) + (1 - p)u(i - 1).$$

For $p = \frac{1}{2}$ the solution of this equation with $u(0) = 0$ and $u(a + b) = 1$ is $\frac{i}{a+b}$. For $p \neq \frac{1}{2}$ we have setting $\theta = \frac{1-p}{p}$ that $u(i) = \frac{1-\theta^i}{1-\theta^{a+b}}$.

For the general case we have to solve a countable system of equations

Theorem 2.1. *Consider a Markov chain on S. Let C be a recurrent class and T the set of all transient states. Introduce for $j \in T$*

$$u(j) = \mathbb{P}(X \text{ absorbed in } C | X_0 = j).$$

Then $u(j)$ is the smallest solution of the set of equations

$$u(j) = \sum_{k \in T} p_{j,k} u(k) + \sum_{k \in C} p_{j,k}.$$

2.2. **Average time.** We revisit the example of the gambler's ruin. We plan to compute $m_i = \mathbb{E}[T | X_0 = i]$ that is the *average duration* of the game. We handle this issue by the first-step analysis. Of course $m_0 = m_{a+b} = 0$ since 0 and $a+b$ are trapping sites.

We have the formula, due to the Markov property and due to translation invariance

$$\begin{aligned} m(i) = \mathbb{E}[T|X_0 = i] = \sum_{n=1}^{\infty} n\mathbb{P}(T = n | X_0 = i) = \\ \sum_{n=1}^{\infty} n \sum_k \mathbb{P}(T = n | X_1 = k, X_0 = i) p_{i,k} = \\ \sum_{n=1}^{\infty} n \sum_k \mathbb{P}(T = n | X_1 = k) p_{i,k} = \\ \sum_{n=1}^{\infty} n \sum_k \mathbb{P}(T = n-1 | X_0 = k) p_{i,k} = \\ 1 + \sum_{n=1}^{\infty} n \sum_k \mathbb{P}(T = n | X_0 = k) p_{i,k} = \\ 1 + pm(i+1) + (1-p)m(i-1). \end{aligned}$$

The solution of this equation is $m(i) = i(a+b-i)$ if $p = \frac{1}{2}$. If $p \neq \frac{1}{2}$, let us observe that a particular solution of the equation is $\frac{i}{1-2p}$ (here we forgot the boundary conditions $m_0 = m_{a+b} = 0$). Then the solution reads

$$m_i = \frac{i}{1-2p} + \alpha + \beta(\frac{1-p}{p})^i.$$

We chose α and β to satisfy the conditions $m_0 = m_{a+b} = 0$. Then

$$m_i = \frac{i}{1-2p} - \frac{a+b}{1-2p}(\frac{1-\theta^i}{1-\theta^{a+b}}),$$

where $\theta = \frac{1-p}{p}$.

Remark 2.2. *Consider the matrix $\tilde{P}$ obtained from P by eliminating the lines corresponding to absorbing states. Then the analysis above shows that* *the vector* $m = (m(i))$ *is solution to the equation*

$$(Id - \tilde{P})m = \mathbb{1}$$

where $\mathbb{1}$ *is the vector whose entries are* 1.

Références

[1] P. Brémaud *Markov chains. Gibbs fields, Monte Carlo simulation, and queues.* Springer texts in applied mathematics. Springer, 1998

[2] D. Chafai and C. Malrieu *Recueil de modèles aléatoires* Mathématiques et applications 78, Springer Berlin Heidelberg New-York 2016.

[3] J-F. Delmas and B. Jourdain, *Modèles aléatoires. Applications aux sciences de l'ingénieur et du vivant.* Mathématiques et applications 57, Springer Berlin Heidelberg New-York 2006.

[4] B. Hajek *Cooling schedules for optimal annealing,* Math. Op. Research, 13(2), 311-329, 1988.

[5] S. Méléard, *Modèles aléatoires en écologie et évolution,* Mathématiques et Applications, Springer-Verlag Berlin Heidelberg (2016)

(Olivier Goubet) Laboratoire Paul Painlevé CNRS UMR 8524, et équipe projet INRIA PARADYSE, Université de Lille, 59 655 Villeneuve d'Ascq cedex.

Email address: `olivier.goubet@univ-lille.fr`