PROBABILITIES: MODELS AND APPLICATIONS.
1. MARKOV CHAINS IN DISCRETE TIME

O. GOUBET

1. Mathematical framework

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 $(\Omega, \mathcal{F}, \mathbb{P})$ to $(S, \mathcal{P}(S))$, where $\mathcal{P}(S)$ is the $\sigma-$ algebra of all subsets of S.

1.1. **Discrete time stochastic process.**

Definition 1.1 (Discrete time stochastic process). *It is a sequence X_n of random variables from Ω to S indexed by discrete time $n \in \mathbb{N}$. The set S is the state space of the process.*

The *distribution (or law)* of the process is characterized by the point probabilities $\mathbb{P}(X_n = i_n, ...X_0 = i_0)$ for i_k in S.
Example 1. In a file (or queue) X_n represents the time (in number of minutes) that the n-th customer waits after arrival before receiving service.

Definition 1.2 (Markov chain). *A* **Markov chain** *is a discrete in time process that satisfies the so-called Markov property*

$$\mathbb{P}(X_n = i_n | X_{n-1} = i_{n-1}, ...X_0 = i_0) = \mathbb{P}(X_n = i_n | X_{n-1} = i_{n-1}).$$

Here $\mathbb{P}(A|B)$ denotes $\mathbb{P}(A \cap B)\mathbb{P}(B)^{-1}$ that is the conditional probability of A knowing B. A Markov process is a process without memory. The point probabilities at n depends only on what happens at $n-1$.

Example 2. (Gambler) You gamble m euros at a roulette. At each time n you play one euro. If the roulette gives an even number, you lose your euro. If the roulette gives an odd number, you win an extra euro. The Markov process is your fortune in euros after n rounds.

Example 3. (Random walk on $\mathbb{Z}$) Consider an i.i.d sequence ε_n that take values in $\{-1, 1\}$ with probability $\frac{1}{2}$. Set $X_n = \sum_{k=1}^{n} \varepsilon_k$. This represents the walk of someone starting from 0 that walks one step ahead or beyond at each round, with probability $\frac{1}{2}$.

Date: November 2020.

1.2. **Transition matrix.** Let X_n be a Markov chain and set

$$p_{i,j}^{(n)} = \mathbb{P}(X_n = j | X_{n-1} = i).$$

Due to the Markov property we have the following statement that can be proved by induction

$$\mathbb{P}(X_n = i_n, ...X_0 = i_0) = p_{i_{n-1},i_n}^{(n-1)} ....p_{i_0,i_1}^{(0)} \mathbb{P}(X_0 = i_0). \tag{1}$$

We now introduce the following

Definition 1.3. *If $p_{i,j}^{(n)}$ does not depend on n we write $p_{i,j}$ and we have an* **homogeneous** *Markov chain.*

In the sequel we consider homogeneous Markov chains and sometimes we forgot the word homogeneous.

Definition 1.4. *The matrix P whose entries are $p_{i,j}$ is the* **transition matrix** *of the Markov chain.*

A transition matrix P is a stochastic matrix, i.e. its entries are nonnegative and we have

$$\sum_{j \in S} p_{i,j} = 1.$$

From (1) we observe that a homogeneous Markov chain is characterized by its initial distribution and its transition matrix.

Example 3. The transition matrix is an infinite matrix that reads

$$P = \begin{pmatrix} .../... \\ 0 \; \frac{1}{2} \; 0 \; \frac{1}{2} ... \\ .. \; 0 \; \frac{1}{2} \; 0 \; \frac{1}{2} \\ ../... \end{pmatrix}.$$

Example 4. (Wright-Fisher model). We follow a population of N individuals to be chosen between two species α and β. At each step there is a death and birth process and X_n, that is the number of individuals of type α at time n is defined through the transition probabilities

$$\mathbb{P}(X_n = j | X_{n-1} = i) = \binom{N}{j} (\frac{i}{N})^j (1 - \frac{i}{N})^{N-j}.$$

The proof of the following statement goes by induction

Proposition 1.5. *Consider an homogeneous Markov chain on a finite state space with transition matrix P and initial distribution Φ. Then*

$$\mathbb{P}(X_n = j) = \sum_{i=1}^{\#S} \Phi_i (P^n)_{i,j}.$$

2. Classification of states

2.1. **Transition diagram.** An alternate way to represent Markov chain comes from graph theory. Consider a graph whose vertices are the states and the arrows represent the probability to move from one state to one another. For instance for the random walk on $\mathbb{Z}$ we have

$$(i-1) \xleftarrow{\frac{1}{2}} (i) \xrightarrow{\frac{1}{2}} (i+1)$$

If we change the rules of the game assuming that you have $\frac{1}{2}$ to stay at the same spot, and $\frac{1}{4}$ to move ahead or beyond then you have

$$\begin{array}{c} (i-1) \xleftarrow{\frac{1}{4}} (i) \xrightarrow{\frac{1}{4}} (i+1) \\ \circlearrowleft \\ \frac{1}{2} \end{array}$$

2.2. **Irreducibility.** Consider here an homogeneous Markov chain with discrete state space.

Definition 2.1 (Path)**.** *There is a* **path** *(or possible path) between states i and j if there exists a finite sequence of states $i_0 = i, i_1, \ldots . i_m = j$ such that $p_{i_{l-1},i_l} > 0$ for any l such that $1 \leq l \leq m$. This is equivalent to there exists a m such that $P^m_{i,j} > 0$. Two states communicates if there is a path between i and j and vice-versa. This relation denoted by $i \leftrightarrow j$ is an equivalence relation.*

Since $\leftrightarrow$ is an equivalence relation then S divides into equivalent classes of states/vertices. Let C be a communication class.

Definition 2.2 (Irreducibility)**.** *A Markov chain is* **irreducible** *if there is only one communication class. In other words : for any pair i, j there exists an integer m such that $P^m_{i,j} > 0$.*

Remark 2.3. *If S is finite a Markov chain is irreducible if and only if the transition matrix P is irreducible. In that case the Perron-Fröbenius theorem applies : 1 is a simple eigenvalue for P^* (the adjoint of P) and there is a nonnegative eigenvector in $Ker(P^* - Id)$.*

Definition 2.4. *A communication class is* **closed** *if and only if for any i in C we have $\sum_{j \in C} p_{i,j} = 1$.*

Indeed the restriction of a Markov chain to a closed communication class is also a Markov chain.

Definition 2.5 (Absorbing state)**.** *A state i is* **absorbing** *if $p_{i,i} = 1$. Absorbing states are trapping states for the Markov chain. If $X_n = i$ for some n then the chain is trapped into $\{i\}$ forever.*

Example. Consider $S = \{1, 2, 3\}$ and the Markov chain associated to the transition matrix $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$. Then we have two closed comunication classes that are $\{1\}$ and $\{2, 3\}$.
Example (Wright-Fisher) There are two absorbing states 0 and N and a non closed communication class $\{1, \ldots .N - 1\}$.

2.3. **Hitting time.**

Definition 2.6 (Hitting time)**.** *For any state i we define the* **hitting time** *of i by*

$$T^i = \inf\{n > 0; X_n = i\}.$$

The hitting time is a stopping time that takes values in $\mathbb{N} \cup \{+\infty\}$.

Let us observe that if $i \leftrightarrow j$ then

$$\mathbb{P}(T^i < +\infty | X_0 = j) > 0,$$

and vice-versa.

Definition 2.7 (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**.

We complete this definition.

Definition 2.8. *A recurrent state i is* **positive recurrent** *if*

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

If not it is **null recurrent**.

We now state

Theorem 2.9. *For a discrete homogeneous Markov chain with transition matrix P*

- *a state i is recurrent if and only if the probability that $X_n = i | X_0 = i$ for an infinite number of n is 1. We then have $\sum_n (P^n)_{i,i} = +\infty$.*
- *a state i is transient if and only if the probability that $X_n = i | X_0 = i$ for an infinite number of n is 0. We then have $\sum_n (P^n)_{i,i} < +\infty$.*

For the sake of simplicity we set $P^n_{i,j} = (P^n)_{i,j}$.
Proof. Let us define the events

$$F_{n,m} = \{X_n = i \text{ and } X_{n+k} \neq i; 1 \leq k \leq m\},$$

and

$$F_n = \cap_m F_{n,m} = \{X_n = i \text{ and } X_{n+k} \neq i; \forall k \geq 1\},$$

and

$$I = \{\#\{n; X_n = i\} < +\infty\}.$$

We have that

$$\mathbb{P}_i(I) = \mathbb{P}(I|X_0 = i) = \sum_n \mathbb{P}_i(F_n). \tag{2}$$

Besides we also have

$$\mathbb{P}_i(F_{n,m}) = \mathbb{P}(X_{n+k} \neq i; 1 \leq k \leq m | X_n = i, X_0 = i)\mathbb{P}_i(X_n = i).$$

Using Markov property and time translation invariance (the chain is homogeneous) we have

$$\mathbb{P}_i(F_{n,m}) = \mathbb{P}(X_{n+k} \neq i; 1 \leq k \leq m | X_n = i)\mathbb{P}_i(X_n = i) = \\ \mathbb{P}_i(F_{0,m})\mathbb{P}_i(X_n = i).$$

Letting $m \to \infty$ we then have

$$\mathbb{P}_i(F_n) = \mathbb{P}_i(F_0)\mathbb{P}_i(X_n = i),$$

and gathering this with (2) we have

$$\mathbb{P}_i(I) = \mathbb{P}(I|X_0 = i) = \mathbb{P}_i(F_0)\left(\sum_n \mathbb{P}_i(X_n = i)\right).$$

We now discuss this equality

— If $\mathbb{P}_i(F_0) > 0$ that implies that i is transient then $\sum_n \mathbb{P}_i(X_n = i) < +\infty$. This means that

$$\int \sum_n \mathbb{1}_{X_n = i} d\mathbb{P}_i < +\infty,$$

then $\sum_n \mathbb{1}_{X_n=i} < +\infty$ is true $\mathbb{P}_i$-almost surely and then $\mathbb{P}_i(I) = 1$

— If $\mathbb{P}_i(F_0) = 0$ then $\mathbb{P}_i(I) = 0$. This implies that $\sum_n \mathbb{1}_{X_n=i} = +\infty$ $\mathbb{P}_i$-almost surely and then $\sum_n \mathbb{P}_i(X_n = i) = +\infty$.

□

Corollary 2.10. *All states in the same communication class are either recurrent or transient.*

Proof. Assume $i \leftrightarrow j$ that is that there exists l, m such that $P^l_{i,j}P^m_{j,i} > 0$. By the Markov property the quantity $P^l_{i,j}P^k_{j,j}P^m_{j,i}$ is the probability that there exists some loop between i and i. Then

$$P^{l+k+m}_{i,i} \geq P^l_{i,j}P^k_{j,j}P^m_{j,i},$$

and then

$$\sum_n P^n_{i,i} \geq P^l_{i,j} \left(\sum_k P^k_{j,j} \right) P^m_{j,i},$$

and the result follows promptly. $\square$

We say that a communication class C is recurrent (respectively transient) if one state in C is recurrent (respectively transient).

Corollary 2.11. *A* finite *communication class is recurrent if and only if it is closed.*

Proof. Assume first that C is closed. This implies that for any i in C

$$\sum_{j \in C} P^k_{i,j} = 1.$$

Therefore by the Fubini-Tonelli theorem

$$\sum_{j \in C} \sum_{k=1}^{\infty} P^k_{i,j} = \sum_{k=1}^{\infty} \sum_{j \in C} P^k_{i,j} = +\infty.$$

Since C is *finite* by the pigeon-hole principle there exists j in C such that

$$\sum_{k=1}^{\infty} P^k_{i,j} = +\infty.$$

Using that there exists l such that $P^l_{j,i} > 0$ we then have

$$+\infty = (\sum_{k=1}^{\infty} P^k_{i,j}) P^l_{j,i} \leq \sum_n P^n_{i,i},$$

then i is recurrent.

Conversely, let us assume that there exists i in C such that

$$\sum_{j \in C} P_{i,j} < 1.$$

Then there exists a state j that is not in C such that $P_{i,j} > 0$. Then we have $\mathbb{P}_i(T^i = +\infty) \geq \mathbb{P}_i(X_1 = j) > 0$. this implies that $\mathbb{P}_i(T^i < +\infty) < 1$ and i is transient. $\square$

Corollary 2.12. *Consider an irreducible* recurrent *Markov chain. Then the probability that there exists a path between two states i and j is* 1.

Proof. Admitted. $\square$

Example (Wright-Fisher) : The Wright-Fisher Markov chain has two recurrent states 0 and N that are absorbing states and one transient class.

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`