Discrete Random Processes (EDRP)

Lecture 10

Continuous-time Markov chains

Continuous-time Markov chains

We will extend the Markov chain model to *continuous* time. A continuous-time process allows one to model not only the transitions between states, but also the duration of time in each state. The central Markov property continues to hold -

given the present, past and future are independent.

The Markov property is a form of memorylessness. This leads to the exponential distribution. In a continuous-time Markov chain, when a state is visited, the process stays in that state for an exponentially distributed length of time before moving to a new state.

If one just watches the sequence of states that are visited, ignoring the length of time spent in each state, the process looks like a discrete-time Markov chain.

Markov property

In discrete time we formulated the Markov property as: for any possible values of $j, i, i_{n-1}, \ldots, i_0$,

$$\mathbb{P}\left(X_{n+1} = j \middle| X_n = i, X_{n-1} = i_{n-1}, \ldots, X_0 = i_0\right) \\ = \mathbb{P}\left(X_{n+1} = j \middle| X_n = i\right).$$

In continuous time:

Definition 1

We say that $(X_t)_{t \geq 0}$ is a **(continuous-time) Markov chain** if for any $t \geq 0$, $0 \leq s_0 < s_1 < \ldots < s_n < s$ and possible states $i_0, \ldots, i_n, i, j$, we have

$$\mathbb{P}\left(X_{t+s} = j \middle| X_s = i, X_{s_n} = i_n, \ldots, X_{s_0} = i_0\right) = \mathbb{P}\left(X_{t+s} = j \middle| X_s = i\right).$$

Markov property - cont'd

So for any $t \geq 0$, $0 \leq s_0 < s_1 < \ldots < s_n < s$ and any states $i_0, \ldots, i_n, i, j$,

$$\mathbb{P}\left(X_{t+s} = j \middle| X_s = i, X_{s_n} = i_n, \ldots, X_{s_0} = i_0\right) = \mathbb{P}\left(X_{t+s} = j \middle| X_s = i\right).$$

In words, given the present state, the rest of the past is *irrelevant* for predicting the future.

Note that built into the definition is the fact that the probability of going from i at time s to j at time $s+t$ only depends on the difference t in the times (this means, we assume **time-homogeneity**).

Transitions

Discrete time Markov chains were described by giving their transition probabilities

$$p_{i,j} = \mathbb{P}\left(X_{n+1} = j \middle| X_n = i\right) = \mathbb{P}\left(X_1 = j \middle| X_0 = i\right)$$

- the probability of jumping from i to j in one step.
In continuous time there is no first time $t > 0$, so we introduce for each $t > 0$ a transition probability

$$p_{i,j}(t) = \mathbb{P}\left(X_{t+s} = j \middle| X_s = i\right) = \mathbb{P}\left(X_t = j \middle| X_0 = i\right).$$

Let also

$$p_i(t) = \mathbb{P}(X_t = i).$$

Transitions - cont'd

So

- $$p_{i,j}(t) = \mathbb{P}\left(X_{t+s} = j \middle| X_s = i\right) = \mathbb{P}\left(X_t = j \middle| X_0 = i\right)$$

  is the probability that a Markov chain that is presently in state i will be in state j after an additional time t,

- $$p_i(t) = \mathbb{P}(X_t = i)$$

  is the probability that a Markov chain is in state i at time t.

Clearly, for all t

$$\sum_i p_i(t) = 1,$$

and

$$\sum_j p_{i,j}(t) = 1.$$

Chapman–Kolmogorov equations

In continuous time, as in discrete time, the transition probability satisfies Chapman-Kolmogorov equation

$$p_{i,j}(s+t) = \sum_k p_{i,k}(s) p_{k,j}(t).$$

An intuition behind this equation is the following: in order for the chain to go from i to j in time $s+t$, it must be in some state k at time s, and the Markov property implies that the two parts of the journey are independent.

Transition function

If we define $\mathbf{P}(t)$ as the matrix of the $p_{i,j}(t)$, that is,

$$\mathbf{P}(t) = \begin{bmatrix} p_{1,1}(t) & p_{1,2}(t) & p_{1,3}(t) & \dots \\ p_{2,1}(t) & p_{2,2}(t) & p_{2,3}(t) & \dots \\ p_{3,1}(t) & p_{3,2}(t) & p_{3,3}(t) & \dots \\ \dots & \dots & \dots & \dots \end{bmatrix},$$

then the Chapman-Kolmogorov equation

$$p_{i,j}(s+t) = \sum_k p_{i,k}(s) p_{k,j}(t).$$

becomes

$$\mathbf{P}(t+s) = \mathbf{P}(t)\mathbf{P}(s).$$

The matrix function $\mathbf{P}(t)$ is called **the transition function**.

An example - Poisson process

A Poisson process $(N_t)_t$ with parameter λ is a continuous-time Markov chain. The Markov property holds as a consequence of stationary and independent increments.

Holding times

By homogeneity, when a Markov chain visits state $i \in S$, its forward evolution from that time onward behaves the same as the process started in i at time $t = 0$. Time-homogeneity and the Markov property characterize the distribution of the length of time that a continuous-time chain stays in state i before transitioning to a new state.

Definition 2

Holding time at state $i \in S$, denoted T_i, is the length of time that a continuous-time Markov chain started in i stays in i before transitioning to a new state.

Theorem 3

Every T_i is exponentially distributed.

Holding times - cont'd

For each state i, let q_i be the parameter of the exponential distribution for the holding time T_i. Thus

$$T_i \sim \mathrm{Exp}\,(q_i)\,, \quad i \in S.$$

We will be assuming that

$$0 < q_i < +\infty.$$

In particular

$$\mathbb{E}\,T_i = \frac{1}{q_i}.$$

Holding times - cont'd

In general, it is possible to define a continuous-time process with $q_i = 0$ or $+\infty$:

- $q_i = 0$ means that when i is visited, the process never leaves (one says that i is an **absorbing state**),

- $q_i = +\infty$ means that the process leaves i immediately upon entering i (this would allow for infinitely many transitions in a finite interval - such a process is called **explosive**.)

Dynamics of continuous-time Markov chain

The evolution of a continuous-time Markov chain can be described as follows.

- Starting from i, the process stays in i for an exponentially distributed length of time, on average $1/q_i$ time units.
- Then, it hits a new state $j \neq i$, with some probability $p_{i,j}$.
- The process stays in j for an exponentially distributed length of time, on average $1/q_j$ time units.
- It then hits a new state $k \neq j$, with some probability $p_{j,k}$,
- ... and so on.

Embedded chain

The transition probabilities $p_{i,j}$ describe the discrete transitions from state to state.

If we ignore time, and just watch state to state transitions, we see a sequence

$$Y_0, Y_1, \dots,$$

where

Y_n is the n-th state visited by the continuous time process $(X_t)_t$.

Definition 4

The sequence $Y_0, Y_1, \dots$ is a discrete-time Markov chain called the **embedded chain**.

Embedded chain - transition matrix

Let $\tilde{\mathbf{P}}$ be the transition matrix for the embedded chain. That is,

$$(\tilde{\mathbf{P}})_{i,j} = p_{i,j}$$

for all $i, j \in S$.

Then, $\tilde{\mathbf{P}}$ is a *stochastic* matrix (all its entries are non-negative and in each row add up to 1), whose diagonal entries are 0.

Examples of embedded chain

Example 1

The embedded chain for the Poisson process

Example 2

The embedded chain for a two-state chain

Jump rates

For $j \neq i$ define

$$q_{i,j} = \lim_{h \to 0} \frac{p_{i,j}(h)}{h}.$$

If this limit exists, it is just the derivative $p'_{i,j}(0)$ of the transition probability $p_{i,j}(t)$ at zero.

Definition 5

We will call $q_{i,j}$ the **jump rate** (or **transition rate**) from i to j.

One can prove that the parameter q_i of the exponential distribution for the holding time T_i (which is the rate at which the chain leaves i) is related to the jump rates through the equation:

$$q_i = \sum_{j \neq i} q_{i,j}.$$

Jump rates - cont'd

One can also prove that for the embedded chain, transition probabilities $p_{i,j}$ satisfy

$$p_{i,j} = \frac{q_{i,j}}{q_i}, \quad i \neq j.$$

Example - the general three-state continuous Markov chain

Generator

Define a matrix $\mathbf{Q}$ by setting:

$$(\mathbf{Q})_{i,j} = q_{i,j} = \lim_{h \to 0} \frac{p_{i,j}(h)}{h} = p'_{i,j}(0), \quad i \neq j,$$

and

$$(\mathbf{Q})_{i,i} = -q_i.$$

Definition 6

The matrix $\mathbf{Q}$ is called the **generator** or **infinitesimal generator**.

Generator - cont'd

The generator is the most important matrix for continuous-time Markov chains.

We derived $\mathbf{Q}$ from the transition function $\mathbf{P}(t)$. However, in a modeling context one typically starts with $\mathbf{Q}$, identifying the transition rates $q_{i,j}$ based on the qualitative and quantitative dynamics of the process. The transition function and related quantities are derived from $\mathbf{Q}$.

Clearly, the generator is not a stochastic matrix:

- diagonal entries are negative,
- entries can be greater than 1,
- and rows sum to 0.

Example - the generator for a Poisson process with parameter λ

Generators - a summary

For discrete-time Markov chains, there is no generator matrix and the probabilistic properties of the stochastic process are captured by the transition matrix $\mathbf{P}$.

For continuous-time Markov chains the generator matrix $\mathbf{Q}$ gives a complete description of the dynamics of the process. The distribution of any finite subset of the $(X_t)_t$, and all probabilistic quantities of the stochastic process, can, in principle, be obtained from the infinitesimal generator and the initial distribution.