# Markov Property

A sequence of random variables $\{X_n\}$ is called a Markov chain if it has the **Markov property**:

$$P\{X_k = i | X_{k-1} = j, X_{k-2}, \cdots, X_1\} = P\{X_k = i | X_{k-1} = j\}$$

States are usually labelled $\{(0,)1, 2, \cdots\}$, and state space can be finite or infinite.

- First order Markov assumption (memoryless):

$$P(q_t = i \,|\, q_{t-1} = j, q_{t-2} = k, \cdots) = P(q_t = i \,|\, q_{t-1} = j)$$

- Homogeneity:

$$P(q_t = i \,|\, q_{t-1} = j) = P(q_{t+l} = i \,|\, q_{t+l-1} = j)$$

# 1-D Random Walk

- Time is slotted.
- The walker flips a coin every time slot to decide which way to go:

$$X(t+1) = \begin{cases} X(t)+1 & \text{with probability p} \\ X(t)-1 & \text{with probability (1-p)} \end{cases}$$

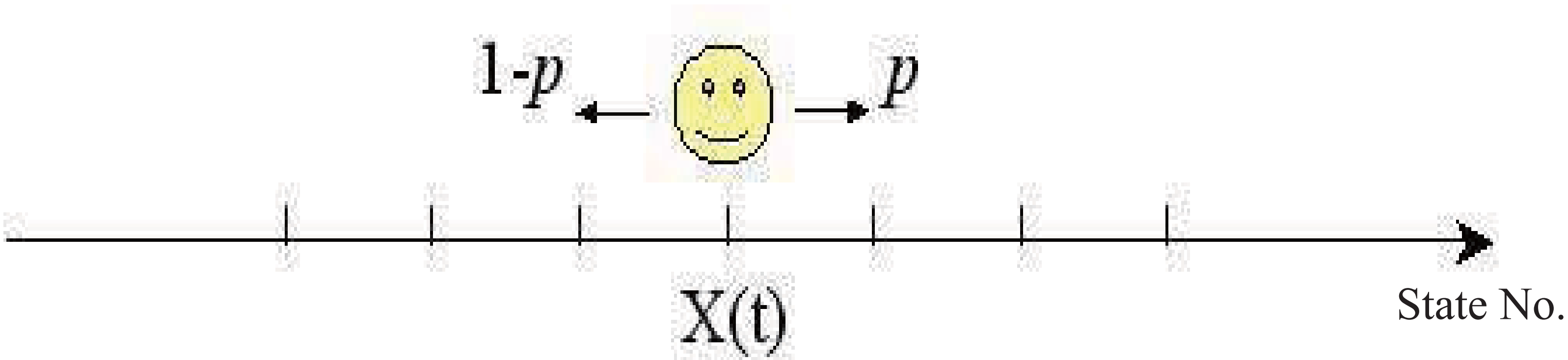


Figure 1: 1-D Random walk.

# Homogeneous Markov chain

**Definition**: an homogeneous Markov Chain has $P_{ij}^{n,n+1} = P_{ij}$ for all $n$. The one-step transition probabilities do not depend on time $n$.

**Definition**: the transition probability matrix is given by

$$\mathbf{P} = [P_{ij}] = \begin{bmatrix} P_{00} & P_{01} & P_{02} & \cdots \\ P_{10} & P_{11} & P_{12} & \cdots \\ P_{20} & P_{21} & P_{22} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

where

$$P_{ij} \geq 0 \quad \text{and} \quad \sum_j P_{ij} = 1$$

The $n$-step transition probability $P_{ij}^n$ is the probability that a process in state $i$ will be in state $j$ after $n$ steps, that is

$$P_{ij}^n = \mathtt{P}\,(X_{n+m} = j \mid X_m = i)$$

## Illustrative example

The No-Claim Discount system in car insurance is a system where the premium charged depends on the policyholder's claim record. This is usually modelled using a Markov chair. In an NCD scheme a policyholder's premium will depend on whether or not they had a claim.

| NCD level | Discount % |
|---|---|
| 0 | 0 |
| 1 | 25 |
| 2 | 50 |

If a policyholder has no claims in a year then they move to the next higher level of discount (unless they are already on the highest level, in which case they stay there). If a policyholder has one or more claims in a year then they move to the next lower level of discount (unless they are already on the lowest level, in which case they stay there). The probability of no claims in a year is assumed to be 0.9.

The state space of the Markov chain is 0,1,2 where state $i$ corresponds to NCD level $i$.

The transition matrix is

$$\mathbf{P} = \begin{bmatrix} 0.1 & 0.9 & 0 \\ 0.1 & 0 & 0.9 \\ 0 & 0.1 & 0.9 \end{bmatrix}$$

Example 1: (Forecasting the weather): suppose that the chance of rain tomorrow depends on previous weather conditions only through whether or not it is raining today and not on past weather conditions. Suppose also that if it rains today, then it will rain tomorrow with probability $\alpha$; and if it does not rain today, then it will rain tomorrow with probability $\beta$.

If we say that the process is in state 0 when it rains and state 1 when it does not rain, then the above is a two-state markov chain whose transition probabilities are given by

$$\mathbf{P} = \begin{bmatrix} \alpha & 1-\alpha \\ \beta & 1-\beta \end{bmatrix}$$

# Classification of States

Some definitions:

- A state $i$ is said to be an *absorbing state* if $P_{ii} = 1$ or, equivalently, $P_{ij} = 0$ for any $j \neq i$.
- State $j$ is *accessible* from state $i$ if $P_{ij}^n > 0$ for some $n \geq 0$. This is written as $i \rightarrow j$, $i$ leads to $j$ or $j$ is accessible from $i$. Note that if $i \rightarrow j$ and $j \rightarrow k$ then $i \rightarrow k$.
- State $i$ and $j$ communicate if $i \rightarrow j$ and $j \rightarrow i$. This is written as $i \leftrightarrow j$. Note that $i \leftrightarrow i$ for all $i$, if $i \leftrightarrow j$ then $j \leftrightarrow i$, and if $i \leftrightarrow j$ and $j \leftrightarrow k$ then $i \leftrightarrow k$.
- The *class* of states that communicate with state $i$ is $C(i) = \{j \in S, i \leftrightarrow j\}$. Two states that communicate are in the same class. Note that for any two states $i, j \in S$ either $C(i) = C(j)$ or $C(i) \cap C(j) = \Phi$. Also if $j \in C(i)$, then $C(i) = C(j)$.
- A Markov chain is said to be *irreducible* if there is only one class. All states therefore communicate with each other.

**Example 3** : Consider the Markov chain consisting of the three states 0,1,2 and having transition probability matrix

$$\mathbf{P} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ 0 & \frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

It is easy to verify that this Markov chain is irreducible. For example, it is possible to go from state 0 to state 2, since

$$0 \rightarrow 1 \rightarrow 2$$

That is, one way of getting from state 0 to state 2 is to go from state 0 to state 1 (with probability 1/2) and then go from state 1 to state 2 (with probability 1/4).

It is also possible to go from state 2 to state 0, since 2→1→0

**Example 4**: Consider a Markov chain consisting of the four states 0,1,2,3 and having transition probability matrix

$$\mathbf{P} = \begin{bmatrix} 0.5 & 0.5 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 \\ 0.25 & 0.25 & 0.25 & 0.25 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

The classes of this Markov chain are $\{0,1\}$, $\{\ 2\ \}$, and $\{3\}$. Note that while state 0 (or 1) is accessible from state 2, the reverse is not true. Since state 3 is an absorbing state, that is, $P_{33} = 1$, no other state is accessible from it.

# Classification of States

Definition: Recurrent and Transient States

In a Markov chain, a state i is transient if there exists a state j such that i→j but j↛i; otherwise, if no such state j exists, then state i is recurrent.

Theorems:

- A transient state will only be visited a finite number of times.
- In a finite state Markov chain, at least one state must be recurrent or that not all states can be transient. For otherwise, after a finite number of times, no state will be visited which is impossible.
- If state $i$ is recurrent and state $i$ communicates with state $j(i \leftrightarrow j)$, then state $j$ is recurrent.
- If state $i$ is transient and state $i$ communicates with state $j(i \leftrightarrow j)$, then state j is transient.

- All states in a finite irreducible Markov chain are recurrent.
- A class of states is recurrent if all states in the class are Recurrent. A class of states is transient if all states in the class are transient

**Example 5**: consider a Markov chain with state space 0, 1, 2, 3, 4 and transition matrix

$$\mathbf{P} = \begin{bmatrix} 0.5 & 0.5 & 0 & 0 & 0 \\ 0.5 & 0.5 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0.5 & 0 \\ 0 & 0 & 0.5 & 0.5 & 0 \\ 0.25 & 0.25 & 0 & 0 & 0.5 \end{bmatrix}$$

Determine which states are recurrent and which states are transient.

**Example 6**: consider a Markov chain with state space 0, 1, 2, 3, 4, 5 and transition matrix

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} & 0 & 0 & 0 \\ 0 & \frac{1}{5} & \frac{2}{5} & \frac{1}{5} & 0 & \frac{1}{5} \\ 0 & 0 & 0 & \frac{1}{6} & \frac{1}{3} & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & \frac{1}{4} & 0 & \frac{3}{4} \end{bmatrix}$$

Determine which states are recurrent and which states are transient.

**Example 7** consider earlier example in which the weather is considered as a two-state Markov chain. If $\alpha = 0.7$ and $\beta = 0.4$, then calculate the probability that it will rain four days from today given that it is raining today.

Solution: The one-step transition probability matrix is given by

$$\mathbf{P} = \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix}$$

Hence

$$\mathbf{P}^{(2)} = \mathbf{P}^2 = \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix} \cdot \begin{bmatrix} 0.7 & 0.3 \\ 0.4 & 0.6 \end{bmatrix} = \begin{bmatrix} 0.61 & 0.39 \\ 0.52 & 0.48 \end{bmatrix}$$

$$\mathbf{P}^{(4)} = (\mathbf{P}^2)^2 = \begin{bmatrix} 0.61 & 0.39 \\ 0.52 & 0.48 \end{bmatrix} \cdot \begin{bmatrix} 0.61 & 0.39 \\ 0.52 & 0.48 \end{bmatrix} = \begin{bmatrix} 0.5749 & 0.4251 \\ 0.5668 & 0.4332 \end{bmatrix}$$

and the desired probability $P_{00}^{(4)}$ equals 0.5749.

## Chapman-Kolmogorov Equation

Define $n$-step transition probabilities

$$P_{ij}^{(n)} = P\{X_{n+k} = j | X_k = i\}, \quad n > 0, i, j \geq 0$$

CK equation states that

$$P_{ij}^{(n+m)} = \sum_{k=0}^{\mathrm{K}} P_{ik}^{n} P_{kj}^{m}, \quad \text{for all } n, m > 0, \text{ all } i, j$$

Matrix notation:

$$\mathbf{P}^{(n+m)} = \mathbf{P}^{(n)} \mathbf{P}^{(m)}$$

In previous example, we calculated $\mathbf{P}^{(4)}$ for a two-state Markov chain; it turned out to be

$$\mathbf{P}^{(4)} = \begin{bmatrix} 0.5749 & 0.4251 \\ 0.5668 & 0.4332 \end{bmatrix}$$

From this it follows that $\mathbf{P}^{(8)} = \mathbf{P}^{(4)}\,\mathbf{P}^{(4)}$ is given by

$$\mathbf{P}^{(8)} = \begin{bmatrix} 0.572 & 0.428 \\ 0.570 & 0.430 \end{bmatrix}$$

Note that the matrix $\mathbf{P}^{(8)}$ is almost identical to the matrix $\mathbf{P}^{(4)}$, and secondly, that each of the rows of $\mathbf{P}^{(8)}$ has almost identical entries. In fact, it seems that $P_{ij}^{n}$ is converging to some value (as $n \to \infty$) which is the same for all $i$. In other words, there seems to exist a limiting probability that the process will be in state $j$ after a large number of transitions, and this value is independent of the initial state.

## Steady-state matrix

- The steady-state probabilities are average probabilities that the system will be in a certain state after a large number of transition periods.
- The convergence of the steady-state matrix is independent of the initial distribution.
- Long-term probabilities of being on certain states

**Theorem**: For an irreducible Markov chain, $\lim_{n\to\infty} P_{ij}^n$ exists and is independent of $i$. Furthermore, if we let

$$\pi_j = \lim_{n\to\infty} P_{ij}^n$$

then $\pi_j$ is the unique non-negative solution to the system of equation

$$\pi_j = \sum_{i=0}^{K} \pi_i P_{ij} \quad (j \geq 0) \quad (\text{matrix form } \pi\mathbf{P} = \pi)$$

and

$$\sum_{j=0}^{K} \pi_j = 1$$

which are expanded as

$$\pi_0 P_{00} + \pi_1 P_{10} + \pi_2 P_{20} + \ldots + \pi_K P_{K0} = \pi_0$$
$$\pi_0 P_{01} + \pi_1 P_{11} + \pi_2 P_{21} + \ldots + \pi_K P_{K1} = \pi_1$$
$$\vdots$$
$$\pi_0 P_{0K} + \pi_1 P_{1K} + \pi_2 P_{2K} + \ldots + \pi_K P_{KK} = \pi_K$$
$$\pi_0 + \pi_1 + \pi_2 + \ldots + \pi_K = 1$$

- The limiting probability $\pi_j$ that the process will be in state $j$ at time $n$ also equals the long-run proportion of time that the process will be in state $j$.
- Consider the probability that the process will be in state $j$ at time $n+1$

$$P(X_{n+1} = j) = \sum_{i=0}^{\infty} P(X_{n+1} = j | X_n = i)\, P(X_n = i) = \sum_{i=0}^{\infty} P_{ij}\, P(X_n = i)$$

  Now, if we let $n \to \infty$, the result of the theorem follows.
- If the Markov chain is irreducible, then $\pi_j$ is interpreted as the long run proportion of time the Markov chain is in state $j$.

## Illustrative example

Example 8: Consider Example 1, in which we assume that if it rains today, then it will rain tomorrow with probability $\alpha$; and if it does not rain today, then it will rain tomorrow with probability $\beta$. If we say that the state is 0 when it rains and 1 when it does not rain, then the limiting probability $\pi_0$ and $\pi_1$ are given by

$$\begin{aligned} \pi_0 &= \alpha\pi_0 + \beta\pi_1 \\ \pi_1 &= (1-\alpha)\pi_0 + (1-\beta)\pi_1 \\ \pi_0 &+ \pi_1 = 1 \end{aligned}$$

which yields that

$$\pi_0 = \frac{\beta}{1+\beta-\alpha} \quad \pi_1 = \frac{1-\alpha}{1+\beta-\alpha}$$

For example, if $\alpha = 0.7$ and $\beta = 0.4$, then the limiting probability of rain is $\pi_0 = \frac{4}{7} = 0.571$.