Hitting Times and Probabilities for Imprecise Markov Chains

Thomas Krak, Natan T'Joens, and Jasper De Bock

Foundations Lab for Imprecise Probabilities, Ghent University, Belgium

Abstract

Markov chains are models for describing dynamical systems under uncertainty. Important inferences for these models are the expected **hitting times and probabilities** of subsets of the state space of the system.

Imprecise Markov chains are a robust generalisation of Markov chains that are based on the theory of imprecise probabilities. However, this generalisation is not unambiguous, with various approaches being proposed in the literature; see **(I),(II),(III)** for details.

We proved that the (lower and upper) expected hitting times and probabilities are **the same for all these types** of imprecise Markov chains. Moreover, we derived a **characterisation of these inferences** that is a direct generalisation of a known characterisation for precise Markov chains.

Markov chains

Stochastic processes describe dynamical systems that evolve in an uncertain way. We consider a system that takes values in a finite **state space** $\mathscr{X}$ and that evolves at discrete steps in time. So, at each point in time $n \in \mathbb{N}_0$, the system is in an uncertain state X_n. A stochastic process is a **probability measure** describing how the system might evolve over time.

Markov Chains. A **Markov chain** is a stochastic process that satisfies a **conditional independence assumption** known as the **Markov property**. In particular, a stochastic process P is called a Markov chain if, for all $n \in \mathbb{N}_0$ and all $x_0, \ldots, x_n, y \in \mathscr{X}$, it holds that

$$P(X_{n+1} = y \,|\, X_{0:n} = x_{0:n}) = P(X_{n+1} = y \,|\, X_n = x_n).$$

A Markov chain is called **homogeneous** if, moreover, for all $n \in \mathbb{N}_0$ and all $x, y \in \mathscr{X}$, it holds that

$$P(X_{n+1} = y \,|\, X_n = x) = P(X_1 = y \,|\, X_0 = x).$$

Transition Matrices. Any homogeneous Markov chain P can be **uniquely characterised**—up to its initial distribution $P(X_0)$, which is irrelevant here—by a single **transition matrix** T. A transition matrix T is an $|\mathscr{X}| \times |\mathscr{X}|$ matrix that is **row-stochastic**, meaning that $\sum_{y \in \mathscr{X}} T(x,y) = 1$ and $T(x,y) \geq 0$ for all $x, y \in \mathscr{X}$. A given transition matrix T determines a homogeneous Markov chain P for which

$$P(X_{n+1} = y \,|\, X_n = x) = T(x,y),$$

for all $x, y \in \mathscr{X}$ and all $n \in \mathbb{N}_0$.

Note that functions of the form $f : \mathscr{X} \to \mathbb{R}$ live in the vector space $\mathbb{R}^{|\mathscr{X}|}$, and any transition matrix T encodes a linear operator on this space. That is, for all $f \in \mathbb{R}^{|\mathscr{X}|}$ and all $x \in \mathscr{X}$, we can write $[Tf](x) := \sum_{y \in \mathscr{X}} T(x,y) f(y)$.

Hitting times and probabilities

We consider the **expected hitting times** and **hitting probabilities** of a set $A \subset \mathscr{X}$ that the system can be in.

Expected Hitting Times. Expected hitting times $\mathbb{E}_P[H_A \,|\, X_0]$ for a Markov chain P are the expected number of steps H_A before the system visits an element of A, if it started in state X_0. For a Markov chain P with transition matrix T, the vector h_A of expected hitting times, defined as $h_A(x) := \mathbb{E}_P[H_A \,|\, X_0 = x]$ for all $x \in \mathscr{X}$, is the **minimal non-negative solution** to the system

$$h_A = \mathbb{I}_{A^c} + \mathbb{I}_{A^c} \cdot T h_A,$$

where $\mathbb{I}_{A^c}$ is the indicator of $A^c := \mathscr{X} \setminus A$.

Hitting Probabilities. Hitting probabilities $\mathbb{E}_P[G_A \,|\, X_0]$ are the probabilities that the process P will **at some time** visit an element of A, if it started in state X_0. Here G_A is a function that takes the value one if the process ever visits A and takes value zero otherwise; it is a function of the complete evolution of the system. For a Markov chain P with transition matrix T, the vector p_A of hitting probabilities, defined as $p_A(x) := \mathbb{E}_P[G_A \,|\, X_0 = x]$ for all $x \in \mathscr{X}$, is the **minimal non-negative solution** to the system $p_A = \mathbb{I}_A + \mathbb{I}_{A^c} \cdot T p_A$.

Imprecise Markov chains

When the transition matrix T is **not known exactly**, or when we think the **independence assumptions might be violated**, we can instead model the underlying system using an **imprecise Markov chain**. There are various ways in which an imprecise Markov chain can be defined—see **(I),(II),(III)**—but they are all parameterised in the same way.

Sets of transition matrices and transition operators. Any imprecise Markov chain is parameterised—up to its initial model—by a **set of transition matrices** $\mathscr{T}$, which must satisfy some technical closure conditions. Colloquially, we can interpret $\mathscr{T}$ to contain all transition matrices T that we deem "plausible".

Dually, we can describe an imprecise Markov chain with the **lower and upper transition operators** $\underline{T}$ and $\overline{T}$. These operate on functions $f : \mathscr{X} \to \mathbb{R}$, and they are defined respectively, for all $x \in \mathscr{X}$, as

$$\big[\underline{T} f\big](x) := \inf_{T \in \mathscr{T}} \big[Tf\big](x) \quad \text{and} \quad \big[\overline{T} f\big](x) := \sup_{T \in \mathscr{T}} \big[Tf\big](x).$$

The required closure conditions ensure that in fact

$$\underline{T} f = \min_{T \in \mathscr{T}} Tf \quad \text{and} \quad \overline{T} f = \max_{T \in \mathscr{T}} Tf.$$

Inferences for Imprecise Markov Chains. In all cases, inferences for imprecise Markov chains are described using **lower and upper expectations**, which can be interpreted as **tight bounds** on the inferential quantities of interest. These satisfy an **imprecise Markov property**, in the sense that

$$\underline{\mathbb{E}}_{\mathscr{T}}\big[f(X_{n+1}) \,\big|\, X_{0:n} = x_{0:n}\big] = \big[\underline{T} f\big](x_n),$$

and similarly for the upper expectations.

(I) Sets of homogeneous MCs

Arguably the simplest imprecise Markov chain is $\mathscr{P}^{\mathrm{H}}_{\mathscr{T}}$, which is the **set of all homogeneous Markov chains** whose transition matrix T is included in $\mathscr{T}$. Its **lower and upper expectation operators** are defined as

$$\underline{\mathbb{E}}^{\mathrm{H}}_{\mathscr{T}}[\cdot\,|\,\cdot] := \inf_{P \in \mathscr{P}^{\mathrm{H}}_{\mathscr{T}}} \mathbb{E}_P[\cdot\,|\,\cdot] \quad \text{and} \quad \overline{\mathbb{E}}^{\mathrm{H}}_{\mathscr{T}}[\cdot\,|\,\cdot] := \sup_{P \in \mathscr{P}^{\mathrm{H}}_{\mathscr{T}}} \mathbb{E}_P[\cdot\,|\,\cdot].$$

(II) Sets of general processes

Another type of imprecise Markov chain is $\mathscr{P}^{\mathrm{I}}_{\mathscr{T}}$, which is called an imprecise Markov chain under **epistemic irrelevance**. It is the **set of all stochastic processes** P for which, for all $n \in \mathbb{N}_0$ and all $x_0, \ldots, x_n \in \mathscr{X}$,

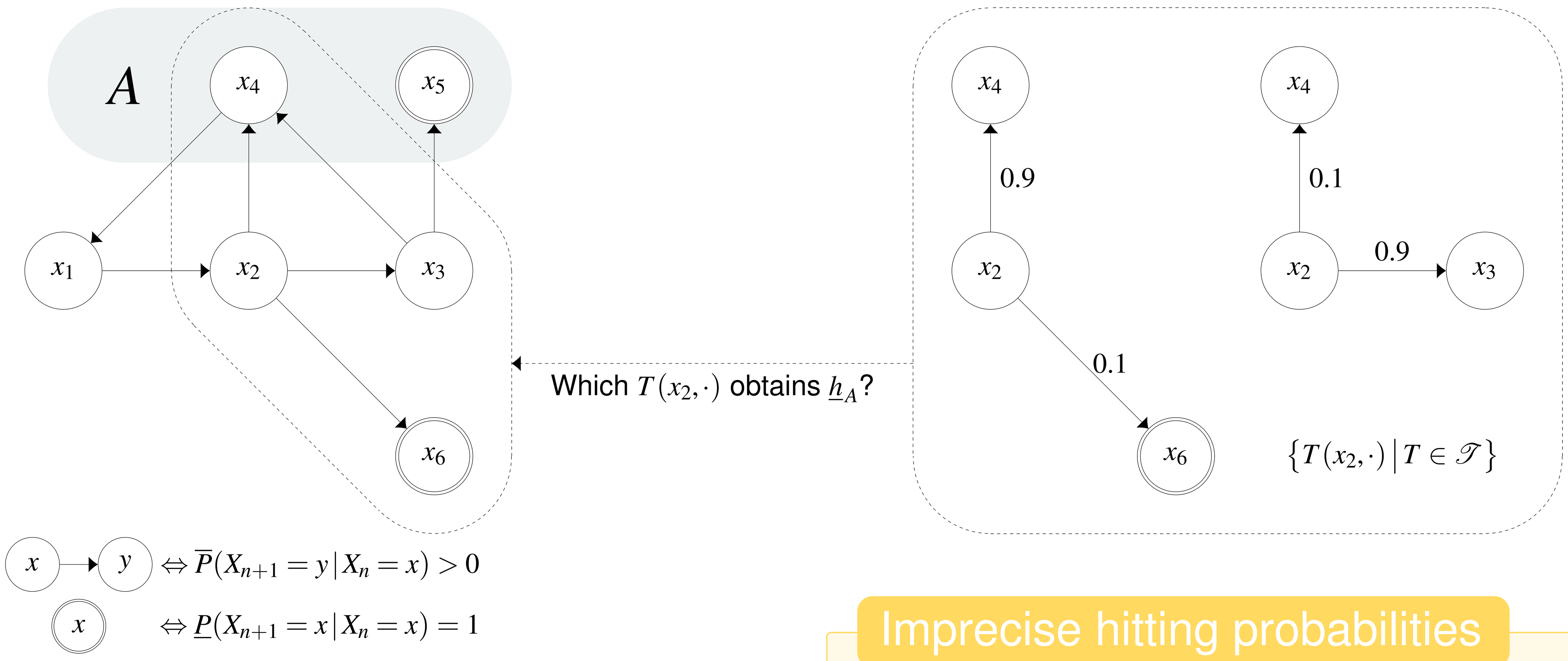
$$\exists T \in \mathscr{T} : \forall y \in \mathscr{X} : \quad P(X_{n+1} = y \,|\, X_{0:n} = x_{0:n}) = T(x_n, y).$$

Its lower expectation and upper expectation operators are defined as

$$\underline{\mathbb{E}}^{\mathrm{I}}_{\mathscr{T}}[\cdot\,|\,\cdot] := \inf_{P \in \mathscr{P}^{\mathrm{I}}_{\mathscr{T}}} \mathbb{E}_P[\cdot\,|\,\cdot] \quad \text{and} \quad \overline{\mathbb{E}}^{\mathrm{I}}_{\mathscr{T}}[\cdot\,|\,\cdot] := \sup_{P \in \mathscr{P}^{\mathrm{I}}_{\mathscr{T}}} \mathbb{E}_P[\cdot\,|\,\cdot].$$

(III) Game-theoretic models

An entirely different formalisation of imprecise Markov chains uses the **game-theoretic probability** framework that was popularised by Shafer and Vovk. Its characterisation is based on the theory of **sub- and supermartingales**, and exceeds the scope of this work. The corresponding (global) upper expectations $\overline{\mathbb{E}}^{\mathrm{V}}_{\mathscr{T}}[\cdot\,|\,\cdot]$ are derived from **local uncertainty models** that agree with $\overline{T}$, and the lower expectations are obtained by the conjugacy relation $\underline{\mathbb{E}}^{\mathrm{V}}_{\mathscr{T}}[\cdot\,|\,\cdot] = -\overline{\mathbb{E}}^{\mathrm{V}}_{\mathscr{T}}[-\cdot\,|\,\cdot]$.



Imprecise hitting times

We proved that the lower and upper expected hitting times are **the same for all types** of imprecise Markov chains. Moreover, they are **obtained by a homogeneous Markov chain** with transition matrix $T \in \mathscr{T}$.

Equality. There exists some $P \in \mathscr{P}^{\mathrm{H}}_{\mathscr{T}}$ such that

$$\underline{\mathbb{E}}^{\mathrm{V}}_{\mathscr{T}}[H_A \,|\, X_0] = \underline{\mathbb{E}}^{\mathrm{I}}_{\mathscr{T}}[H_A \,|\, X_0] = \underline{\mathbb{E}}^{\mathrm{H}}_{\mathscr{T}}[H_A \,|\, X_0] = \mathbb{E}_P[H_A \,|\, X_0],$$

and, similarly, there exists some $P \in \mathscr{P}^{\mathrm{H}}_{\mathscr{T}}$ such that

$$\overline{\mathbb{E}}^{\mathrm{V}}_{\mathscr{T}}[H_A \,|\, X_0] = \overline{\mathbb{E}}^{\mathrm{I}}_{\mathscr{T}}[H_A \,|\, X_0] = \overline{\mathbb{E}}^{\mathrm{H}}_{\mathscr{T}}[H_A \,|\, X_0] = \mathbb{E}_P[H_A \,|\, X_0].$$

Characterisation. For any imprecise Markov chain, the vector $\underline{h}_A$ of lower expected hitting times, defined as $\underline{h}_A(x) := \underline{\mathbb{E}}[H_A \,|\, X_0 = x]$ for all $x \in \mathscr{X}$, is the **minimal non-negative solution** to the (non-linear) system

$$\underline{h}_A = \mathbb{I}_{A^c} + \mathbb{I}_{A^c} \cdot \underline{T}\, \underline{h}_A.$$

Similarly, the vector $\overline{h}_A$ of upper expected hitting times, defined as $\overline{h}_A(x) := \overline{\mathbb{E}}[H_A \,|\, X_0 = x]$ for all $x \in \mathscr{X}$, is the minimal non-negative solution to the system

$$\overline{h}_A = \mathbb{I}_{A^c} + \mathbb{I}_{A^c} \cdot \overline{T}\, \overline{h}_A.$$

Imprecise hitting probabilities

We proved that the lower and upper hitting probabilities are **the same for all types** of imprecise Markov chains.

Equality. There exists some $P \in \mathscr{P}^{\mathrm{H}}_{\mathscr{T}}$ such that

$$\underline{\mathbb{E}}^{\mathrm{V}}_{\mathscr{T}}[G_A \,|\, X_0] = \underline{\mathbb{E}}^{\mathrm{I}}_{\mathscr{T}}[G_A \,|\, X_0] = \underline{\mathbb{E}}^{\mathrm{H}}_{\mathscr{T}}[G_A \,|\, X_0] = \mathbb{E}_P[G_A \,|\, X_0].$$

Moreover, it holds that

$$\overline{\mathbb{E}}^{\mathrm{V}}_{\mathscr{T}}[G_A \,|\, X_0] = \overline{\mathbb{E}}^{\mathrm{I}}_{\mathscr{T}}[G_A \,|\, X_0] = \overline{\mathbb{E}}^{\mathrm{H}}_{\mathscr{T}}[G_A \,|\, X_0],$$

but these are **not always obtained** by a $P \in \mathscr{P}^{\mathrm{H}}_{\mathscr{T}}$.

Characterisation. For any imprecise Markov chain, the vector $\underline{p}_A$ of lower hitting probabilities, defined as $\underline{p}_A(x) := \underline{\mathbb{E}}[G_A \,|\, X_0 = x]$ for all $x \in \mathscr{X}$, is the **minimal non-negative solution** to the (non-linear) system

$$\underline{p}_A = \mathbb{I}_A + \mathbb{I}_{A^c} \cdot \underline{T}\, \underline{p}_A.$$

Similarly, the vector $\overline{p}_A$ of upper hitting probabilities, defined as $\overline{p}_A(x) := \overline{\mathbb{E}}[G_A \,|\, X_0 = x]$ for all $x \in \mathscr{X}$, is the minimal non-negative solution to the system

$$\overline{p}_A = \mathbb{I}_A + \mathbb{I}_{A^c} \cdot \overline{T}\, \overline{p}_A.$$

We refer to `arXiv:1905.08781` for technical details.

UTOPIÆ Uncertainty Treatment and Optimisation in Aerospace Engineering

Handling the unknown at the edge of tomorrow

HORIZON 2020

MARIE CURIE ACTIONS

This work was partially supported by H2020-MSCA-ITN-2016 UTOPIAE, grant agreement 722734.