Random Walks: WEEK 1

1 Markov chains - an introduction

1.1 Basic definitions

Definition 1.1. A Markov chain is a discrete-time stochastic process $(X_n, n \geq 0)$ such that each random variable X_n takes values in a discrete set S (the state space) and such that

$$\mathbb{P}(X_{n+1} = j \mid X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) = \mathbb{P}(X_{n+1} = j \mid X_n = i)$$

for all $n \geq 0$ and all states $i, j, i_0, \ldots, i_{n-1}$. In other words, the Markov property states that for a Markov process, given the present, the future does not depend on the past.

Remarks. - With a discrete countable state space S, we can typically identify S with $\mathbb{N}$ or $\mathbb{Z}$.

- Note that a continuous state space can also be considered.

Definition 1.2. - We say that the chain "is in state i at time n", and makes a "transition from state i to state j" between time n and $n+1$, and we denote the transition probability as

$$\mathbb{P}(X_{n+1} = j \mid X_n = i) \triangleq p_{ij}(n) \triangleq p_{i \to j}(n).$$

- The matrix $(P)_{ij} = p_{ij}$ is called the transition matrix.

Basic properties. Note the following properties on transition probabilities:

(a) $0 \leq p_{ij} \leq 1, \ \forall\, i, j \in S$.

(b) $\sum_{j \in S} p_{ij} = 1$.

Proof. $\sum_{j \in S} p_{ij} = \sum_{j \in S} \mathbb{P}(X_{n+1} = j \mid X_n = i) = \sum_{j \in S} \frac{\mathbb{P}(X_{n+1} = j, X_n = i)}{\mathbb{P}(X_n = i)} = \frac{\mathbb{P}(X_n = i)}{\mathbb{P}(X_n = i)} = 1.$

Definition 1.3. The initial distribution of a Markov chain is given by $\mathbb{P}(X_0 = i) = \pi_i^{(0)} \ \forall\, i \in S$.

Definition 1.4. An homogeneous Markov chain is a Markov chain such that

$$\mathbb{P}(X_{n+1} = j \mid X_n = i) = p_{ij} = \mathbb{P}(X_1 = j \mid X_0 = i)$$

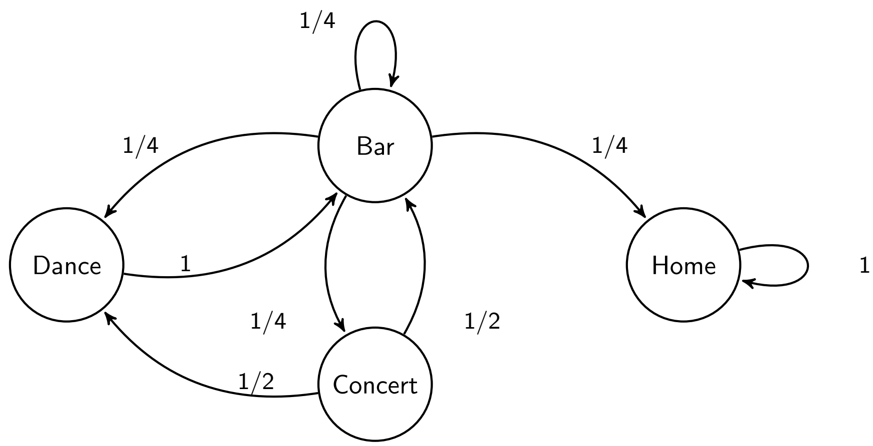
is independent of the time n.

1.2 Graphical representation

Example 1.5. (Music festival)
The four states of a student in a music festival are $S = \{$"dancing", "at a concert", "at the bar", "back home"$\}$. Let us assume that the student changes state during the festival according to the following transition matrix:

$$P \quad = \quad \begin{array}{cl} \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1/2 & 0 & 1/2 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 \\ 0 & 0 & 0 & 1 \end{pmatrix} & \begin{matrix} \leftarrow \textit{Dance} \\ \leftarrow \textit{Concert} \\ \leftarrow \textit{Bar} \\ \leftarrow \textit{Home} \end{matrix} \\ \begin{matrix} \textit{Dance} \uparrow & \textit{Concert} \uparrow & \textit{Bar} \uparrow & \textit{Home} \uparrow \end{matrix} & \end{array}$$

Then this Markov chain can be represented by the following transition graph:



Example 1.6. (Simple symmetric random walk)
Let $(X_n, n \geq 1)$ be i.i.d. random variables such that $\mathbb{P}(X_n = +1) = \mathbb{P}(X_n = -1) = \frac{1}{2}$, and let $(S_n, n \geq 0)$ be defined as $S_0 = 0, S_n = X_1 + \ldots + X_n,\ \forall n \geq 1$. Then $(S_n, n \geq 0)$ is a Markov chain with state space $S = \mathbb{Z}$. Indeed:

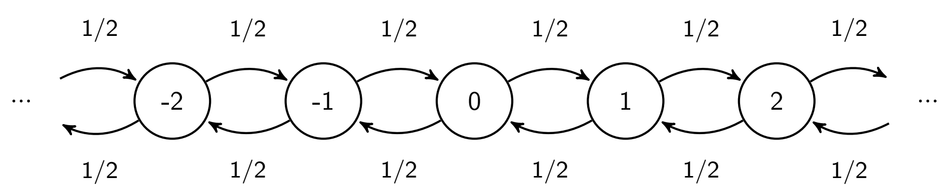
$$\begin{aligned}
&\mathbb{P}(S_{n+1} = j | S_n = i, S_{n-1} = i_{n-1}, \ldots, S_0 = i_0) \\
&= \mathbb{P}(X_{n+1} = j - i | S_n = i, S_{n-1} = i_{n-1}, \ldots, S_0 = i_0) \\
&= \mathbb{P}(X_{n+1} = j - i) \\
&= \mathbb{P}(S_{n+1} = j | S_n = i)
\end{aligned}$$

The third equality follows from the assumption that variables X_n are independent. The last equality shows that the Markov property holds. The chain is moreover time-homogeneous as

$$\mathbb{P}(X_{n+1} = j - i) = \begin{cases} \frac{1}{2} & \text{if } |j - i| = 1 \\ 0 & \text{otherwise} \end{cases}$$

does not depend on n.

Here is the transition graph of the chain:



1.3 Time evolution and the Chapman-Kolmogorov equation

We restrict our attention to homogeneous chains and are interested in the study of the evolution of such Markov chains in their long time behavior (i.e. what happens when $n \to +\infty$). The evolution of a Markov chain can be computed since we know the initial distribution $\mathbb{P}(X_0 = i) = \pi_i^{(0)}$, and the transition probabilities $\mathbb{P}(X_1 = j \mid X_0 = i) = p_{ij}$.

We want to compute the distribution of a state at time n:

$$\pi_i^{(n)} = \mathbb{P}(X_n = i),$$

and compute the probability of possible paths (think of the random walk):

$$\mathbb{P}(\text{“path through } i_0, i_1, \ldots, i_n\text{”}) = \mathbb{P}(X_0 = i_0, \ldots, X_n = i_n).$$

To compute this last probability, the Markov property becomes important for simplification:

$$\begin{aligned}
\mathbb{P}(X_n = i_n, \ldots, X_0 = i_0) &= \mathbb{P}(X_n = i_n \mid X_{n-1} = i_{n-1}, \ldots, X_0 = i_0)\, \mathbb{P}(X_{n-1} = i_{n-1}, ..., X_0 = i_0) \\
&\overset{(a)}{=} \mathbb{P}(X_n = i_n \mid X_{n-1} = i_{n-1})\, \mathbb{P}(X_{n-1} = i_{n-1}, \ldots, X_0 = i_0) \\
&\ldots \\
&= \mathbb{P}(X_n = i_n \mid X_{n-1} = i_{n-1})\, \mathbb{P}(X_{n-1} = i_{n-1} \mid X_{n-2} = i_{n-2}) \ldots \\
&\quad \ldots \mathbb{P}(X_1 = i_1 \mid X_0 = i_0) \mathbb{P}(X_0 = i_0) \\
&= \pi_{i_0}^{(0)} p_{i_0 i_1} p_{i_1 i_2} p_{i_2 i_3} \ldots p_{i_{n-1} i_n}
\end{aligned}$$

where (a) comes from the Markov property.

$$\begin{aligned}
\Longrightarrow \pi_{i_n}^{(n)} = \mathbb{P}(X_n = i_n) &= \sum_{i_0, \ldots, i_{n-1} \in S} \mathbb{P}(X_n = i_n, \ldots, X_0 = i_0) \\
&= \sum_{i_0, \ldots, i_{n-1} \in S} \pi_{i_0}^{(0)} p_{i_0 i_1} p_{i_1 i_2} p_{i_2 i_3} \ldots p_{i_{n-1} i_n} \\
&= \sum_{i_0 \in S} \pi_{i_0}^{(0)} (P^n)_{i_0 i_n} \\
\Longrightarrow \pi_{i_n}^{(n)} &= (\pi^{(0)} P^n)_{i_n}
\end{aligned}$$

where P^n is the n^{th} power of matrix P, and $\pi^{(0)}$ is the vector of initial probabilities.

To summarize, the study of Markov chains is essentially the study of Equation 1.3:

$$\pi^{(n)} = \pi^{(0)} P^n$$

This equation gives the time evolution of the probability distribution in the state. Notice also that $\pi^{(n)} = \pi^{(0)} P^n$ and $\pi^{(n-1)} = \pi^{(0)} P^{n-1}$ so that $\pi^{(n)} = \pi^{(n-1)} P$ or more generally:

$$\pi^{(n+m)} = \pi^{(n)} P^m$$

Theorem 1.7. The Chapman-Kolmogorov equation is simply the equation $P^{m+n} = P^m P^n$ expressed in components (with the transition probabilities):

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

or

$$p_{ij}^{n+m} = \sum_{k \in S} p_{ik}(n) p_{kj}(m)$$

Remarks. The Chapman-Kolmogorov equation expresses a consistency condition that the transition probabilities of the Markov chain have to satisfy. This consistency condition is necessary to have a Markov chain but not sufficient (see e.g. from Grimmett page 219 example 14).

The study of Markov chains is mainly the study of long time behavior. Multiple questions come out:

- When does $\pi^* = \pi^* P$ has a solution ? This is the question of existence of a “stationary distribution” that does not evolve with time (also referred to as equilibrium distribution).

- If π^* exists, when is it the case that $\pi^{(n)} \underset{n\to+\infty}{\longrightarrow} \pi^*$? This is called ergodicity.

- And how fast does $\pi^{(n)}$ approach π^* ? This question is important in algorithmic applications. For example, we will see that the spectrum of eigenvalues of P plays an important role.

1.4 Classification of states

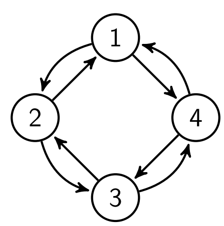
These questions all concern the long time behavior of the Markov chain. It is useful before turning to these questions to classify the type of states (and chains). We begin with a few easy definitions and terminology and then go over the very important notions of recurrent states and transient states.

Definition 1.8. A state $j \in S$ is said to be accessible from state i if $p_{ij}(n) > 0$ for some $n \geq 0$.

Definition 1.9. A state $i \in S$ is said to be absorbing if $p_{ii} = 1$. It means that once you reach an absorbing state, you stay there forever (e.g. state "Home" in Example 1.5).

Definition 1.10. A state $i \in S$ is said to be periodic with period d if $d = \mathrm{GCD}(n : p_{ii}(n) > 0)$. If $d = 1$, we say that the state is aperiodic. (Note that in general d may depend on i, but more on this later on.)

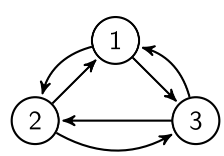
Example 1.11. Let us consider the following Markov chain:



Then $p_{11}(2) > 0$, $p_{11}(4) > 0$, $p_{11}(6) > 0, \ldots$ and $p_{11}(3) = p_{11}(5) = p_{11}(7) = \ldots = 0$. State 1 is thus a periodic state with period $d = 2$. And by symmetry all states have period $d = 2$.

Remarks. When all states are periodic, we say that the "chain is d-periodic".

Example 1.12. Note that for the chain



we have $p_{11}(2) > 0$, $p_{11}(3) > 0$, $p_{11}(4) > 0$, $p_{11}(5) > 0, \ldots$. Here State 1 is aperiodic and we say that $d = 1$.