

# Stochastic Processes and Markov's chains

Up to now we had put a lot of effort studying sequences of random variables that were independent. So now the next step will be to study sequences that are dependent in some sense.

Def 1 An stochastic process (or random process) is sequence of random variables  $\{X_i\}_{i \in I}$ .

In general,  $I$  could be  $\mathbb{R}$  (continuous-time stochastic process) or  $\mathbb{N}$  (discrete-time stochastic process). In this course we will only deal with the first case: let  $\{X_n\}_{n \geq 0}$  be a random process where each  $X_i$  takes value on a countable set  $S$ .

Def 1 The set  $S$  is called the state space.

In general, the random variables in an stochastic process may be dependent. However, we will start studying some very important type of process:

Def 1 An stochastic process  $\{X_i\}_{i \geq 0}$  is a Markov chain if  $\forall n$  we have that

$$P(X_{n+1} = s_{n+1} | X_n = s_n, \dots, X_0 = s_0) = P(X_{n+1} = s_{n+1} | X_n = s_n)$$

for all choice of  $s_{n+1}, s_n, \dots, s_0 \in S$  such that  $P(X_n = s_n, \dots, X_0 = s_0) \neq 0$ .

This definition tells us the following: the result of the random variable  $X_{n+1}$  only depends on what happened at time  $n$ . Indeed, this tells us that this definition is equivalent to having the property:

$$\textcircled{A} \quad P(X_{n+1} = s_{n+1} | X_{n_1} = s_{n_1}, X_{n_2} = s_{n_2}, \dots, X_{n_r} = s_{n_r}) = P(X_{n+1} = s_{n+1} | X_{n_1} = s_{n_1})$$

$n_1 < n_2 < \dots < n_r, \forall s_{n_1}, \dots, s_{n_r}, s_{n+1}$

$$\textcircled{B} \quad P(X_{n+m} = s_{n+m} | X_0 = s_0, X_1 = s_1, \dots, X_n = s_n) = P(X_{n+m} = s_{n+m} | X_n = s_n) \quad \forall n, m$$

$\forall s_0, s_1, \dots, s_n, s_{n+m}$

As we assume that  $S$  is countable, we may assume that  $S \subseteq \mathbb{N}$ . So, all the information of the Markov's chain  $\{X_i\}_{i \geq 0}$  is encoded by the probabilities  $P(X_{n+1} = i | X_n = j)$ . We may restrict to the case where invariance is satisfied:

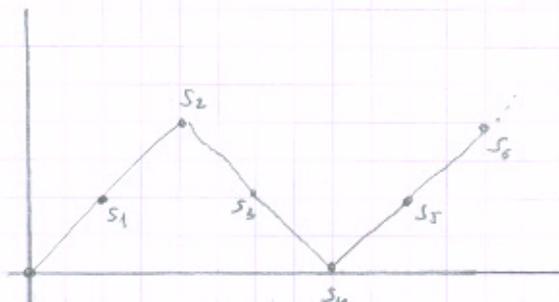
Def 1 A Markov chain  $\{X_i\}_{i \geq 0}$  is homogeneous if  $P(X_{n+1} = i | X_n = j) = P(X_1 = i | X_0 = j)$ .  
(where  $P(X_0 = j) \neq 0$ )

Ex 1. Random walk: realize that at the  $d$ -dimensional lattice we defined a random walk with  $T = \{(1, 0, \dots, 0), (0, \pm 1, \dots, 0), \dots, (0, \dots, \pm 1)\}$ , each with probability  $1/2d$ . The set of states is  $\mathbb{Z}^d = S$ , which is countable. Then the random process  $\{S_n\}_{n \geq 0}$ ,  $S_n = \sum_{i=1}^n X_i$ ,  $S_0 = 0$ ,  $X_i \sim X$ ,  $a \in T$ ,  $P(X=a) = 1/2d$  is a Markov chain:

$$P(S_{n+1} = s_{n+1} | S_n = s_n, S_{n-1} = s_{n-1}, \dots, S_0 = s_0) = P(X_{n+1} = s_{n+1} - s_n) = P(X_{n+1} = s_{n+1} - s_n)$$

Also,  $P(S_n = s | S_0 = a) = P(\sum_{i=1}^n X_i = s - a) = P(\sum_{i=1}^n X_i = s - a) = P(S_n = s | S_0 = a)$ , so the Markov chain is homogeneous.

Random walk with barrier: consider random walks starting at  $(0, 0)$ , with steps  $\uparrow = (1, 1)$  and  $\downarrow = (1, -1)$ . Assume that my random walk does NOT cross the  $y$ -axis. Then the model is slightly different:



- It is clearly a Markov chain because the position at state  $n$  only depends of the position at state  $n-1$ .  
 - Observe now that  $S_n$  now is a sum of random variables, which are NOT identically distributed:  $P(S_3 = (3,1) | S_2 = (2,1)) = 1 \neq P(S_3 = (3,1) | S_2 = (2,2)) = \frac{1}{2}$ .

- The chain is also time homogeneous.

③ Branching process: a branching process  $\{Z_n\}_{n \geq 0}$  with offspring distribution  $X$  satisfies the Markov's property as  $Z_{n+1}$  only depends on the number of elements in the  $n$ -th generation ( $Z_n$ ). Additionally, it is also time homogeneous.

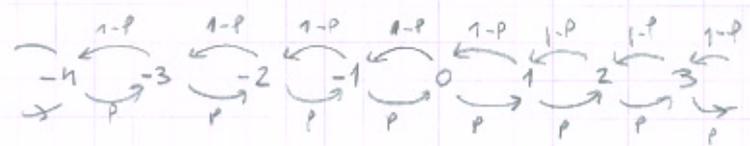
Under the homogeneous condition, we have the following property:  $P(X_{n+1} = j | X_n = i) = P(X_1 = j | X_0 = i) \equiv P_{ij} \equiv$  coefficient of the transition matrix.

Def/ The (possibly infinite) matrix  $P = (P_{ij})_{i,j}$ , with  $P_{ij} = P(X_1 = j | X_0 = i)$  is called the transition matrix of the Markov chain  $\{X_n\}_{n \geq 0}$ .

A transition matrix satisfies 2 main properties:

- Ⓐ  $P$  has non-negative entries
- Ⓑ  $\sum_{j \in S} P_{ij} = \sum_{j \in S} P(X_1 = j | X_0 = i) = \frac{1}{P(X_0 = i)} \cdot \sum_{j \in S} P(X_1 = j, X_0 = i) = \frac{1}{P(X_0 = i)} P(\bigcup_{j \in S} X_1 = j \cap X_0 = i) \stackrel{1}{=} 1$

Ex/ 1-dim random walk with  $p$  and  $1-p$  ( $\leftarrow$ , and  $\rightarrow$ , respectively):



$$P_{ij} = \begin{cases} p & j = i+1 \\ 1-p & j = i-1 \\ 0 & \text{in all other case} \end{cases}$$

$$P(X_1 = j | X_0 = i)$$

Once we have the matrix  $P$ , which encodes the information for 1 step, it is natural to try to get the information for  $n$  steps:  $P(X_{n+m} = j | X_m = i) = P(X_n = j | X_0 = i) \equiv P_{ij}^{(n)}$ . Define the corresponding matrix by  $P_{ij}^{(n)}$ . Can we compute it?

Theorem/ (Chapman-Kolmogorov's Equation)  $P_{ij}^{(n+m)} = \sum_{k \in S} P_{ik}^{(m)} P_{kj}^{(n)}$ .

Proof/ We have that

$$P_{ij}^{(n+m)} = P(X_{n+m} = j | X_0 = i) = \sum_{k \in S} P(X_{n+m} = j, X_m = k | X_0 = i) =$$

$$P(A \cap B | C) = \frac{P(A \cap B \cap C)}{P(C)} = \sum_{k \in S} P(X_{n+m} = j | X_m = k, X_0 = i) P(X_m = k | X_0 = i) =$$

$$\stackrel{11}{=} P(A|B \cap X) \cdot P(B|C) = \frac{P(A \cap B \cap C)}{P(B \cap C)} \cdot \frac{P(B \cap C)}{P(C)} = \sum_{k \in S} P(X_{n+m} = j | X_m = k) P(X_m = k | X_0 = i) = \sum_k P_{ik}^{(m)} P_{kj}^{(n)}$$

So, this result in particular gives us that  $P_{n+m} = P_n P_m$ , and then  $P_n = P^n$ . So this tells us how to relate long-term development in terms of short-term developments, and also how  $X_n$  depends on the initial choice of  $X_0$ .

vector  $\mu_i^{(n)} = P(X_n=i)$  for the mass function of  $X_n$ , and  $\bar{\mu}^{(n)}$  the row vector  $(\mu_0^{(n)}, \mu_1^{(n)}, \mu_2^{(n)}, \mu_3^{(n)}, \dots)$ .

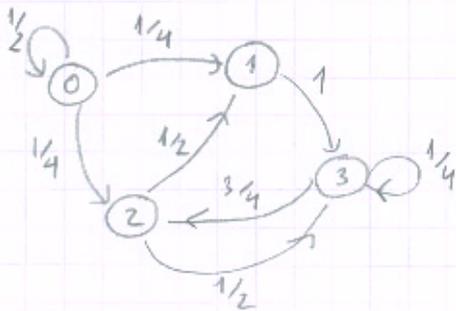
Lemma /  $\bar{\mu}^{(m+n)} = \bar{\mu}^{(m)} P^n$ . In particular,  $\bar{\mu}^{(n)} = \bar{\mu}^{(0)} P^n$ .

Proof / Just expand this term: denoting by  $(\bar{\mu}^{(m)} P)_j^n$  the  $j$ -th component of  $\bar{\mu}^{(m)} P$ , we get that:

$$\begin{aligned} \mu_j^{(m+n)} &= P(X_{m+n}=j) = \sum_{i \in S} P(X_{m+n}=j | X_m=i) P(X_m=i) \\ &= \sum_{i \in S} P(X_n=j | X_0=i) P(X_m=i) = \sum_{i \in S} \mu_i^{(m)} P_{ij}^{(n)} = (\bar{\mu}^{(m)})_j \end{aligned}$$

So this tells us that  $\bar{\mu}^{(n)}$  is determined by the matrix  $P$  and  $\bar{\mu}^{(0)}$ .

Ex / Consider the finite state Markov's chain with transition probabilities:



$$P = \begin{pmatrix} P_{00} & P_{01} & P_{02} & P_{03} \\ P_{10} & P_{11} & P_{12} & P_{13} \\ P_{20} & P_{21} & P_{22} & P_{23} \\ P_{30} & P_{31} & P_{32} & P_{33} \end{pmatrix} = \begin{pmatrix} 1/2 & 1/4 & 1/4 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1/2 & 0 & 1/4 \\ 0 & 0 & 3/4 & 1/4 \end{pmatrix}$$

taking  $\bar{\mu}^{(0)} = (P(X_0=0), P(X_0=1), P(X_0=2), P(X_0=3))$  we can see the evolution of the probabilities.

## Classification of the states

We saw that in random walks we had (in some cases) a certain probability to return to the origin, and we said that the random walks were transient or recurrent depending on the fact that this probability was 1 or  $< 1$ . We will develop similar ideas in this general context.

We can see the development of a Markov chain as the motion of a particle jumping between states. So one may wonder if some of the states are visited or not.

Def / Let  $X$  be a Markov chain with state space  $S$ . We say that the state  $i \in S$  is recurrent if

$$P\left(\bigcup_n X_n=i \mid X_0=i\right) = 1$$

otherwise, we say that the state is transient.

So in the study of random walks, we are also interested in the first passage through this state under study  $j$ , when starting at state  $i$ .

$$f_{ij}^{(n)} = P(X_n=j, X_{n-1} \neq j, \dots, X_1 \neq j \mid X_0=i)$$

and  $f_{ij} = \sum_{n=1}^{\infty} f_{ij}^{(n)}$ . In particular,  $f_{ij} = 1$  iff the state  $j$  is recurrent. Our objective now is to get a criterion to relate the different transition probabilities  $P_{ij}^{(n)}$  and  $f_{ij}^{(n)}$ . Write now

$$P_{ij}^{(n)} = \sum_m P_{ij}^{(n)} \delta^m; \quad F_{ij}^{(n)} = \sum_m f_{ij}^{(n)} \delta^m; \quad P_{ij}^{(0)} = \delta_{ij}; \quad f_{ij}^{(0)} = 0$$

Theorem / a)  $P_{ii}^{(n)} = 1 + F_{ii}^{(n)} P_{ii}^{(n)}$       c)  $F_{ij}^{(1)} = f_{ij}$   
 b)  $P_{ij}^{(n)} = F_{ij}^{(n)} P_{jj}^{(n)}$  if  $i \neq j$

Proof / Similar as we did for random walks! (HW).

As a consequence we have that:

Corollary / a) The state  $j$  is recurrent if  $\sum_n P_{jj}^{(n)} = \infty$ ; if this is true, then  $\sum P_{ij}^{(n)} = \infty$  for all  $i$  such that  $f_{ij} > 0$ .

b) The state  $j$  is transient if  $\sum_n P_{jj}^{(n)} < +\infty$ ; if this is true, then  $\sum_n P_{ij}^{(n)} < +\infty$  for all  $i$  such that  $f_{ij} > 0$ .

Proof / Write equation  $P_{jj}^{(s)} = 1 + F_{jj}^{(s)} P_{jj}^{(s)}$  as:

$$P_{jj}^{(s)} = (1 - F_{jj}^{(s)})^{-1} \Rightarrow P_{jj}^{(1)} = (1 - F_{jj}^{(1)})^{-1} \Rightarrow \begin{cases} \sum_n P_{jj}^{(n)} = \infty \Rightarrow f_{jj} = 1 \\ \sum_n P_{jj}^{(n)} < +\infty \Rightarrow f_{jj} < 1 \end{cases}$$

Now, to see the other statement, observe that

$$\begin{cases} \sum_n P_{jj}^{(n)} = +\infty \Rightarrow P_{jj}^{(1)} = +\infty \Rightarrow P_{ij}^{(1)} = F_{ij}^{(1)} P_{jj}^{(1)} \stackrel{f_{ij} > 0}{\Rightarrow} P_{ij}^{(1)} = +\infty \\ \sum_n P_{jj}^{(n)} < +\infty \Rightarrow P_{jj}^{(1)} < +\infty \Rightarrow P_{ij}^{(1)} = F_{ij}^{(1)} P_{jj}^{(1)} \stackrel{f_{ij} > 0}{=} 0 \quad F_{ij}^{(1)} < +\infty \end{cases}$$

In particular, if the state  $j$  is transient, then  $p_{jj}^{(n)} \rightarrow 0$  for all choice of  $j$ .

Let us move now to a series of definitions. For a Markov chain with  $X_0 = i$ , let  $T_{ij} = \min \{n \geq 1 : X_n = j\}$ , with the convention that  $T_{ij} = +\infty$  if  $X_n = j$  never occurs.

Def / The mean recurrence time  $\mu_i$  of the state  $i$  is:

$$\mu_i = \mathbb{E}[T_{ii}] = \begin{cases} \sum_n n f_{ii}^{(n)}, & i \text{ recurrent} \\ \infty, & i \text{ transient} \end{cases}$$

In particular,  $\mu_i$  could be infinite despite  $i$  being a recurrent state.

Def / An state  $i$  is null (non-null exp.) iff  $\mu_i = +\infty$  (or  $\mu_i < +\infty$ )

Def / An state  $i$  is absorbing if  $P_{ii} = 1$   $\forall n$ .

Def / The period  $d(i)$  of the state  $i$  is  $d(i) = \gcd \{n : P_{ii}^{(n)} > 0\}$ .

Ex / The period of each state in a random walk is 2.

Def / An state  $i$  is said to be ergodic if it is recurrent, non-null and with period 1 (aperiodic)

## Classification of chains

Once we have how an state could be, next step is based on looking on how states interact between them.

Def / We say that the state  $i$  communicate with  $j$  ( $i \leftrightarrow j$ ) if  $P_{ij}^{(n)} > 0$  for some  $n$ . We also say that the state  $i$  intercommunicate with  $j$  ( $i \leftrightarrow j$ ) if both  $P_{ij}^{(n)}, P_{ji}^{(n)} > 0$ .

Then, the relation " $\leftrightarrow$ " is an equivalence relation between states:  $P_{ii}^{(0)} = 1$  so  $i \leftrightarrow i$ . Also, by definition,  $i \leftrightarrow j$  iff  $j \leftrightarrow i$ . Finally, it is easy to check that  $i \leftrightarrow j$  and  $j \leftrightarrow k$  implies that  $i \leftrightarrow k$  (apply Chapman-Kolmogorov)

So one may wonder which are the common properties shared by states of the same equivalence class.

Theorem / If  $i \leftrightarrow j$  then

- $i$  is transient iff  $j$  is transient
- $d(i) = d(j)$

Proof / let us start proving a). If  $i \leftrightarrow j$ , then  $\exists m, n \geq 0$  such that  $\alpha = p_{ij}^{(m)} p_{ji}^{(n)} > 0$ . Then, by applying Chapman-Kolmogorov we have that:

$$p_{ii}^{(m+r+n)} \geq p_{ij}^{(m)} p_{jj}^{(r)} p_{ji}^{(n)} = \alpha p_{jj}^{(r)} \quad \forall r$$

So, now summing over  $r$  we get that  $\sum_r p_{jj}^{(r)} < +\infty$  if  $\sum_r p_{ii}^{(r)} < +\infty$ . So  $i$  transient  $\Rightarrow j$  transient. The converse is satisfied by changing  $i$  and  $j$ .

b) Show that  $d(i) | d(j)$  and  $d(j) | d(i)$  by using that  $\forall n, p_{ii}^{(n)} = \sum_{r+m+s=n} p_{ii}^{(r)} p_{ii}^{(m)} p_{ii}^{(s)}$

Def / A set  $C \subseteq S$  of states is closed if  $p_{ij} = 0 \quad i \in C, j \notin C$ .

Def / A set  $C \subseteq S$  of states is irreducible if  $i \leftrightarrow j \quad \forall i, j \in C$ .

Observe that a closed state with 1 element is indeed an absorbing element. From the previous theorem, all state in an irreducible set are aperiodic (for instance) if one of them is. So we can extend different notions to the whole family  $C$ .

Theorem / (Decomposition Theorem) let  $\mathbb{X} = \{X_i, i \geq 0\}$  be a Markov chain with state space  $S$ . Then  $S$  can be partitioned as:

$$S = T \cup C_1 \cup C_2 \cup \dots$$

where  $T$  is the set of transient states, and each  $C_i$  of irreducible closed of recurrent states.

Proof / Each  $C_i$  are the equivalence classes by ' $\leftrightarrow$ ' of recurrent states. let us see that each class is closed. Assume that  $i \in C_r$  and that  $i \rightarrow j$ . let us show that there is a positive probability to go from  $i$  to  $j$  without revisiting  $i$ :

$$0 < \sum_n p_{ij}^{(n)} = \sum_n \sum_r p_{ii}^{(r)} f_{ij}^{(n-r)} \Rightarrow \exists \text{ choice of } n \text{ and } r : f_{ij}^{(n-r)} > 0$$

Then, if  $j \rightarrow i$ , then  $p_{ji}^{(n)} = 0 \quad \forall n$ . However,  $i$  is recurrent, but with probability  $> 0$  the return is impossible (as  $f_{ij}^{(n-r)} > 0$ ), and this is a contradiction.

This theorem reads in the following way: take  $\mathbb{X} = \{X_i, i \geq 0\}$ ,  $S = T \cup C_1 \cup C_2 \cup \dots$  and  $X_0 \in C_1$ . Then all the sequence of state remains in  $C_1$ . If  $X_0 \in T$ , eventually the state goes through one of the  $C_i$ , and then it remains there.

In the special case of finite state, we have that:

Prop / If  $S$  is finite, then at least one state is recurrent.

Proof / Assume that all states are transient. Then for all  $j \in S$ ,  $p_{ij}^{(n)} \rightarrow 0 \quad \forall i, n \rightarrow \infty$ . But:

$$1 = \sum_{j \in \text{finite!}} p_{ij}^{(n)} \xrightarrow{n \rightarrow \infty} 0 \Rightarrow \text{Contradiction!}$$

So in all our theory we should assume that the Markov chain under consideration is irreducible. This is what we will explicit now.