

Unit 7. Stochastic processes

Jordi Villà i Freixa

Universitat de Vic - Universitat Central de Catalunya
Study Abroad. Operations Research

jordi.villa@uvic.cat
<https://mon.uvic.cat/cbbl>

last updated: December 11, 2025

December 10-17th, 2025

Preliminary

This course is strongly based on the monography on Operations Research by Carter, Price and Rabadi [1], and in material obtained from different sources (quoted when needed through the slides).

Markov chains: Summary

1 Markov chains

- State Transitions
- State Probabilities
- First Passage Probabilities
- State Properties
- Steady-State Analysis
- Expected First Passage Times
- Absorbing Chains
- Illustrative Applications

2 Hidden Markov Models

State Transitions

A Markov process is a discrete-time stochastic process in which the next state depends only on the current state:

$$P(X_{t+1} = s_{t+1} \mid X_t = s_t) = P(X_{t+1} = s_{t+1} \mid X_t = s_t, \dots, X_0 = s_0).$$

- Finite number of states: $1, \dots, N$
- System occupies one state at each time
- Transitions occur randomly according to probabilities

State Probabilities

Let $p^{(0)}$ be the initial state probability vector. State probabilities evolve as:

$$p^{(1)} = p^{(0)}P, \quad p^{(2)} = p^{(0)}P^2, \dots$$

To compute n -step transition probabilities:

$$P^{(n)} = P^n.$$

Examples include:

- computing probability of a sequence of weather conditions
- identifying that probabilities tend to stabilize in many cases

First Passage Probabilities

First passage probability $f_{ij}^{(n)}$ is the probability that the process enters j for the first time exactly at step n .

Properties:

$$p_{ij}^{(n)} = f_{ij}^{(n)} + \sum_k f_{ij}^{(k)} p_{jj}^{(n-k)}.$$

Used to compute:

- probability of reaching a state for the first time
- probability of return to a state after some steps

Properties of States

- **Reachable:** j reachable from i if $p_{ij}^{(n)} > 0$ for some n
- **Irreducible:** every state reachable from every other state
- **Closed set:** no transitions from the set to outside
- **Absorbing state:** once entered, never left ($p_{ii} = 1$)
- **Transient state:** possible to leave forever
- **Periodic state:** returns occur only at multiples of some $t > 1$
- **Ergodic chain:** irreducible and aperiodic

Steady-State Analysis

For an ergodic chain:

$$P^n \rightarrow \begin{bmatrix} \pi \\ \pi \\ \vdots \\ \pi \end{bmatrix}, \quad \pi P = \pi, \quad \sum \pi_i = 1.$$

Interpretation:

- π_i = long-run probability of being in state i
- independent of initial state

Expected First Passage Times

First passage time T_{ij} : steps needed to reach j from i .

For $i \neq j$:

$$m_{ij} = 1 + \sum_k p_{ik} m_{kj}.$$

For recurrence:

$$m_{ii} = \frac{1}{\pi_i}.$$

Applications include computing expected times between weather changes.

Absorbing Chains

A Markov chain with at least one absorbing state.

Reorder states so transient states come first:

$$P = \begin{bmatrix} Q & R \\ 0 & I \end{bmatrix}.$$

- Q : transitions among transient states
- R : transitions from transient to absorbing

Fundamental Matrix and Absorption Probabilities

Fundamental matrix:

$$F = (I - Q)^{-1}.$$

Meaning:

- element f_{ij} = expected number of visits to transient state j starting from i

Absorption probabilities:

$$A = F R.$$

Water Reservoir Operations

- States represent storage volumes
- Transitions derived from historical inflow data and release decisions
- Objectives: flood control, power generation, low-season flow
- Markov analysis improves long-term expected reward and performance

Dynamic Memory Allocation

- States represent sequences of allocated/free memory block lengths
- Transition matrix derived from allocation/release rules
- Steady-state probabilities used to evaluate memory usage
- Applicable to “first-fit”, “best-fit”, “worst-fit”, buddy systems

Manufacturing Production Capability

- States track failures in a multi-unit system
- Elements may be catastrophic, dependent, or independent
- Steady-state analysis gives:
 - expected production rate
 - probability of shutdown
 - optimal repair policies

Virtual Memory Systems

- States represent page configurations (FIFO, LRU)
- Transitions occur on memory references
- Steady-state analysis yields expected page fault rates
- Allows comparison of replacement algorithms

Exercise 1

Here is a simple game:

- a player bets on coin tosses, a dollar each time,
- the game ends either when the player has no money left or is up to five dollars.

If the player starts with three dollars, what is the chance that the game takes at least five flips? Twenty-five flips?

At any point, this player has either \$0, or \$1, ..., or \$5. We say that the player is in the state s_0, s_1, \dots, s_5 . A game consists of moving from state to state. For instance, a player now in state s_3 has on the next flip a .5 chance of moving to state s_2 and a .5 chance of moving to s_4 . The boundary states are a bit different; once in state s_0 or state s_5 , the player never leaves.

Let $p_i(n)$ be the probability that the player is in state s_i after n flips. Then, for instance, we have that the probability of being in state s_0 after flip $n + 1$ is $p_0(n + 1) = p_0(n) + 0.5 \cdot p_1(n)$. This matrix equation summarizes.

$$\begin{pmatrix} 1 & .5 & 0 & 0 & 0 & 0 \\ 0 & 0 & .5 & 0 & 0 & 0 \\ 0 & .5 & 0 & .5 & 0 & 0 \\ 0 & 0 & .5 & 0 & .5 & 0 \\ 0 & 0 & 0 & .5 & 0 & 0 \\ 0 & 0 & 0 & 0 & .5 & 1 \end{pmatrix} \begin{pmatrix} p_0(n) \\ p_1(n) \\ p_2(n) \\ p_3(n) \\ p_4(n) \\ p_5(n) \end{pmatrix} = \begin{pmatrix} p_0(n + 1) \\ p_1(n + 1) \\ p_2(n + 1) \\ p_3(n + 1) \\ p_4(n + 1) \\ p_5(n + 1) \end{pmatrix}$$

With the initial condition that the player starts with three dollars, calculation gives this.

$n = 0$	$n = 1$	$n = 2$	$n = 3$	$n = 4$	\dots	$n = 24$
$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0 \\ 0 \\ .5 \\ 0 \\ .5 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 \\ .25 \\ 0 \\ 0 \\ .5 \\ 0 \\ .25 \end{pmatrix}$	$\begin{pmatrix} .125 \\ 0 \\ .375 \\ 0 \\ 0 \\ .25 \\ .25 \end{pmatrix}$	$\begin{pmatrix} .125 \\ .1875 \\ 0 \\ .3125 \\ 0 \\ 0 \\ .375 \end{pmatrix}$	\dots	$\begin{pmatrix} .39600 \\ .00276 \\ 0 \\ .00447 \\ 0 \\ 0 \\ .59676 \end{pmatrix}$

- The player quickly ends in either state s_0 or state s_5 . For instance, after the fourth flip there is a probability of 0.50 that the game is already over: The boundary states are said to be **absorptive**.
- *The Markov chain model is historyless:* with a fixed transition matrix, the next state depends only on the current state, not on any prior states. More formally:

$$P(x_{t+1} = s_{t+1} | x_t = s_t, x_{t-1} = s_{t-1}, \dots, x_1 = s_1, x_0 = s_0) = \\ P(x_{t+1} = s_{t+1} | x_t = s_t)$$

- In addition, it is clear that the probability of a transition from state i to state j is independent of time (**stationarity property**):

$$P(x_{t+1} = j | x_t = i) = p_{ij}$$

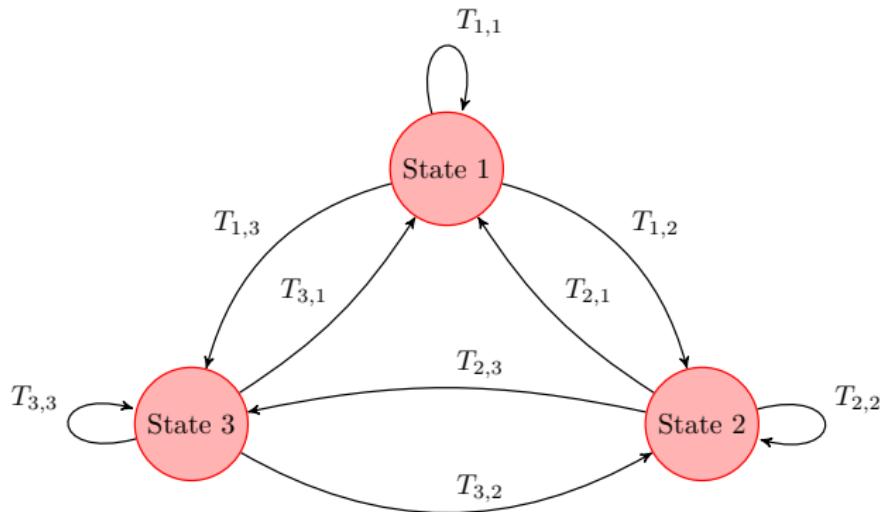
- Each column vector is a probability vector (and it sums up 1) and the matrix is a transition matrix.

Exercise 2

A study divided occupations in the United Kingdom into upper level (executives and professionals), middle level (supervisors and skilled manual workers), and lower level (unskilled).[2] To determine the mobility across these levels in a generation, about two thousand men were asked, “At which level are you, and at which level was your father when you were fourteen years old?”

This equation and transition diagram summarizes the results.

$$\begin{pmatrix} T_{1,1} & T_{1,2} & T_{1,3} \\ T_{2,1} & T_{2,2} & T_{2,3} \\ T_{3,1} & T_{3,2} & T_{3,3} \end{pmatrix}^t \rightarrow \begin{pmatrix} .60 & .29 & .16 \\ .26 & .37 & .27 \\ .14 & .34 & .57 \end{pmatrix} \begin{pmatrix} p_U(n) \\ p_M(n) \\ p_L(n) \end{pmatrix} = \begin{pmatrix} p_U(n+1) \\ p_M(n+1) \\ p_L(n+1) \end{pmatrix}$$



For instance, a child of a lower class worker has a .27 probability of growing up to be middle class. Notice that the Markov model assumption about history seems reasonable— we expect that while a parent's occupation has a direct influence on the occupation of the child, the grandparent's occupation has no such direct influence. With the initial distribution of the respondents's fathers given below, this table lists the distributions for the next five generations.

$n = 0$	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n = 5$
$\begin{pmatrix} .12 \\ .32 \\ .56 \end{pmatrix}$	$\begin{pmatrix} .23 \\ .34 \\ .42 \end{pmatrix}$	$\begin{pmatrix} .29 \\ .34 \\ .37 \end{pmatrix}$	$\begin{pmatrix} .31 \\ .34 \\ .35 \end{pmatrix}$	$\begin{pmatrix} .32 \\ .33 \\ .34 \end{pmatrix}$	$\begin{pmatrix} .33 \\ .33 \\ .34 \end{pmatrix}$

Exercise 3

The World Series of American baseball is played between the team winning the American League and the team winning the National League. The series is won by the first team to win four games. That means that a series is in one of twenty-four states: 0-0 (no games won yet by either team), 1-0 (one game won for the American League team and no games for the National League team), etc. If we assume that there is a probability p that the American League team wins each game then we have the following transition matrix.

$$\begin{pmatrix} 0 & 0 & 0 & 0 & \dots \\ p & 0 & 0 & 0 & \dots \\ 1-p & 0 & 0 & 0 & \dots \\ 0 & p & 0 & 0 & \dots \\ 0 & 1-p & p & 0 & \dots \\ 0 & 0 & 1-p & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \end{pmatrix} \begin{pmatrix} p_{0-0}(n) \\ p_{1-0}(n) \\ p_{0-1}(n) \\ p_{2-0}(n) \\ p_{1-1}(n) \\ p_{0-2}(n) \\ \vdots \end{pmatrix} = \begin{pmatrix} p_{0-0}(n+1) \\ p_{1-0}(n+1) \\ p_{0-1}(n+1) \\ p_{2-0}(n+1) \\ p_{1-1}(n+1) \\ p_{0-2}(n+1) \\ \vdots \end{pmatrix}$$

An especially interesting special case is $p = 0.50$; this table lists the resulting components of the $n = 0$ through $n = 7$ vectors. Check what happens with other values of p .

	$n = 0$	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n = 5$	$n = 6$	$n = 7$
0 – 0	1	0	0	0	0	0	0	0
1 – 0	0	0.5	0	0	0	0	0	0
0 – 1	0	0.5	0	0	0	0	0	0
2 – 0	0	0	0.25	0	0	0	0	0
1 – 1	0	0	0.5	0	0	0	0	0
0 – 2	0	0	0.25	0	0	0	0	0
3 – 0	0	0	0	0.125	0	0	0	0
2 – 1	0	0	0	0.375	0	0	0	0
1 – 2	0	0	0	0.375	0	0	0	0
0 – 3	0	0	0	0.125	0	0	0	0
4 – 0	0	0	0	0	0.0625	0.0625	0.0625	0.0625
3 – 1	0	0	0	0	0.25	0	0	0
2 – 2	0	0	0	0	0.375	0	0	0
1 – 3	0	0	0	0	0.25	0	0	0
0 – 4	0	0	0	0	0.0625	0.0625	0.0625	0.0625
4 – 1	0	0	0	0	0	0.125	0.125	0.125
3 – 2	0	0	0	0	0	0.3125	0	0
2 – 3	0	0	0	0	0	0.3125	0	0
1 – 4	0	0	0	0	0	0.125	0.125	0.125
4 – 2	0	0	0	0	0	0	0.15625	0.15625
3 – 3	0	0	0	0	0	0	0.3125	0
2 – 4	0	0	0	0	0	0	0.15625	0.15625
4 – 3	0	0	0	0	0	0	0	0.15625
3 – 4	0	0	0	0	0	0	0	0.15625

Evenly-matched teams are likely to have a long series— there is a probability of 0.625 that the series goes at least six games.

In summary, a system can be modeled as a Markov process if it has the following four properties:

- Property 1: A finite number of states can be used to describe the dynamic behavior of the system.
- Property 2: Initial probabilities are specified for the system.
- Property 3: Markov property—We assume that a transition to a new state depends only on the current state and not on past conditions.
- Property 4: Stationarity property—The probability of a transition between any two states does not vary in time.

If Markov analysis is appropriate for the system being studied one can try answering questions like:

- How many transitions (steps) will it likely take for the system to move from some specified state to another specified state?
- What is the probability that it will take some given number of steps to go from one specified state to another?
- In the long run, which state is occupied by the system most frequently?
- Over a long period of time, what fraction of the time does the system occupy each of the possible states?
- Will the system continue indefinitely to move among all N states, or will it eventually settle into a certain few states?

Hidden Markov Models: Summary

1 Markov chains

- State Transitions
- State Probabilities
- First Passage Probabilities
- State Properties
- Steady-State Analysis
- Expected First Passage Times
- Absorbing Chains
- Illustrative Applications

2 Hidden Markov Models

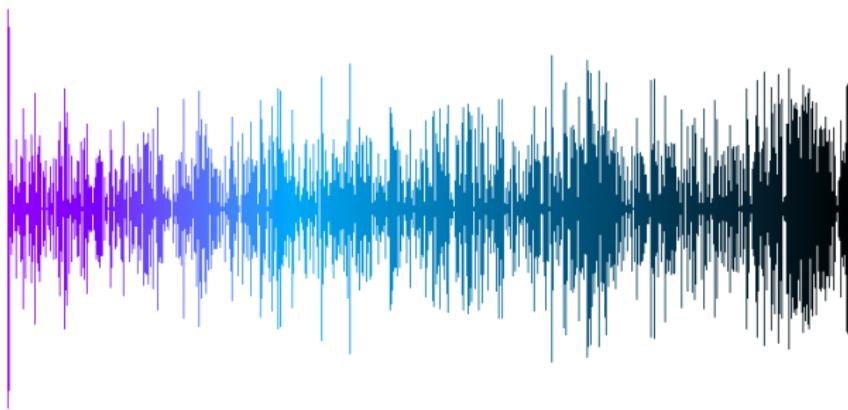
Hidden Markov Models

HMMs are probably the most popular directed graphical model out there. They are used in many sequential and temporal domains:

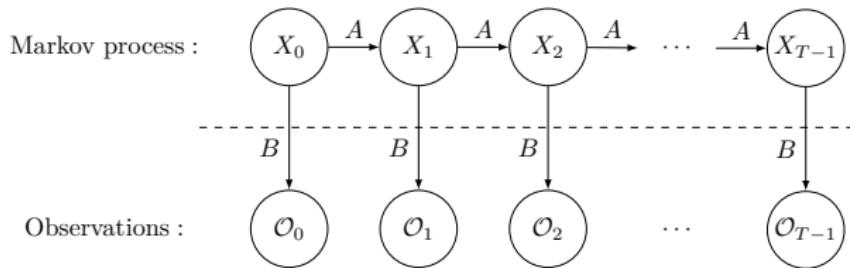
- speech recognition,
- handwriting recognition,
- visual target tracking and localization,
- machine translation,
- robot localization,
- gene prediction,
- time-series analysis,
- natural language processing and part-of-speech recognition,
- stochastic control,
- protein folding, ...

In HMMs, the random variables are divided into hidden states (phonemes, letters, target location) and observations (audio signal, pen strokes, target image). The goal is to predict the states from observations.

What is this?



We can start wth the graphical representation of the HMM



HMM

An HMM is defined by:

- A set of N hidden states $S = \{s_1, s_2, \dots, s_N\}$.
- A set of M observation symbols $O = \{o_1, o_2, \dots, o_M\}$.
- State transition probability distribution $A = \{a_{ij}\}$, where $a_{ij} = P(q_{t+1} = s_j | q_t = s_i)$.
- Observation symbol probability distribution $B = \{b_j(k)\}$, where $b_j(k) = P(v_k | q_t = s_j)$.
- Initial state distribution $\pi = \{\pi_i\}$, where $\pi_i = P(q_1 = s_i)$.

The complete parameter set of the model is $\lambda = (A, B, \pi)$.

References I

- [1] Michael Carter, Camille C Price, and Ghaith Rabadi. *Operations Research. A Practical Introduction. Second Edition.* CRC Press, 2019. ISBN: 978-1-4987-8010-0.
- [2] A. H. Halsey, ed. *British Social Trends since 1900.* London: Palgrave Macmillan UK, 1988. ISBN: 978-0-333-34522-1 978-1-349-19466-7. DOI: 10.1007/978-1-349-19466-7. URL: <http://link.springer.com/10.1007/978-1-349-19466-7>.