Markov Chain

Defn and Examples.

Consider a simple coin tossing experiment repeated for a number of times. The possible outcomes at each trial are two: head with probability say p and tail with probability q, $p+q=1$. Let us denote head by 1 and tail by 0 and the random variable denoting the result of the nth toss by X_n. Then for $n=1,2,3,\ldots$

$$\Pr\{X_n = 1\} = p \qquad \Pr\{X_n = 0\} = q$$

Thus we have a seq. of random variables $X_1, X_2, \ldots$ The trials are independent and the result of the nth trials does not depends in any way on the previous trials numbered $1,2,\ldots,(n-1)$. The r.v. are independent

Consider now the r.v. given by the partial sum $S_n = X_1 + X_2 + \cdots + X_n$. The sum S_n gives the accumula no. of heads in the first trails and its possible values are $0, 1, \ldots, n$

We have, $S_{n+1} = S_n + X_{n+1}$

given that $S_n = j \; (j=0,1,\ldots,n)$ the r.v S_{n+1} can assume only two possible values: $S_{n+1} = j$ with prob. q and $S_{n+1} = j+1$ with prob. p; these probabilities are not at all affected by the values of the variables $S_1, \ldots,$

Thus

$$\Pr\{S_{n+1} = j+1 \mid S_n = j\} = p$$

$$\Pr\{S_{n+1} = j \mid S_n = j\} = q$$

We have here an example of a Markov chain, a case of simple dependence that the outcome of $(n+1)$th trial depends directly on the nth trial and only on it.

The conditional prob. of S_{n+1} given S_n depends on the value of S_n and the manner in which the value of S_n was reached is of no consequence.

Def^n The stochastic process* $\{X_n : n = 0, 1, 2, \ldots\}$ is called a Markov chain if for $j, k, j_1, \ldots, j_{n-1} \in N$ (or any subset of I)

$$P_r\{X_n = k \mid X_{n-1} = j, X_{n-2} = j_1, \ldots, X_0 = j_{n-1}\}$$

$$= P_r\{X_n = k \mid X_{n-1} = j\} = P_{jk}$$

whenever the first member is defined.

The outcomes are called the states of the Markov chain; if X_n has the outcome j (ie, $X_n = j$) the process is said to be at state j at nth trial. To a pair of states (j, k) at the two successive trials (nth and $(n+1)$th) there is an associated conditional probability P_{jk}. It is the probability of transition from the state j at nth trials to the state k at $(n+1)$th trials. The transition probabilities P_{jk} are basic to study of the structure of the Markov chain.

* A stochastic process: Let (Ω, P, S) be a given probability space. A collection of r.v. $\{X(t), t \geq 0\}$ defined on the probability space is called stochastic process.

State Transition Matrix and Diagram:

We often list the transition probabilities in a matrix. The matrix is called the state transition matrix or transition probability matrix and is usually shown by P. Assuming the states are $1, 2, \ldots, r$, then the state transition matrix is given by

$$P = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1r} \\ p_{21} & p_{22} & \cdots & p_{2r} \\ \vdots & & & \\ p_{r1} & p_{r2} & \cdots & p_{rr} \end{bmatrix}$$

Note $p_{ij} \geq 0$ and for all i, we have

$$\sum_{k=1}^{r} p_{ik} = \sum_{k=1}^{r} P(X_{m+1} = k \mid X_m = i) = 1$$

This is because, given that we are in state i, the next state must be one of the possible states. Thus the rows of any state transition matrix must sum to one.

State Transition Diagram:

A Markov chain is usually shown by a state transition diagram. Consider a Markov chain with

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

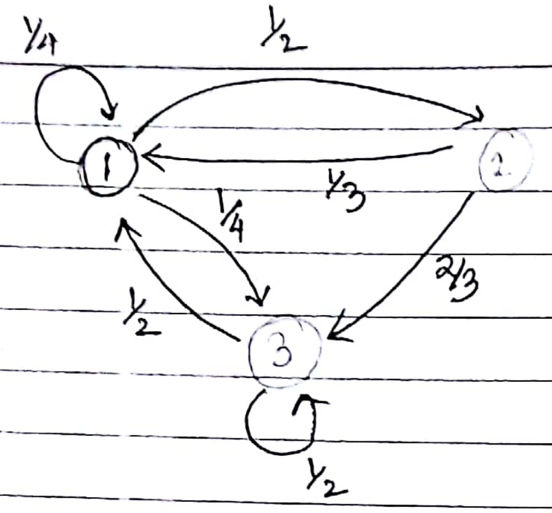


Fig shows the state transition diagram for the above Markov chain. There are Three possible states 1, 2, and 3. and the arrows from each state to other states shows the transition probabilities p_{ij}. When there is no arrow from state i to state j, it means $p_{ij} = 0$.

Examples:

Consider the Markov chain shown in the fig-

a. Find $P(X_4 = 3 \mid X_3 = 2)$.
b. Find $P(X_3 = 1 \mid X_2 = 1)$
c. If we know $P(X_0 = 1) = \frac{1}{3}$. Find $P(X_0 = 1, X_1 = 2)$
d. If we know $P(X_0 = 1) = 1/3$, Find $P(X_0 = 1, X_1 = 2, X_2 = 3)$

Solution:

a. $P(X_4 = 3 \mid X_3 = 2) = p_{23} = \frac{2}{3}$.

b. $P(X_3 = 1 \mid X_2 = 1) = p_{11} = 1/4$

c. $P(X_0 = 1, X_1 = 2) = P(X_0 = 1) P(X_1 = 2 \mid X_0 = 1)$
$= \frac{1}{3} p_{12} = \frac{1}{3} \times \frac{1}{2} = \frac{1}{6}$

We can ~~rewir~~ rewrite the above result in the form of matrix multiplication

$$\pi^{(1)} = \pi^{(0)} P$$

where P is the state transition matrix.

Similarly

$$\pi^{(2)} = \pi^{(1)} P = \pi^{(0)} P^2$$

More generally,

$$\pi^{(n+1)} = \pi^{(n)} P \quad , \text{ for } n = 0, 1, 2, \ldots$$

$$\pi^{(n)} = \pi^{(0)} P^n \quad , \text{ for } n = 0, 1, 2, \ldots$$

Example:

Consider the system that can be in one of two possible states, $S = \{0, 1\}$. In particular, suppose the transition matrix is given by

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

Suppose that the system is in state 0 at time $n=0$ ie, $X_0 = 0$

a. Draw the state transition diagram.

b. Find the probability that the system is in state 1 at time $n=3$.

d. $P(X_0 = 1, X_1 = 2, X_2 = 3) = P(X_0 = 1)\, P(X_1 = 2 | X_0 = 1)\, P(X_2 = 3 | X_1 = 2, X_0 = 1)$

$= \frac{1}{3}\, p_{12}\, P(X_2 = 3 | X_1 = 2)$ [by Markov Prop.]

$= \frac{1}{3} \cdot p_{12}\, p_{23} = \frac{1}{3} \frac{1}{2} \frac{2}{3} = \frac{1}{9}$.

Probability Distribution :

State Probability Distribution :

Consider a Markov chain $\{X_n, n = 0, 1, 2, \ldots\}$ where $X_n \in S = \{1, 2, \ldots, r\}$ Suppose we know the probability distribution of X_0. More specifically, define the row vector $\pi^{(0)}$ as

$$\pi^{(0)} = [\, P(X_0 = 1) \quad P(X_0 = 2) \quad \cdots \quad P(X_0 = r) \,]$$

How can we obtain the probability distribution of $X_1, X_2, \ldots$?

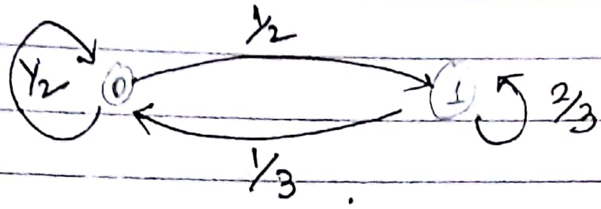
We can use the law of total probability. More specifically, for any $j \in S$, we can write:

$$P(X_1 = j) = \sum_{k=1}^{r} P(X_1 = j | X_0 = k)\, P(X_0 = k)$$

$$= \sum_{k=1}^{r} p_{kj}\, P(X_0 = k)$$

If we generally define $\pi^{(n)} = [P(X_n = 1)\; P(X_n = 2) \ldots P(X_n = r)]$

(a)



(b) Here $\pi^{(0)} = [P(X_0=0) \quad P(X_0=1)]$
$= [1 \quad 0]$

Thus $\pi^{(3)} = \pi^{(0)} P^3$

$$= [1 \quad 0] \cdot \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{2}{3} \end{bmatrix}^3 = \begin{bmatrix} \frac{29}{72} & \frac{43}{72} \end{bmatrix}$$

Thus the probability that the system is in state 1 at time $n=3$ is $\frac{43}{72}$.

n-step transition Probabilities :

Consider a Markov chain $\{X_n, n=0,1,\dots\}$ where $X_n \in S$. If $X_0 = i$, then $X_1 = j$ with probability p_{ij} ie, p_{ij} gives us the probability of going from state i to state j in one step. Now suppose that we are interested in finding the probability of going from i th state to j th state in two steps. ie,

$$p_{ij}^{(2)} = P(X_2 = j \mid X_0 = i).$$

We can find this probability by applying the law of total probability. In particular, we argue that X_1 can take one of the possible values in S. Thus we can write

$$p_{ij}^{(2)} = P(X_2 = j \mid X_0 = i) = \sum_{k \in S} P(X_2 = j \mid X_1 = k, X_0 = i)\, P(X_1 = k \mid X_0 = i)$$

$$= \sum_{k \in S} P(X_2 = j \mid X_1 = k)\; P(X_1 = k \mid X_0 = i)$$
(by Markov Prop)

$$= \sum_{k \in S} p_{kj}\, p_{ki}$$

$$\therefore \quad p_{ij}^{(2)} = P(X_2 = j \mid X_0 = i) = \sum_{k \in S} p_{kj}\, p_{ki}$$

In order to get to state j, we need to pass through some intermidiate state k. The probability of this event is $p_{ik} p_{kj}$. To obtain $p_{ij}^{(2)}$, we sum over all possible intermidiate states. Accordingly, we can define the two-step transition matrix as follows:

$$P^{(2)} = \begin{bmatrix} p_{11}^{(2)} & p_{12}^{(2)} & \cdots & p_{1r}^{(2)} \\ p_{21}^{(2)} & p_{22}^{(2)} & \cdots & p_{2r}^{(2)} \\ \vdots & & & \\ p_{r1}^{(2)} & p_{r2}^{(2)} & \cdots & p_{rr}^{(2)} \end{bmatrix}$$

Clearly,

$$P^2 = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1r} \\ p_{21} & p_{22} & \cdots & p_{2r} \\ \vdots & & & \\ p_{r1} & p_{r2} & \cdots & p_{rr} \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1r} \\ p_{21} & p_{22} & \cdots & p_{2r} \\ \vdots & & & \\ p_{r1} & p_{r2} & \cdots & p_{rr} \end{bmatrix} = P^{(2)}$$

More generally we can define the n-step transition probabilities $\{p_{ij}^{(n)}\}$ as

$$p_{ij}^{(n)} = P(X_n = j \mid X_0 = i) \quad \text{for } n = 0, 1, 2, \ldots \quad \text{--- (1)}$$

and n-step transition matrix.

$$p^{(n)} = \begin{bmatrix} p_{11}^{(n)} & p_{12}^{(n)} & \cdots & p_{1r}^{(n)} \\ p_{21}^{(n)} & p_{22}^{(n)} & \cdots & p_{2r}^{(n)} \\ \vdots & & & \\ p_{r1}^{(n)} & p_{r2}^{(n)} & \cdots & p_{rr}^{(n)} \end{bmatrix}$$

We can generalized Eqn (1). Let m, n be two positive integers and, assume $X_0 = i$. In order to get to state j in $(m+n)$ steps, the chain will be at some intermidiate state k after m steps. To obtain $p_{ij}^{(m+n)}$, we sum over all possible intermidiate states:

$$p_{ij}^{(m+n)} = P(X_{m+n} = j \mid X_0 = i)$$

$$= \sum_{k \in S} p_{ik}^{(m)} p_{kj}^{(n)}$$

The above equation is called the Chapman-Kolmogorov Equation. Similar to the case of two step transition probabilities, we can show.

$$p^{(n)} = p^{n} \quad \text{for } n = 1, 2, \ldots$$

The Chapman - Kolmogorov Equation can be written as

$$p_{ij}^{(m+n)} = P(X_{m+n} = j \mid X_0 = i)$$

$$= \sum_{k \in S} p_{ik}^{(m)} \, p_{kj}^{(n)}$$

The n-step transition matrix is given by

$$P^{(n)} = P^n \quad , \text{ for } n = 1, 2, 3, \ldots$$

Examples:

A ~~m~~ Markov chain $\{X_n\}$ on the state 0, 1, 2 has the probability matrix

$$P = \begin{bmatrix} 0.1 & 0.2 & 0.7 \\ 0.2 & 0.2 & 0.6 \\ 0.6 & 0.1 & 0.3 \end{bmatrix}$$

a. Compute the two-step transition matrix P^2.
b. What is $P(X_3 = 1 \mid X_1 = 0)$?
c. What is $P(X_3 = 1 \mid X_0 = 0)$?

a. $$P^2 = \begin{bmatrix} 0.1 & 0.2 & 0.7 \\ 0.2 & 0.2 & 0.6 \\ 0.6 & 0.1 & 0.3 \end{bmatrix} \begin{bmatrix} 0.1 & 0.2 & 0.7 \\ 0.2 & 0.2 & 0.6 \\ 0.6 & 0.1 & 0.3 \end{bmatrix}$$

$$= \begin{bmatrix} 0.47 & 0.13 & 0.4 \\ 0.42 & 0.14 & 0.44 \\ 0.26 & 0.17 & 0.57 \end{bmatrix}$$

b. $P(X_2 = 1 | X_1 = 0) = p_{01}^{2} = 0.13$

c. $P(X_3 = 1 | X_0 = 0) = p_{01}^{3}$

Now first calculating $P^3 = P^2 \cdot P$

$$= \begin{bmatrix} 0.47 & 0.13 & 0.4 \\ 0.42 & 0.14 & 0.44 \\ 0.26 & 0.17 & 0.57 \end{bmatrix} \begin{bmatrix} 0.1 & 0.2 & 0.7 \\ 0.2 & 0.2 & 0.6 \\ 0.6 & 0.1 & 0.3 \end{bmatrix}$$

$$= \begin{bmatrix} 0.313 & 0.16 & 0.527 \\ 0.334 & 0.156 & 0.51 \\ 0.402 & 0.143 & 0.455 \end{bmatrix} \quad \cancel{p_{00}^{3}}$$

$\therefore P(X_3 = 1 | X_0 = 0) = p_{01}^{3} = 0.16.$

Classification of States:

Accessible

The state j is accessible from the state i, written as $i \to j$ if $p_{ij}^{(n)} > 0$ for some n. We assume every state is accessible from itself since $p_{ii}^{(0)} = 1$.

Communicate

Two states i and j are said to communicate, written as $i \leftrightarrow j$, if they are accessible from each other.

ie, $i \leftrightarrow j$ means $i \to j$ and $j \to i$

Communication is an equivalence relation. That means

* Every state communicates with itself, $i \leftrightarrow i$
* If $i \leftrightarrow j$ then $j \leftrightarrow i$
* If $i \leftrightarrow j$ & $j \leftrightarrow k$ then $i \leftrightarrow k$.

Note:

state j is accessible from state i iff, starting in state i, it is possible that the process will ever enter state j.

If j is not accessible from i, then

$$P\{\text{ever be in } j \mid \text{start in } i\} = P\left\{\bigcup_{n=0}^{\infty}\{X_n = j\} \,\Big|\, X_0 = i\right\}$$

$$\leq \sum_{n=0}^{\infty} P\{X_n = j \mid X_0 = i\}$$

$$= \sum_{n=0}^{\infty} P_{ij}^{(n)} = 0$$