BASICS ON MARKOV CHAINS

We consider successive trials with possible outcomes $E_1, E_2, \ldots$ .Let us define $(X_n)_{n\geq 1}$ the sequence of random variables such as:

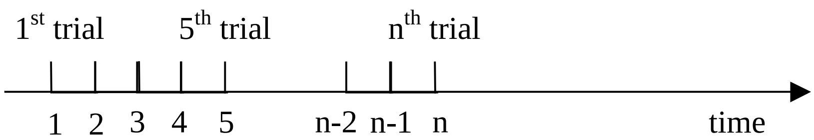
$$\{X_n = k\} = \{E_k \text{ is the outcome of the } n^{th} \text{ trial}\}$$

$(X_n)_{n\geq 1}$ is said to be a Markov chain if and only if:

$$\forall\, n, i_1, i_2, \ldots, i_n \quad P\left\{X_n = i_n / \bigcap_{k=1}^{n-1} \{X_k = i_k\}\right\} = P\{X_n = i_n / X_{n-1} = i_{n-1}\}$$

This relation is known as the Markov property.

If we associate a time scale to the sequence of trials,



n corresponds to the future
n-1 corresponds to the present
1 to(n-2) corresponds to the past

Then, the Markov property can be stated as follows:

$$P\{\text{Future/Present and Past}\} = P\{\text{Future / Present}\}$$

Furthermore $(X_n)_{n\geq 1}$ is homogeneous if and only if:

$\forall\, n, j, i \quad P\{X_n = j / X_{n-1} = i\}$ does not depend on n..

So we can denote $p_{ij} = P\{X_n = j / X_{n-1} = i\}$

These conditional probabilities p_{ij} are called transition probabilities. If the number of states is finite (for instance n_0), they can be arranged in a transition probability matrix T so that the first subscript (i) stands for row and the second (j) for column. T is a square matrix ($n_0 \times n_0$) with non negative elements and unit row sums.

$$\forall\, i, j \quad 0 \le p_{ij} \le 1 \quad and \quad \forall\, i \quad \sum_{j=1}^{n_0} p_{ij} = 1$$

$$T = \begin{array}{c} \\ E_1 \\ E_2 \\ \\ E_{n0} \end{array} \begin{array}{cccc} E_1 & E_2 & & E_{n0} \end{array} \\ \begin{pmatrix} p_{11} & p_{12} & & p_{1n0} \\ p_{21} & p_{22} & & p_{2n0} \\ & & & \\ p_{n01} & p_{n02} & & p_{n0n0} \end{pmatrix}$$

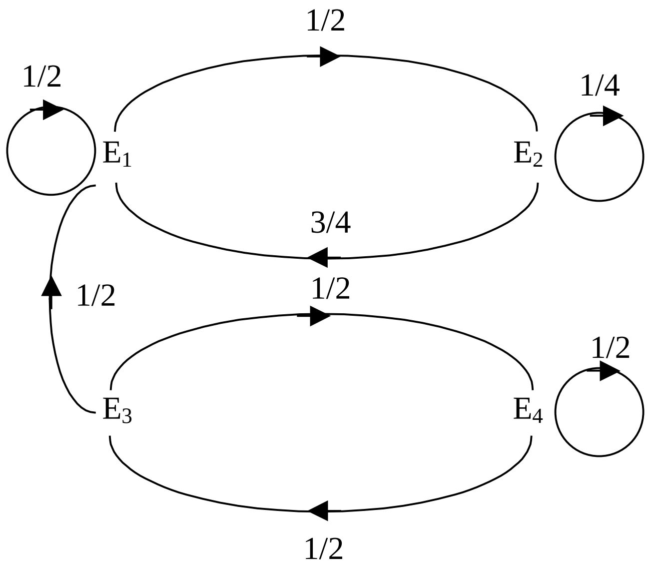
Another way of describing a Markov chain consists of sketching the transition graph: an array links a state E_i to a state E_j when a transition (in one step) from E_i to E_j is possible, i.e. when $p_{ij} > 0$. So each term different from zero in the transition probability matrix is associated with an array in the graph.

Example

Let us consider a Markov chain whose probability transition matrix is given by:

$$T = \begin{array}{c|cccc} & E_1 & E_2 & E_3 & E_4 \\ \hline E_1 & 1/2 & 1/2 & 0 & 0 \\ E_2 & 3/4 & 1/4 & 0 & 0 \\ E_3 & 1/2 & 0 & 0 & 1/2 \\ E_4 & 0 & 0 & 1/2 & 1/2 \end{array}$$

The graph can be sketched as follows:



From the graph, we can notice that:

- starting from E_1, we can reach E_2. So E_1 **leads to** E_2.And starting from E_2, we can reach E_1, so E_1 and E_2 **communicate**.

- E_3 leads to E_1 but as E_1 does not lead to E_3, E_1 and E_3 do not communicate.

- Starting from E_1, we are sure to return to it as

  $P\{not\ to\ return\ to\ E_1\ /\ we\ start\ \ from\ E_1\} = \frac{1}{2} \times \lim_{n \to +\infty} \left(\frac{1}{4}\right)^n = 0$

  E_1 is said to be **recurrent**

- starting from E_3, we are not sure to return to it since if we reach E_1, it will be impossible to return to E_3. E_3 is said to be **transient**.

As E_1 and E_2 (respectively E_3 and E_4) communicate, they are of the same type, i.e. E_2 (respectively E_4) is recurrent (respectively transient).

- The states which communicate with each other form a subset called a **class**.

- A class is said to recurrent (respectively transient) if its elements are recurrent (respectively transient).

- A recurrent class R is an **absorbing class** in the sense that no element outside R can be reached by any of the elements of R.

(E_1, E_2) forms a **recurrent class** while (E_3, E_4) is a **transient class**.

ASYMPTOTIC BEHAVIOUR OF A MARKOV CHAIN

The asymptotic behaviour of a Markov chain is described by the following theorem:

Let $(X_n)_{n\geq 1}$ be a Markov chain having a finite number of states n_0 with T as transition probability matrix. We assume that $(X_n)_{n\geq 1}$ has one single recurrent class. Then:

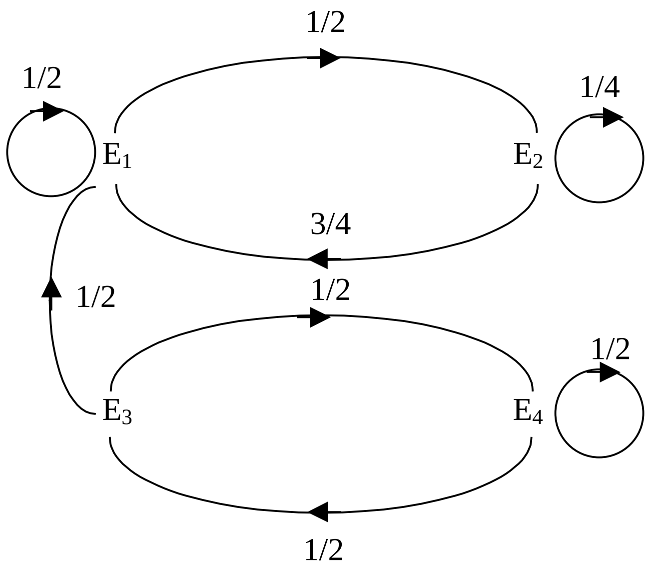
a stationary (or invariant) probability distribution μ exists and satisfies the equation $\mu = \mu \times T$. μ is a row vector $(\mu_1, \mu_2, ..., \mu_{n_0})$ such as:

$$\forall\, k \quad 1 \leq k \leq n_0 \qquad \mu_k = \lim_{n \to +\infty} P\{X_n = k\}$$

In other words, it means that after a sufficiently long time, the probability distribution will be approximately invariant.

The stationary probabilities of the transient states are null since, after a sufficiently long time, we are sure not to return to a transient state. Consequently, to solve the equation $\mu = \mu \times T$ for μ, we can limit ourselves to the probability transition sub-matrix T' which concerns only the recurrent states.

Let us resume our example:



There is only one recurrent class (E_1, E_2), so we know from the theorem that a stationary probability distribution exists.

As E_3 and E_4 are transient states, we limit ourselves to the sub-matrix T':

$$T' = \begin{array}{c} \\ E_1 \\ E_2 \end{array} \begin{array}{c} \begin{array}{cc} E_1 & E_2 \end{array} \\ \begin{pmatrix} 1/2 & 1/2 \\ 3/4 & 1/4 \end{pmatrix} \end{array}$$

Let us denote $\mu = (x, y)$ the stationary probability distribution. We have to solve the system:

$$\begin{cases} (x, y) = (x, y)\begin{pmatrix} 1/2 & 1/2 \\ 3/4 & 1/4 \end{pmatrix} \\ x + y = 1 \quad \textit{(as } \mu \textit{ is a probability distribution)} \end{cases}$$

$$\begin{cases} x = \dfrac{x}{2} + \dfrac{3y}{4} \\ y = \dfrac{x}{2} + \dfrac{y}{4} \\ x + y = 1 \end{cases}$$

Eventually, we get:

$$\begin{cases} x = \dfrac{3}{5} \\ y = \dfrac{2}{5} \end{cases}$$