Stochastic Modeling and Simulation

Dr. Nawel ARRAR
National School of Artificial Intelligence

2023-2024

Part I

Stochastic Processes and their Classification

Chapter 1

Discrete-Time Markov Chains

We cover Markov chain in depth, starting with Discrete-Time Markov Chains (DTMCs). In a DTMC, the world is broken up into synchronized time step. An event (arrival and departure) can only occur at the end of a time step. This property makes DTMCs a little odd for modeling computer science. However, there are many other problems that are well modeled by DTMCs.

In Continuous-Time Markov Chains (CTMCs) events can happen at any moment in time. This makes CTMCs convenient for modeling system.

1.1 Stochastic Processes

Definition 1.1 *A* ***stochastic process*** *(or random process) is a sequence of random variables* X_t :

$$\{X_t, t \in T\}$$

and defined on the same probabilized space $\{\Omega, \mathcal{A}, \mathbb{P}\}$. *The parameter* t *is generally interpreted as time and belongs to a given set.*

A ***continuous-time process*** *is one in which the set* T *is uncountable (usually* $\mathbb{R}_+$*). It is denoted by* $\{X(t), t \geq 0\}$.

A ***discrete-time process*** *occurs when* T *is finite or at least countable (most often* $T = \mathbb{Z}_+$*). It is denoted by* $\{X_n, n \geq 0\}$ *, we then speak of a stochastic sequence, while the term process is reserved for the continuous case.*

Definition 1.2 *The set of all possible values of the variables defining a stochastic process is called the* ***state space*** *of the process, and is denoted* S. *If* *this set is finite or infinite countable containing many values (or states), the process is called a* ***chain****.*

Definition 1.3 *A random process* $\{X_n, n = 0, 1, 2, ...\}$ *is* ***Markovian*** *where* X_n *denotes the state at (discrete) time step* n *and such that for all states* $i_0, ..., i_{n-1}, i_n, i_{n+1}$ *and all integer* $n \geq 0$

$$\mathbb{P}\left[X_{n+1} = i_{n+1}/X_n = i_n, X_{n-1} = i_{n-1}, ..., X_0 = i_0\right] = \mathbb{P}\left[X_{n+1} = i_{n+1}/X_n = i_n\right], \tag{1.1}$$

it is independent of the time step and of past history.

Definition 1.4 *The* ***Markovian property*** *states that the conditional distribution of any future state* X_{n+1}*, given that the past states* $X_0, X_1, ..., X_{n-1}$*, and given the present state* X_n*, is independent of past states and depends only on present states* X_n*.*

Remark 1.5 *The process is also said to be* ***memoryless****.*

1.2 Discrete-Time Markov Chains

Gathering the different notions presented above, we finally obtain the following definition

Definition 1.6 *A* ***Discrete-Time Markov Chain*** *is a stochastic process* $\{X_n, n \geq 0\}$ *satisfying the following three restrictions:*

- *The process has a discrete time;*
- *the states space* S *is finite or countable;*
- *the process satisfies the Markov property (1.1).*

In the following, we impose two additional restrictions on the DTMC. The first concerns the state space S, which we consider to be of finite cardinality. The second restriction corresponds to the time homogeneity property.

Definition 1.7 *A Markov chain MC is* ***homogeneous*** *(over time) if the probability of making a transition from one state to another is independent* *of the time at which the transition takes place. In other words, for all instant $n \geq 0$ and for any pair (i, j),*

$$\mathbb{P}\left[X_n = j / X_{n-1} = i\right] = \mathbb{P}\left[X_{n+k} = j / X_{n+k-1} = i\right], \qquad (1.2)$$

for all $k \geq 0$.

Remark 1.8 *Note, however, that the results presented remain valid for homogeneous chains with an infinite but countable number of states.*

1.2.1 Representative Graphs

A MC or more precisely its transition matrix, can be represented by an oriented graph G whose vertices correspond to the states of the chain and where an arc connects the vertices associated with the states i and j if the transition probability from i to j is positive, the graph thus defined is called the **representative graph** of the Markov chain.

1.2.2 Probabilities and matrix transition

Conditional probabilities do not vary over time. We can therefore simplify the notations and define the probabilities

$$p_{ij} = \mathbb{P}\left[X_n = j / X_{n-1} = i\right] \qquad (1.3)$$

for any pair of (i, j) independently of n.

Definition 1.9 *The probability p_{ij} is called **probability of transition** (or **of passage**) from the state i to the state j in one step and is equal to the conditional probability that the system will be in state j at the next step, given that it is currently in state i.*

If a MC has $s = |S|$ states, it exists s^2 probabilities of transition which can be arranged in a square matrix $s \times s$. We suppose that states of the chain are numbered from 1 to s, and we then speak of state k rather than the k^{th} state of the chain in the selected order.

Definition 1.10 *The **transition probability matrix** associated with any DTMC is a matrix*

$$P = (P_{ij}) = (p_{ij}),$$

whose entry in row i and column j is equal to the probability of moving to state j on the next transition, given that the current state is i.

$$P_{ij} = (p_{ij}).$$

Transition matrices are also called **stochastic matrices** and satisfy the following two conditions:

1. their elements are non-negative:

$$p_{ij} \geq 0, \forall i, j \in S,$$

2. the sum of the elements of each row is equal to 1 :

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

Example 1.11 (Repair facility problem) *A machine is either working or in the repair center. If it is working today, then there is a 95% chance that it will be working tomorrow. If it is in the repair center today, then there is a 40% chance that it will be working tomorrow. We are interested in questions like "what fraction of time does my machine spend in the repair shop?"*

Question 1. *Describe the DTMC for the repair facility problem.*

Question 2. *Now soppose that after the machine remains broken for 4days, the machine is replaced with a new machine. How does the DTMC diagram change?*

1.2.3 Power of P: n-Step Transition Probabilities

The transition probabilities p_{ij} and more generaly P, describe the evolution of the chain state step by step. It is also possible to calculate transition probabilities in several steps. We introduce the notation

$$p_{ij}^{(m)} = \mathbb{P}\left[X_{n+m} = j / X_n = i\right]$$

for the conditional probability of going from i to$j$ in exactly m steps.

The probabilities $p_{ij}^{(m)}$ are called **transition probabilities in** m **steps** and the matrix $P^{(m)}$ whose element (i, j) is equal to $p_{ij}^{(m)}$ is called the **transition matrix in** m **steps.**

To calulate transition probabilities in m steps, the probability that a process goes from state i to state j in two steps,

$$\begin{aligned} p_{ij}^{(2)} &= \mathbb{P}\left[X_2 = j/X_0 = i\right] \\ &= \sum_{k\in S} \mathbb{P}\left[X_2 = j/X_1 = k\right]\mathbb{P}\left[X_1 = k/X_0 = i\right] \\ &= \sum_{k\in S} p_{kj}p_{ik} = \sum_{k\in S} p_{ik}p_{kj}. \end{aligned}$$

So $p_{ij}^{(2)}$ is simply equal to the element (i,j) of the matrix P^2.

Theorem 1.12 *The probability $p_{ij}^{(m)}$ that a MC can be in state j after m states, if it is currently in state i, is given by the element (i,j) of the matrix P^m.*

In matrix form, this property becomes

$$P^{(m)} = \left(p_{ij}^{(m)}\right) = P^m$$

$P^m = P.P...P$, multiplied m time. We can use the notation P_{ij}^m to denote $(P^m)_{ij}$.

Example 1.13 ***Repair facility problem***

Now we consider again the simple repair facility problem, with general transition probability matrix P :

$$P = \begin{bmatrix} 1-a & a \\ b & 1-b \end{bmatrix}, 0 < a < 1, 0 < b < 1.$$

We prove by induction that

$$\begin{aligned} P^n &= \begin{bmatrix} \frac{b+a(1-a-b)^n}{a+b} & \frac{a-a(1-a-b)^n}{a+b} \\ \frac{b-b(1-a-b)^n}{a+b} & \frac{a+b(1-a-b)^n}{a+b} \end{bmatrix} \\ \lim_{n\to\infty} P^n &= \begin{bmatrix} \frac{b}{a+b} & \frac{a}{a+b} \\ \frac{b}{a+b} & \frac{a}{a+b} \end{bmatrix}. \end{aligned}$$

1.2.4 Classification of Chain States

There are several types of states. The classification of the states of a Mrkov chain plays an essential role in the study of its long-term behavior.

Definition 1.14 *Let i, j be two states $(i, j \in S)$*

- *We say that i **leads** to j if there is an integer $n \geq 0$ such that $p_{i,j}^n > 0\,(\neq 0)$.*
- *We say that i and j **communicate** if i leads to j and j leads to i.*

Note that i and j do not communicate if one (or both) of the equalities is satisfied.

$$\begin{aligned} p_{ij}^n &= 0 \quad \forall n \geq 0 \\ p_{ji}^n &= 0 \quad \forall n \geq 0. \end{aligned}$$

Recurrence

Definition 1.15 *Let (X_n) be a Markov chain and let i be a fixed state. For any integer $n \geq 1$*

$$f_{ii}^n = \mathbb{P}\left[X_n = i, X_k \neq i, k = 1, 2, \cdots, n-1 / X_0 = i\right], \left(f_{i,i}^0 = 0\right)$$

the probability that if the initial state is i, the first transition to state i occurs at the n^{th} transition.

Definition 1.16 *A state i is **recurrent** if*

$$f_{ii} = \mathbb{P}\left(\text{returning to state } i/\text{starting from state } i \text{ (initial)}\right) = 1.$$

*Otherwise $(f_{ii} < 1)$, the state is **transient**.*

Theorem 1.17 *A state i is recurrent if and only if $\sum_{n=0}^{\infty} p_{ii}^n = \infty$.*

Corollary 1.18 *If i and j communicate and i is recurrent, then j is recurrent.*

Proposition 1.19 *$f_{ii}^1 = p_{ii}$ and $p_{ii}^n = \sum_{k=0}^{n} f_{ii}^k p_{ii}^{n-k}$.*

Definition 1.20 *A recurrent state i is **zero recurrent** if*

$$\mu_i = E\left[\text{time of first return to } i/\text{departure from } i\right] = \infty$$

*Otherwise $(\mu_i < \infty)$, the state is **positive recurrent**. Here, time is measured in number of transitions.*

Periodicity

Definition 1.21 *The* ***period*** $d(i)$ *of a state* i *is defined as*

$$d(i) = \gcd\{n \geq 1 : p_{ii}^n > 0\}$$

Conventionally, $d(i) = 0$ if the above set is empty. A state i is **periodic** if $d(i) > 1$.

Otherwise ($d(i) = 0$ or 1), a state is **aperiodic** if it has period 1. A chain (X_n) is said to be aperiodic if all of its states are aperiodic.

Remark 1.22 *If* i *and* j *communicate, so* $d(i) = d(j)$.

So aperiodicity is clearly necessary for the limiting probabilities to exit. However in an aperiodic Markov chain, it could still turn out that the limiting probabilities depend on the start state, wherease we want p_{ij} be the same for all i.

If we also want the limiting probabilities to be independent of the start state, we need one more condition, known as irreducibility, which says that from any state one can get any other state.

Definition 1.23 *A Markov chain is said to be* ***irreducible*** *if all states communicate with each other*

Ergodicity

Definition 1.24 *An aperiodic and positive recurrent state is said to be* ***ergodic****. This is particularly true for a state* i *such as* $p_{ii} = 1$*, which is said to be* ***absorbing****.*

Example 1.25 *Let a DTMC on the states* $0, 1, 2, 3, 4$ *and* 5 *with the transition probabilities on one step given by the matrix*

$$P = \begin{pmatrix} 0 & 1/3 & 1/3 & 0 & 1/3 & 0 \\ 0 & 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 1/2 & 0 & 0 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1/2 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Describe the type of its states.

1.2.5 Stationary distribution

Limiting Probabilities

We now move on to looking at the limit. Consider the (i, j) th entry of the power matrix P^n for large n :

$$\lim_{n\to\infty} P_{ij}^n = \left(\lim_{n\to\infty} P^n\right)_{ij}.$$

This quantity represents the limiting probability if being in state j infinitely far into the future, given that we started in state i.

Definition 1.26 *Let*

$$\pi_j = \lim_{n\to\infty} P_{ij}^n = \lim_{n\to\infty} p_{ij}^{(n)}.$$

π_j *represents the **limiting probability** that the chain is in state* j *(independent of the starting state* i*). For an* n*-state DTMC, with states* $0, 1, ..., n-1$,

$$\overrightarrow{\pi} = (\pi_0, \pi_1, ..., \pi_{n-1}), \ \textit{where} \sum_{i=0}^{n-1} \pi_i = 1$$

*represents the **limiting distribution** of being in each state.*

For the rest of this chapter, we assume that the limiting probabilities exist.

Stationary Distribution

Let's consider a system evolving randomly over time, and assume that the evolution of the state of this system can be modeled by a Markov chain. Among the key questions for evaluating the performance of such a system, we can cite the following:

1. What is the probability of the system returning to state i after n transitions?
2. What is the proportion of time that the system spends in state i?
3. If the system is initially in state i, how many transitions will it make, on average, before visiting state j for the first time?

First of all, let's remember that the state of a Markov chain at a given time obviously depends on the transition probabilities p_{ij}, but also on the initial state such that:

$$\pi_i^{(0)} = \mathbb{P}\left(X_0 = i\right), i \in S.$$

Remark 1.27 *The special case where the initial state is known, corresponds to a vector $\pi^{(0)}$ whose components are all zero, except for the component associated with the initial state, which is* 1.

Starting from an initial distribution $\pi^{(0)}$, the probability $\pi_i^{(1)}$ that the system wil be in state i after a transition is as follows

$$\begin{aligned}\pi_i^{(1)} &= \mathbb{P}\left(X_1 = i\right) = \sum_{j\in S} \mathbb{P}\left(X_1 = i / X_0 = j\right) \mathbb{P}\left(X_0 = j\right) \\ &= \sum_{j\in S} p_{ji}\pi_j^{(0)} = \sum_{j\in S} \pi_j^{(0)} p_{ji}.\end{aligned}$$

In matrix form, this equality can be written as

$$\pi^{(1)} = \pi^{(0)} P.$$

In addition, using the homogeneity of transition probabilities, we also have

$$\pi^{(n)} = \pi^{(n-1)} P, \qquad n = 1, 2, ...$$

hence the following result:

Theorem 1.28 *If the distribution of the initial state of a Markov chain is given by the probability vector $\pi^{(0)}$, the distribution of the chain states after n steps is*

$$\pi^{(n)} = \pi^{(0)} P^n \tag{1.4}$$

where P is the transition matrix of the chain.

This study seeks to determine, as the number n of steps goes to infinity, if the distribution $\pi^{(n)}$ converges to a distribution π^*, the latter must not be affected by a transition. Indeed, if $\pi^{(n)}$ converges to $\pi^*, \pi^{(n+1)}$ converges to the same limit,

$$\pi^{(n+1)} = \pi^{(n)} P$$

and the matrix P has all its components bounded, so π^* satisfies

$$\pi^* = \pi^* P.$$

Definition 1.29 *A probability distribution* π *on the states of a Markov chain is* ***invariant*** *or* ***stationary*** *if it satisfies*

$$\pi = \pi P. \tag{1.5}$$

Definition 1.30 *A Markov chain for wich the limiting probabilities exist is said to be* ***stationary*** *or in* ***steady state*** *if the initial state is chosen according to the stationary probabilities.*

If a Markov Chain has long-term distribution, it depends a priori not only on P, but also on $\pi^{(0)}$. Indeed, let n goes to infinity in the equation (1.4), we obtain

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

where P^* designes, if it exist, the limit of P^n when n goes to infinity.

Theorem 1.31 *The distribution* $\pi^{(n)}$ *of a Markov chain goes to a limit* π^* *independent of the distribution* $\pi^{(0)}$ *if and only if the matrix* P^n *goes to a limit* P^* *whose all rows are identical to each other and, moreover, identical to* π^*.

Remark 1.32 *Given the limiting distribution* $\{\pi_j, j = 0, 1, 2, \cdots, n-1\}$ *exists, we can obtain it by solving the stationary equations*

$$\pi = \pi P \text{ and } \sum_{i=0}^{n-1} \pi_i = 1$$

where $\pi = (\pi_0, \pi_1, ..., \pi_{n-1})$.

Remark 1.33 *The same result holds for infinite state DTMCs.*

Example 1.34 (Repair facility problem with cost) *Consider again the Repair facility problem represented by the finite-state DTMC. The help desk is trying to figure out how much to charge me for maintaining my machine. They figure that it costs them 3000 DA every day that my machine is in repair. What will my annual repair bill be?*

Example 1.35 *We have to describe, if possible, the behavior of long-term distributions of the states of three Markov chains, their transition matrices are:*

1.

$$P = \begin{pmatrix} 1/4 & 0 & 3/4 \\ 1/4 & 1/4 & 1/2 \\ 1/4 & 1/2 & 1/4 \end{pmatrix}$$

2.

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

3.

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

The three previous examples illustrate the possible behavior of the long-term distribution of states in a Markov chain:

1. The powers of P converge to a stochastic matrix P^* whose rows are all equal to the same probability vector P^*, and the distribution of chain states also converges to π^*, independently of the initial distribution $\pi^{(0)}$.

2. The powers of P do not converge, nor does the distribution of chain states.

3. The powers of P converge to a stochastic matrix P^* whose rows are not all equal. The distribution of chain states admits a limit when n tends to infinity, but this limit depends on the initial distribution $\pi^{(0)}$.