

12.2: General Description

12.2.1 Markov Processes

More generally, suppose we have a system possessing a discrete set of states, which can be labeled by an integer $0, 1, 2, \dots$. A **Markov process** is a set of probabilistic rules that tell us how to choose a new state of the system, based on the system's current state. If the system is currently in state n , then the probability of choosing state m on the next step is denoted by $P(m|n)$. We call this the "transition probability" from state n to state m . By repeatedly applying the Markov process, we move the system through a random sequence of states, $\{n^{(0)}, n^{(1)}, n^{(2)}, n^{(3)}, \dots\}$, where $n^{(k)}$ denotes the state on step k . This kind of random sequence is called a **Markov chain**.

There is an important constraint on the transition probabilities of the Markov process. Because the system must transition to *some* state on each step,

$$\sum_m P(m|n) = 1 \quad \text{for all } n \in \{0, 1, \dots\}. \quad (12.2.1)$$

Next, we introduce the idea of **state probabilities**. Suppose we look at the ensemble of all possible Markov chains which can be generated by a given Markov process. Let $\{p_0^{(k)}, p_1^{(k)}, p_2^{(k)}, \dots\}$ denote the probabilities for the various states, $n = 0, 1, 2, \dots$, on step k . Given these, what are the probabilities for the various states on step $k+1$? According to Bayes' theorem, we can write $p_m^{(k+1)}$ as a sum over conditional probabilities:

$$p_m^{(k+1)} = \sum_n P(m|n) p_n^{(k)}. \quad (12.2.2)$$

This has the form of a matrix equation:

$$\begin{bmatrix} p_0^{(k+1)} \\ p_1^{(k+1)} \\ \vdots \end{bmatrix} = \begin{bmatrix} P(0|0) & P(0|1) & \cdots \\ P(1|0) & P(1|1) & \cdots \\ \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} p_0^{(k)} \\ p_1^{(k)} \\ \vdots \end{bmatrix}, \quad (12.2.3)$$

where the matrix on the right-hand side is called the **transition matrix**. Each element of this matrix is a real number between 0 and 1; furthermore, because of the aforementioned conservation of transition probabilities, each column of the matrix sums to 1. In mathematics, matrices of this type are called "left stochastic matrices".

12.2.2 Stationary Distribution

A **stationary distribution** is a set of state probabilities $\{\pi_0, \pi_1, \pi_2, \dots\}$, such that passing through one step of the Markov process leaves the probabilities unchanged:

$$\pi_m = \sum_n P(m|n) \pi_n. \quad (12.2.4)$$

By looking at the equivalent matrix equation, we see the vector $[\pi_0; \pi_1; \pi_2; \dots]$ must be an eigenvector of the transition matrix, with eigenvalue 1. It turns out that there is a mathematical theorem (the [Perron–Frobenius theorem](#)) which states every left stochastic matrix has an eigenvector of this sort. Hence, every Markov process possesses a stationary distribution. Stationary distributions are the main reasons we are interested in Markov processes. In physics, we are often interested in using Markov processes to model thermodynamic systems, such that a stationary distribution represents the distribution of thermodynamic micro-states under thermal equilibrium. (We'll see an example in the next section.) Knowing the stationary distribution, we can figure out all the thermodynamic properties of the system, such as its average energy.

In principle, one way to figure out the stationary distribution is to construct the transition matrix, solve the eigenvalue problem, and pick out the eigenvector with eigenvalue 1. The trouble is that we are often interested in systems where the number of possible states is *huge*—in some cases, larger than the number of atoms in the universe! In such cases, it is not possible to explicitly generate the transition matrix, let alone solve the eigenvalue problem.

We now come upon a happy and important fact: for a huge class of Markov processes, the distribution of states within a sufficiently long Markov chain will converge to the stationary distribution. Hence, in order to find out about the stationary distribution, we

simply need to generate a long Markov chain, and study its statistical properties.

This page titled [12.2: General Description](#) is shared under a [CC BY-SA 4.0](#) license and was authored, remixed, and/or curated by [Y. D. Chong](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.