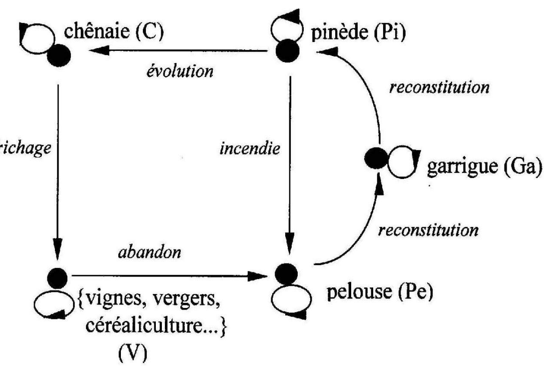
Course 2 : Evolution towards a stationary distribution

We have already encountered the notion of a *stationary distribution* chain Markov is a distribution system between its various states having the property to remain unchanged over time. Specifically, if the proportions of the different states at the initial moment are those of a stationary distribution, then these proportions are not later modified by the dynamics of the Markov chain. One can therefore understand the stationary distributions as a kind of balance to the system. A Markov chain may not have any sort of balance, or have one or more. But these equilibria, where they exist, do not all have the same relevance. Those who will really count are the balances to which one approaches ineluctably as time elapses, and no matter what the initial distribution system.

In this lesson we will first observe the existence of such a balance on a example and then tell us how we can ensure the existence of such a stationary distribution to which the Markov chain evolves and how it can be calculated.

1 Example of a Mediterranean ecosystem

Consider the following dynamics that models the evolution of a Mediterranean ecosystem [1]. Originally the Mediterranean forest, on limestone at low altitude, was certainly dominated by oaks (white oak). But the action of man has eradicated these primeval forests for their substitute rangeland, orchards, ... Then the abandonment of all farm work instead of driving to the restoration of these natural oak has often favored the establishment of another species, the Aleppo pine, after passing through a state of scrubland. But these bits of substitution, highly flammable, suffer recurrently the passage of fire (arson or not) and get stuck in a perpetual rebuilding. Here is the diagram corresponding points and arrows and the transition matrix of this Markov chain with five states $S = \{C, V, Pe, Ga, Pi\}$ and transition matrix $\mathbb{P}$.



$$\mathbb{P} = \begin{pmatrix} 0,8 & 0,2 & 0 & 0 & 0 \\ 0 & 0,7 & 0,3 & 0 & 0 \\ 0 & 0 & 0,4 & 0,6 & 0 \\ 0 & 0 & 0 & 0,2 & 0,8 \\ 0,1 & 0 & 0,25 & 0 & 0,65 \end{pmatrix}$$

Simple observation of the diagram points and arrows, which shows the trajectories that pass indefinitely from one state to another, not enough to understand the long-term ecosystem. Or it may be interesting to understand this evolution for example, be able to influence it (to try to limit the spread of the pine forest for example). Given an initial distribution π_0 that would have such determined from the observation of a specific parcel (by calculating the proportion of each state), we saw that we can calculate the evolution of this distribution after step π_1 by performing the product $\pi_0\mathbb{P}$, after two steps, by multiplying the result obtained again by $\mathbb{P}$, $\pi_2 = \pi_1\mathbb{P}$,, and so on. But this does not necessarily give the idea of limiting behavior, if any. Note that $\pi_2 = \pi_1\mathbb{P} = (\pi_0\mathbb{P})\mathbb{P} = \pi_0(\mathbb{P}\mathbb{P}) = \pi_0\mathbb{P}^2$ and more generally $\pi_k = \pi_0\mathbb{P}^k$.

Now, if we observe the successive powers of the transition matrix $\mathbb{P}$ (what can be done easily with Scilab example), we can see that when k increases, the matrix $\mathbb{P}^k$ is gradually stabilizing and takes the form of a

[1] This example is from the book Modeling and Simulation of Ecosystems, P. Coquillard and D. Hill, Masson 1997.

matrix whose rows are all (almost) equal. For example we obtain the following matrix for $\mathbb{P}^{40}$:

$$\mathbb{P}^{40} = \begin{pmatrix} 0,17520 & 0,11680 & 0,20437 & 0,15327 & 0,35034 \\ 0,17517 & 0,11680 & 0,20438 & 0,15328 & 0,35035 \\ 0,17551 & 0,11678 & 0,20438 & 0,15328 & 0,35037 \\ 0,17517 & 0,11678 & 0,20438 & 0,15328 & 0,35037 \\ 0,17518 & 0,11678 & 0,20437 & 0,15328 & 0,36036 \end{pmatrix}.$$

One can then make the following general observation: a stochastic matrix (necessarily square) $\mathbb{P} = \begin{pmatrix} L_1 \\ ... \\ L_n \end{pmatrix}$ which all lines are equal to a single row vector L^* satisfies the equality $L^*\mathbb{P} = L^*$.

Since L^* is also a vector of coefficients whose sum is 1 (as a line of a stochastic matrix), this property is a good candidate for L^* to be the stationary distribution that is sought is to say the equilibrium distribution towards which the system evolves.

We will explain why this is indeed the desired stationary distribution.

In the example, note that the lines of $\mathbb{P}^{40}$ is not exactly identical, it is not going to experience as well as an approximation to the equilibrium distribution. We can nevertheless say that, whatever the proportions of the five states at the initial time, this model predicts that over time, these proportions will settle at around $17,5\%$ oak, $11,7\%$ of vines , orchards, ..., $20,5\%$ grass, $15,3\%$ of scrubland and finally 35% of pine forest. in particular, we note that according to this model, the oak will be reduced to less than one fifth of the space occupied while in the pine forest conquers almost a third.

2 The theory of Perron-Frobenius

How to find the stationary distribution limit of a Markov chain when there is applied the mathematical theory named after its two German inventors, Ferdinand Georg Frobenius (1949-1917) and Oskar Perron (1880-1975) and for positive matrix which one of the powers is strictly positive. These matrices are called *primitive* matrices. We will study them in detail during the next class.

According to this theory, any stochastic matrix $\mathbb{P}$ is primitive has a strictly positive stochastic vector, $\pi^\infty > 0$, which verifies, for all π_0 stochastic (ie positive, sum to 1),

$$\lim_{k \to +\infty} \pi_0 \mathbb{P}^k = \pi^\infty.$$

Applied to the case of a China Markov, this result ensures that if the transition matrix of the chain is primitive, then we are assured that whatever the initial distribution π_0 system, the distribution of states will evolve under the action the dynamics of the Markov chain, to a stationary distribution π^∞ which is a balance to the system.

The Perron-Frobenius theory does not apply only to stochastic matrices, but more generally to all positive matrix. We will see other applications later.

One of the difficulties of its application is to check its main hypothesis, the property for the matrix to be primitive. Indeed, if found, by computing powers of the matrix, one of them is strictly positive, we know that the matrix is primitive. But when all the forces calculated contain zero coefficients, then we do not know if the matrix is primitive or not A recipe can still be useful:
A matrix $\mathbb{P}$ of size $n \times n$ is primitive if and only if $\mathbb{P}^{n^2-2n+2} > 0$.

In practice, if one wants to determine the future evolution of a system modeled by a Markov chain, we proceed as follows:

1. Establish (if that is the case) that the transition matrix is a primitive matrix.
2. With the help of a scientific computing software (eg Scilab), high compute power of the transition matrix, for example $\mathbb{P}^{40}$ $\mathbb{P}^{100}$.
3. This power with all its lines (almost) equal, this line is a stationary distribution (you can verify if you will).
4. The Perron-Frobenius theory then allows to assert that the dynamics considered is changing the system, whatever the initial distribution to equilibrium given by the stationary distribution.