

CHAPTER I:

MARKOV CHAINS AS MATHEMATICAL MODELS

University of Mila

Master 1, Artificial Intelligence and Its Applications

Academic Year: 2026/2027

Course: Modeling and Simulation

Lecturer: DR. SADEK BENHAMMADA

1. Stochastic Processes

We introduce just enough vocabulary to define Markov chains.

- A **random variable** describes a random phenomenon observed at a single instant; it does not evolve in time.

Example: X represents the result of one die roll, with possible values $\{1,2,3,4,5,6\}$.

- Many systems (weather, server load, a user browsing) must be described over time, so we need a *family* of random variables indexed by time.

Examples : Weather forecasting (rain, wind, temperature), Prices in Stock market

Definition 1.1 (Stochastic process). A **stochastic process** is a family of random variables $(X_t)_{t \in T}$ defined on the same probability space, where T is a set of times. All X_t take values in a common set E called the **state space**.

1. Stochastic Processes

Elements of a Stochastic Process

- **Time Space (T)**
 - Represents the set of possible observation times.
 - Example: $T = \{0,1,2,3, \dots\}$ for a system monitored every minute.
- **State Space (E)**
 - Represents all possible states of the system.
 - Example: $E = \{sunny, cloudy, rainy\}$ for a weather
- **A family of random variables $(X_t)_{t \in T}$.**
 - X_t represents the state of the system at time t
 - Each random variable X_t taking values in the **state space** (E).
 - The value of X_t is not known in advance; it depends on the outcome of a random phenomenon.

1. Stochastic Processes

- **Process with memory: the current state depends on previous states.**
 - **Example 1 (Weather).** $X_n \in \{\text{Sunny, Cloudy, Rainy}\}$ is the weather on day n . Sunny days tend to be followed by sunny days, so the X_n are *not* independent.
 - **Example 2 (Computer server).** $X_n \in \{\text{Low, Medium, High}\}$ is the load level measured every minute. Load typically changes gradually: from Low one goes to Low or Medium, rarely directly to High, so the X_n are *not* independent.
- **Independent process : successive states are independent; the past does not influence the future.**
 - **Example.** Successive tosses of a coin: $X_n \in \{\text{Head, Tails}\}$ with all X_n independent. This is a stochastic process with *no* memory at all and thus no dynamics to model. X_n are **independent**. This process has **no memory** and therefore does not capture dependence between successive states.

2. Discrete-Time Markov Chains

Definition (Markov property). A process $(X_n)_{n \geq 0}$ with values in a discrete state space E has the **Markov property** if for every $n \geq 0$ and all states $i_0, \dots, i_{n-1}, i, j \in E$ (such that the conditioning event has positive probability),

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

- The **probability of the next state** X_{n+1} depends only on the current state X_n , and not on the previous states.
- This means that the **current state at time (n)** contains all the information needed to determine the probability of the next state at time $(n + 1)$.
- Markov chains are therefore often described as “**memoryless**” **processes**: the future does not depend on the entire past, but only on the current state.

2. Discrete-Time Markov Chains

Definition (Homogeneous Markov Chain)

A Markov chain is said to be **homogeneous** if, for every $n \in \mathbb{N}$ and every $i, j \in E$, it satisfies:

$$P(X_{n+1} = i | X_n = j) = P(X_1 = i | X_0 = j)$$

- In other words, a Markov chain is **homogeneous** if the transition probability from one state to another does not depend on the time index (n), but only on the states involved.
- Therefore, the evolution of the process does not depend on **when we start observing it**; the transition probabilities remain the same over time.

2. Discrete-Time Markov Chains

- **Markov chains** are stochastic processes used to model the **evolution of a system**, where **we assume** that the state at time $t + 1$ depends only on the state at time t , and where the system has a finite number of possible states.
- At each step, the system **evolves by moving from one state to another**.
- The probabilities of moving to each possible state at time $t + 1$, given a particular state at time t , can be organized into a **square matrix** called the **transition matrix**.
- The **algebraic properties of the transition matrix** provide information about the **evolution and long-term behavior** of the system.

Remark. The Markov property is a property of a *model*, not of reality.

The art of modeling is to choose states so that the Markov assumption is a **reasonable approximation**.

2. Discrete-Time Markov Chains

Transition Probabilities

- The **transition probability** from i to j is the probability of moving to j in one step knowing that we are in i :

$$P(X_{t+1} = j | X_t = i) = P_{i,j}$$

- Example (Two-state weather chain).** $E = \{S, R\}$ (Sunny, Rainy). After a sunny day, the next day is sunny with probability 0.8 and rainy with probability 0.2. After a rainy day, it is sunny with probability 0.4 and rainy with probability 0.6. So $p_{SS} = 0.8$, $p_{SR} = 0.2$, $p_{RS} = 0.4$, $p_{RR} = 0.6$.

2. Discrete-Time Markov Chains

Transition Matrix

Definition (Transition matrix). The **transition matrix** of a chain with N states is the $N \times N$ matrix $\mathbf{P}$ whose entry in row i and column j is $p_{ij} = \mathbb{P}(X_{n+1} = j \mid X_n = i)$.

$$\mathbf{P} = (p_{ij})_{i,j \in E}$$

Example: Weather Transition Matrix

- Consider a simple weather model with three states: $E = \{\text{Sunny, Cloudy, Rainy}\} = \{0, 1, 2\}$
- Suppose the weather transition probabilities are:

$$\mathcal{P} = \begin{pmatrix} 0.7 & 0.2 & 0.1 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.3 & 0.5 \end{pmatrix}$$

	Sunny	Cloudy	Rainy
Sunny	0.7	0.2	0.1
Cloudy	0.3	0.4	0.3
Rainy	0.2	0.3	0.5

2. Discrete-Time Markov Chains

Properties of the Transition Matrix

- **Property (Stochastic matrix).** A transition matrix satisfies:
 1. $0 \leq P_{i,j} \leq 1$ for all i, j ;
 2. $\forall i \in E \sum_{j \in E} P_{i,j} = 1$ for every row i .
- A square matrix with these two properties is called a **row-stochastic matrix**.
- **Interpretation : Row i** is the probability distribution of the next state, given that we are currently in i . It sums to 1.

2. Discrete-Time Markov Chains

Graph Representation of a Markov Chain

- A **Markov chain** with state space E and transition matrix P can be represented by a **directed labeled graph**:
 $G = (E, A)$,

where each directed edge represents a transition with a **non-zero probability**.

- The set of directed edges is defined as:

$$A = \{(i, j) \mid P_{i,j} > 0\}$$

- Thus, there is a directed edge from state i to state j if and only if the transition probability $P_{i,j}$ is positive.

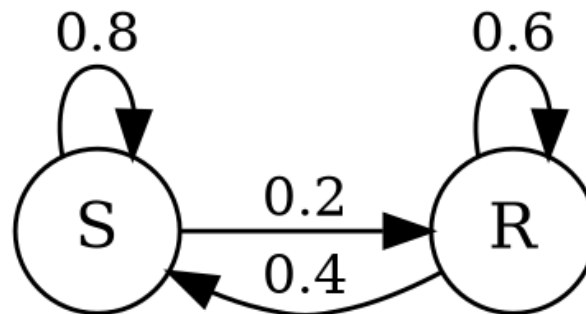
Example: Weather Markov Chain

Consider a simple weather model with two possible states: $E = \{\text{Sunny, Rainy}\}$

Therefore, the transition matrix is:

$$P = \begin{pmatrix} P_{SS} & P_{SR} \\ P_{RS} & P_{RR} \end{pmatrix} = \begin{pmatrix} 0.8 & 0.2 \\ 0.4 & 0.6 \end{pmatrix}$$

Transition graph:



2. Discrete-Time Markov Chains

Graph Representation of a Markov Chain

Example

$$P = \begin{matrix} & \begin{matrix} 0 & 1 & 2 & 3 & 4 & 5 \end{matrix} \\ \begin{matrix} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0.1 & 0.3 & 0.6 & 0 & 0 & 0 \\ 0.2 & 0.1 & 0.3 & 0 & 0.4 & 0 \\ 0 & 0.5 & 0 & 0.3 & 0 & 0.2 \\ 0 & 0 & 0 & 0.5 & 0 & 0.5 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix}$$

