This is a draft version of the lecture notes. We aim to keep improving it but at the current stage it is most likely far from perfect. Please contact us if you notice any typos, errors, subtle points, or if you have any questions or suggestions for improvements.

`maciej.lisicki[at]fuw.edu.pl`
`milosz.panfil[at]fuw.edu.pl`

Maciej Lisicki & Miłosz Panfil
October 29, 2022

6 Markov processes

Markov processes are a particularly important class of stochastic processes for two reasons. Firstly, as we will discuss, in most cases it is possible to describe a physical system by a Markov process with appropriately chosen variables. Secondly, they are mathematically tractable due to the specific structure of the constituting distribution functions, which simplify the description considerably. On an intuitive level, Markov processes are those with short memory — the state of the system is determined by its previous configuration but not by its entire history. In this Chapter, we shall formalise this notion to develop mathematical tools for practical treatment of such processes.

6.1 The Markov property

Consider the stochastic process $Y(X,t)$, in which X is a random variable and the parameter $t \in \mathbb{R}$ is ordered and effectively becomes time. The process will be called Markovian, if for a sequence of times $t_1 < t_2 < \ldots < t_{r+1}$ for $r = 2, 3, \ldots$ the conditional probability $P_{1|r}$ reduces to a two-point conditional probability $P_{1|1}$, namely

$$P_{1|r}(y_{r+1}, t_{r+1}|y_1, t_1; \ldots; y_r, t_r) = P_{1|1}(y_{r+1}, t_{r+1}|y_r, t_r). \tag{6.1}$$

Practically, this means that only the last point preceding the event of interest matters for the conditional probability and all higher-order functions simplify to $P_{1|1}$. Therefore, to fully characterise a Markov process, we only need $P_1(y_1, t)$, called now the *probability of the process*, and $P_{1|1}(y_1, t_1|y_2, t_2)$, for $t_2 > t_1$, called the *transition function*. The transition function is sometimes denoted simply as $T(y_2, t_2|y_1, t_1)$. Note that all other probability distributions can be constructed out of these two functions.

6.2 Chapman-Kolmogorov-Smoluchowski equation

The probability densities of a Markov process still have to satisfy the consistency conditions, mentioned in Sec. 5. In this case, the relevant conditions read

(a) $P_1 \geq 0$, $P_{1|1} \geq 0$,

(b) is satisfied automatically,

(c) $\displaystyle\int P_1(y,t)\mathrm{d}y = 1.$

The last condition, (d), is the most interesting, since it will give us a relationship between the transition probabilities. Let us first examine it in the case of $r = 2$. On one hand, it defines the probability of the process as a marginal distribution

$$P_1(y_2, t_2) = \int P_2(y_2, t_2; y_1, t_1)\mathrm{d}y_1. \tag{6.2}$$

On the other hand, for $t_2 \geq t_1$ we can rewrite P_2 under the integral in terms of the transition probability

$$P_1(y_2, t_2) = \int P_{1|1}(y_2, t_2|y_1, t_1)P_1(y_1, t_1)\mathrm{d}y_1. \tag{6.3}$$

It is now clear from this equation that having P_1 at some point $t = t_1$ and the transition probability $P_{1|1}$, we can construct P_1 at a later time $t = t_2$. On the other hand, if $t_2 < t_1$, we have to change the order in $P_{1|1}$ to obtain

$$P_1(y_2, t_2) = \int P_{1|1}(y_1, t_1|y_2, t_2)P_1(y_2, t_1)\mathrm{d}y_1, \tag{6.4}$$

both sides of which we can divide by $P_1(y_2, t_2)$ to obtain

$$\int P_{1|1}(y_1, t_1|y_2, t_2)\mathrm{d}y_1 = 1, \tag{6.5}$$

another condition to be satisfied by the transition probability. Its interpretation is clear – if we add all outcomes (all trajectories passing through a given point), their probabilities must sum to unity.

Now take $r = 3$ and $t_3 > t_2 > t_1$. The two-point distribution is again written as a marginal distribution of P_3 as

$$P_2(y_1, t_1; y_3, t_3) = \int P_3(y_1, t_1; y_2, t_2; y_3, t_3)\mathrm{d}y_2. \tag{6.6}$$

We can rewrite the left-hand side as

$$P_2(y_1, t_1; y_3, t_3) = P_{1|1}(y_3, t_3|y_1, t_1)P_1(y_1, t_1), \tag{6.7}$$

while the right-hand side is expanded as

$$\begin{aligned} \int P_3(y_1, t_1; y_2, t_2; y_3, t_3)\mathrm{d}y_2 &= \int P_{1|2}(y_3, t_3|y_2, t_2; y_1, t_1)P_2(y_2, t_2; y_1, t_1)\mathrm{d}y_2 \\ &= \int P_{1|1}(y_3, t_3|y_2, t_2)P_{1|1}(y_2, t_2|y_1, t_1)\mathrm{d}y_2, \end{aligned} \tag{6.8}$$

where in the last equality we used the Markov property (6.1) to decompose all distributions into the product of transition probabilities. As a result, we arrive at an integral equation relating transition probabilities at different points,

$$P_{1|1}(y_3, t_3|y_1, t_1) = \int P_{1|1}(y_3, t_3|y_2, t_2)P_{1|1}(y_2, t_2|y_1, t_1)\mathrm{d}y_2, \tag{6.9}$$

called the Chapman-Kolmogorov-Smoluchowski (CKS) equation. It can be given the following interpretation – For a Markov process, the transition probability of jumping from

(y_1, t_1) to (y_3, t_3) is given by a sum of probabilities of jumping through *any* intermediate point (y_2, t_2). Note that this is exactly the meaning of Eq. (4.25) in the discussed model of a 3d random walk, so the assumption of Markovianity lies buried in this equation already. When $t_2 = t_1$, we find that $P_{1|1}(y_2, t_2|y_1, t_1) = \delta(y_2 - y_1)$.

In physics, we often use the homogeneous Markov process, in which case we write

$$P_{1|1}(y_2, t_2|y_1, t_1) = T_{t_2 - t_1}(y_2|y_1). \tag{6.10}$$

A homogeneous process satisfies

(i) $T_0(y_2|y_1) = \delta(y_2 - y_1)$,

(ii) $\int T_\tau(y_2|y_1)\mathrm{d}y_2 = 1$,

(iii) for $t_1, t_2 > 0$ we write the CKS equation as

$$T_{t_1+t_2}(y_2|y_1) = \int T_{t_2}(y_2|y_3)T_{t_1}(y_3|y_1)\mathrm{d}y_3, \tag{6.11}$$

which is sometimes abbreviated to the shorthand notation

$$T_{t_1+t_2} = T_{t_2}T_{t_1},$$

where there is no time reversal symmetry.

We now see that to fully describe a Markov process, we need the transition probability T_t and the initial distribution $P_1(t = 0)$. Then we obtain the distribution at any time by simply

$$P_1(t) = T_t P_1(0).$$

We note here for future reference that for a process to be stationary, we need to have $P_1 = T_t P_1$.

Markov chains Consider a simple case, in which the system has a discrete number of states N and evolves in discrete time steps $t = 0, 1, 2, \ldots$. The transition probability becomes then an $N \times N$ matrix $T_t(y_i|y_j)$, with $T_0(y_i|y_j) = \delta_{ij}$. Such a process is called a *finite Markov chain*. The Markov property (6.11) becomes then simply

$$T_\tau = (T_1)^\tau, \qquad \tau = 0, 1, 2, \ldots$$

The study of finite Markov chains boils down to the study of powers of $N \times N$ matrices, such that (a) their elements are non-negative, and (b) each column adds up to unity. Such matrices are called *stochastic matrices*, as opposed to *random matrices*, where each element of the matrix is a random variable. Stochastic matrices are an interesting research branch of mathematics but in this course we move on to the more general case of continuous time, where we will discuss further properties of Markov processes.