Lecture 2:Random Walks

The Lecture Contains:

- N-stage transition probability
- Few relevant information which is important
- Classification of the states of a Markov Chain
- Recurrence
- Wald's Equations
- Examples
- Transition probability matrices of a Markov chain
- Periodicity
- Generating functions

Previous Next

N-stage transition probability

Consider the following, i.e., at time $t = m$ the process is at state i, (say state number I), while at time $t = (m+n)$, it is at state j, (say state number II). Then the probability that from state number I to state number II is given by $P\left[X_{m+n} = j \middle| X_m = i\right] = p_{ij}^{(n)}$, where

$$p_{ij}^{(1)} = p_{ij} = P\left[X_{m+1} = j \middle| X_m = i\right]$$
$$p_{ij}^{(2)} = P\left[X_{m+2} = j \middle| X_m = i\right]$$
$$\vdots$$

Pictorially we have the following diagram (Figure 1.16) to illustrate the one dimensional stochastic process.

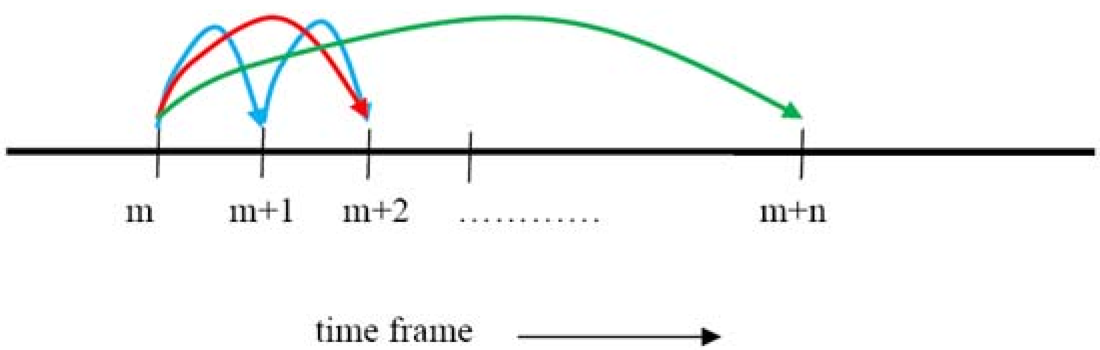

Figure 1.16: Illustration of concept of transition matrix

Here the blue lines are $p_{ij}^{(1)} = p_{ij} = P\left[X_{m+1} = j \middle| X_m = i\right]$ or $p_{jk}^{(1)} = p_{jk} = P\left[X_{m+2} = k \middle| X_{m+1} = j\right]$ as the case may be, while the red line denotes, $P\left[X_{m+2} = k \middle| X_m = i\right]$, which can be expressed as one stage transition probability values as the case may be. Finally the green line denotes A, and in order to avoid confusion we consider the initial and final states as *i* and *j* ONLY.

Previous Next

Few relevant information which is important

Consider Bi are mutually exclusive and exhaustive events, such that $\sum_{\forall i} B_i = \Omega$, which is the sure event, i.e., technically the whole of the sample space and also assume $B_i \cap B_j = \phi$, $\forall i, j \in I$ is true. Now suppose A and C are events, such that we write A as $A = A \cap \sum_{\forall i} B_i = \sum_{\forall i}(A \cap B_i)$. Then what may be of interest to us is to find $P[A|C] = P\left[\left\{\sum_{\forall i}(A \cap B_i)\right\}\middle|C\right] = \sum_{\forall i} P[\{A \cap B_i\}|C]$

Now due to the fact that $A = A \cap \sum_{\forall i} B_i = \sum_{\forall i}(A \cap B_i)$ is true, we can write the following, i.e., $\sum_{\forall i} P[A \cap B_i] = \sum_{\forall i} P[A|B_i] \times P[B_i]$, so with conditional probability $P[A|C]$ now becomes

$$P[A|C] = \sum_{\forall i} P[A|B_i, C] \times P[B_i|C]$$

Pictorially we can denote this as following as shown in Figure 1.17.

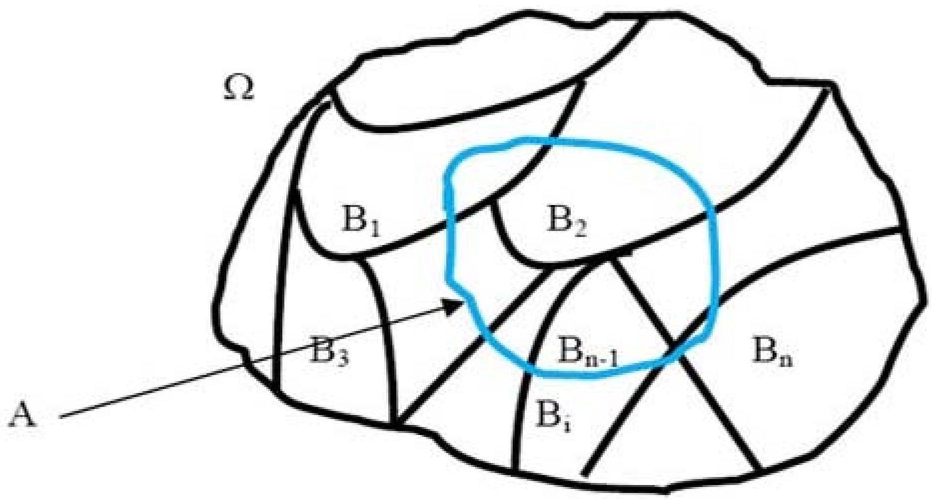

Figure 1.17: Illustration of the concept of conditional probability and joint distribution

Hence we can write

$$P[X_{m+n} = j|X_m = i] = p_{ij}^{(n)} = \sum_k P[X_{m+n} = j, X_{m+1} = k|X_m = i] \text{as}$$

$$= \sum_k P[X_{m+n} = j|X_m = i, X_{m+1} = k] \times P[X_{m+1} = k|X_m = i]$$

$$= \sum_k P[X_{m+n} = j|X_m = i, X_{m+1} = k] \times p_{ik} = \sum_k p_{kj}^{(n-1)} \times p_{ik}$$

$$= \sum_k p_{kj}^{(n-1)} \times p_{ik}$$

Hence: $p_{ij}^{(n)} = \sum_{k \in I} p_{kj}^{(n-1)} \times p_{ik}$.

Module 1:Concepts of Random walks, Markov Chains, Markov Processes

Lecture 2:Random Walks

So generalizing we can write

$p_{ij}^{(1)} = p_{ij}$, $p_{ij}^{(2)} = \sum_{k \in I} p_{kj} \times p_{ik}$, $p_{ij}^{(3)} = \sum_{k \in I} p_{kj}^{(2)} \times p_{ik}$ and so on.

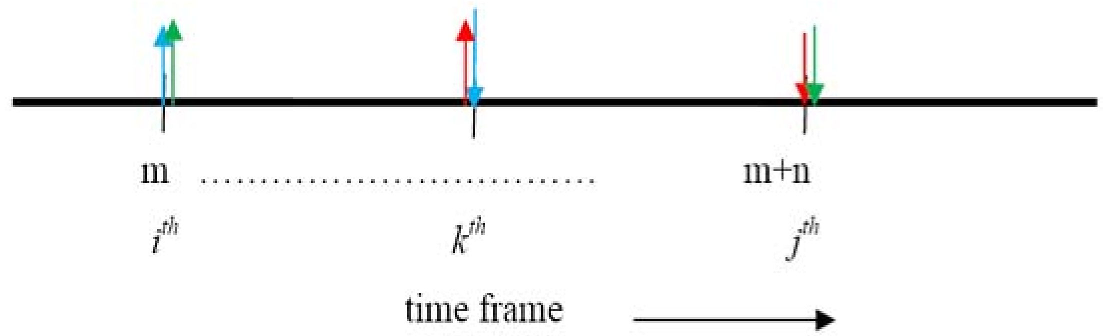

Figure 1.18: Movement from i^{th} to j^{th} state through an arbitrary k^{th} position at any point of time between t such that $m \le t \le (m+n)$

Figure 1.18 thus shows how the stochastic process starts at i^{th} state at time m and finally goes to j^{th} state at time $m+n$. What is interesting to note it how does this transition takes place. A general understanding of the stochastic process movement would make it clear that the process could have arbitrarily been at k^{th} state at say any point of time t between m and $m+n$, i.e., $m < t < m+n$, and the colour scheme of red, blue and green should make this arbitrary movement clear.

Using matrix notation we already have $P = (p_{ij}) = \begin{pmatrix} p_{11} & p_{12} & \cdots \\ p_{21} & p_{22} & \cdots \\ \vdots & \vdots & \cdots \end{pmatrix}$, such that $P^2 = \left(p_{i,j}^{(2)}\right)$, where it means the element in the P^2 matrix and not that we simply multiply pij X pij. Thus it means that $P^2 = P \times P$. Similarly we have $p_{ij}^{(3)} = \sum_{k \in I} p_{ik} \times p_{kj}^{(2)}$, hence $P^2 = P \times P \times P$.Thus generalizing we have $p^n = \underbrace{p \times \cdots \times p}_{n\ times}$ and also $p^{m+n} = p^m \times p^n = \left[\underbrace{p \times \cdots \times p}_{n\ times}\right] \times \left[\underbrace{p \times \cdots \times p}_{m\ times}\right]$.Here it must be remember that we are considering there is no structural breaks or change in the underlying distribution i.e., $F_X(x) \ne F_Y(y)$. If that occurs then some where we would have the transition probability matrix general structures as different, i.e., for t = m, m+1,…, m* we have $F_X(x)$ as the underlying distribution, and afterwards from t = m*+1, m*+2,…, m*+n the distribution is $F_Y(y)$

Hence the basic thing required is to know the one step transition matrix is pij.

If $P[X_0 = i] = a_i$ and $P = (a_{ij})$ are given then $P[X_n = i] = \sum_{j \in I} P[X_n = i, X_0 = j]$, i.e.,

$$P[X_n = i] = \sum_{j \in I} P[X_0 = j] \times P[X_n = i | X_0 = j],\ \text{i.e.,}$$
$$= \sum_{j \in I} a_j \times p_{ji}^{(n)},$$

Thus we have the probability in terms of intermittent probability and the transition matrix

Classification of the states of a Markov Chain

In this classification we will mention state ***j*** as being ***accessible*** from another state ***i***, i.e., $i \rightarrow j$ if $p_{01}^{(n)} = p_{ij}^{(n)} > 0$ for some $n \geq 0$, here n denotes the n stage transition. In case if it not, then the state ***j*** is ***not accessible*** from state ***i***.

Example 1.12

Let us consider a conditional probability matrix as given below

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ q & 0 & p & 0 \\ 0 & q & 0 & p \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Few observations from Example 1.12

- Then if the question is asked, is state 3 accessible from state 1, the answer is No.
- No state can be reached from state 1, i.e., state 1 is like a sink as $p_{11} = 1$ and it is called the absorbing state. Similarly once you reach state 4 you remain in that state for ever as $p_{44} = 1$ and it is also called the absorbing state.
- Similarly $p_{41} = p_{42} = p_{43} = 0$.
- It is possible to go from state 3 to state 1, i.e., 3 to 2 to 1, but reverse is not possible.
- In case we have $i \rightarrow j$ as well as $j \rightarrow i$, then the two states are communicating between themselves. So in the example given above we have
- $p_{23} > 0$ as 2 leads to 3 and also $p_{32} > 0$ as 3 leads to 2, hence state 2 and state 3 are communicating states.
- A set of states in a Markov chain is said to be closed if pij = 0 for *i* Î C and *j* Ï C, what ever the set C be.
- If a subset of a Markov chain is closed then the subset also forms a Markov chain
- If a Markov chain has no closed subset, i.e., if the Markov chain is itself not closed, then the Markov chain is said to be irreducible.
- Further more if $i \rightarrow j$ and $j \rightarrow k$, it implies that $i \rightarrow k$, which means we are only required to show that $p_{ik}^{(n)} > 0$ for some $n \geq 1$.

Note

Now suppose we know that $p_{ij}^{(n_1)} > 0$ and $p_{jk}^{(n_2)} > 0$ and since $p_{ij}^{(n_1)} \times p_{jk}^{(n_2)} > 0$, then from ***Chapman Kolmorogrov*** equations we can show that $i \rightarrow k$, is true, i.e., $p_{ij}^{(n_1)} > 0$ as $i \rightarrow j$ and $j \rightarrow k$ are both true and this statement that $i \rightarrow k$ is true will hold for some *n* ³ *1*. For the interested readers we would advice that rather than panic they should wait till we cover the simple proof of ***Chapman Kolmorogrov*** equations.

Consider a fixed, but arbitrary state i and suppose at time $t = 0$ it is at state i, i.e., X_0 = *i*, and also define $f_i^{(n)} = P\{\text{first time return to } i \text{ state occurs in } n \text{ steps}\}$, i.e., at the n^{th} step I am back at state ***i***.

Previous Next

Recurrence

Consider an arbitrary state, i, which is fixed, such that for all integers, $n \geq 1$, we define: $f_{ii}^{n} = P\left[X_n = i, X_\nu \neq i, \nu = 1,2,\ldots,n-1 \middle| X_0 = i\right]$, which is basically the probability that after having started from the i^{th} state it comes back to $P\left[X_{m+n} = j \middle| X_m = i\right]$ state for the ***first time*** ONLY at the n^{th} transition. The definition makes it clear that $f_{ii}^{1} = p_{ii}$ and $p_{ii}^{n} = \sum_{k=0}^{n}\left(f_{ii}^{k} \times p_{ii}^{n-k}\right)$ for $n \geq 1$, and we also define $f_{ii}^{0} = 0$ Now let us consider the simple illustrations (Figure 1.19 and Figure 1.20).

Figure 1.19: Movement from i^{th} state back to i^{th} state considering that there may or may not be any reoccurrence or visits to i^{th} state during the time period n

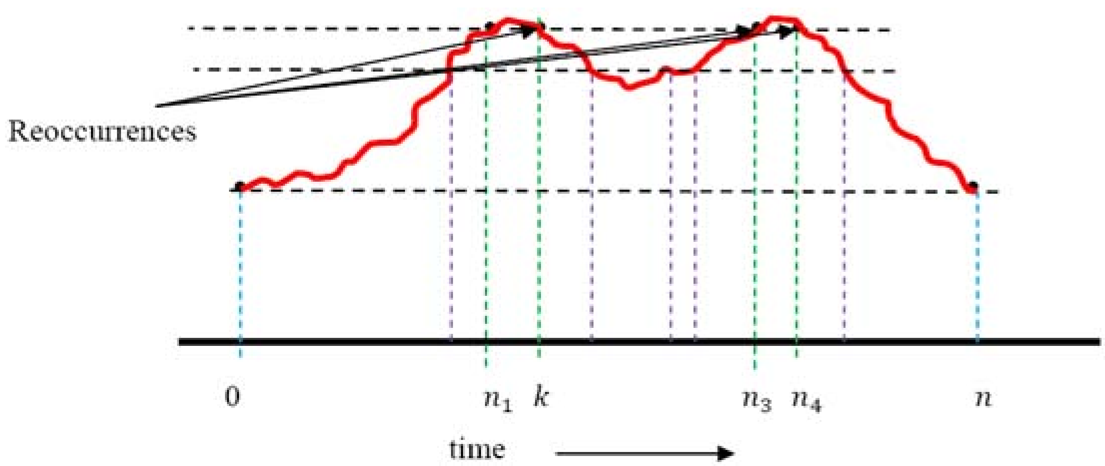

Figure 1.20: Movement from i^{th} state back to i^{th} state considering that there may or may not be any reoccurrences or visits at i^{th} state in between

From both the figures (Figure 1.19 and Figure 1.20) it is clear that $X_0 = i$ and $X_n = i$ and the first return to state i occurs at k^{th} transition. If this return is denoted by event E_k, then the events E_k $(k = 1,2,\ldots,n)$ are *mutually exclusive*.

Now the probability of the event that the first return is at the k^{th} transition is f_{ii}^{k} which we already know. Now for the remaining $(n-k)$ transitions we will only deal with those such that $X_n = i$ holds true. In case it does not, we will not consider that.

So we have

$$P[E_k] = P\left[\text{first return is at } k^{th} \text{ transition} \middle| X_0 = i\right] \times P\left[X_n = i \middle| X_0 = i\right]$$

$\because P[E_k] = f_{ii}^k \times P_{ii}^{n-k}$ for $1 \le k \le n$ and we also have $P_{ii}^0 = 1$

Hence

$$P\left[X_n = i \middle| X_0 = i\right] = \sum_{k=1}^{n} P(E_k) = \sum_{k=1}^{n} f_{ii}^k P_{ii}^{n-k} = \sum_{k=0}^{n} f_{ii}^k P_{ii}^{n-k} \text{ as } f_{ii}^0 = 0$$

$$p_{ii}^n = Pr\{X_n = i | X_0 = i\}$$

$$= \sum_{m=1}^{n} P\{\text{first time return to } i \text{ state occurs in } m^{th} \text{ step}\} \times Pr\{X_n = i | X_m = i\}$$

Previous Next

Let us pictorially illustrate the concept in Figure 1.21, where we start at i^{th} state at time $n=0$ and after $n=n$ time it come back to i^{th} state. If one observes closely the main difference with Figure 1.18 is the fact that here we denote the states while in Figure 1.18 it was the time which was depicted along the line.

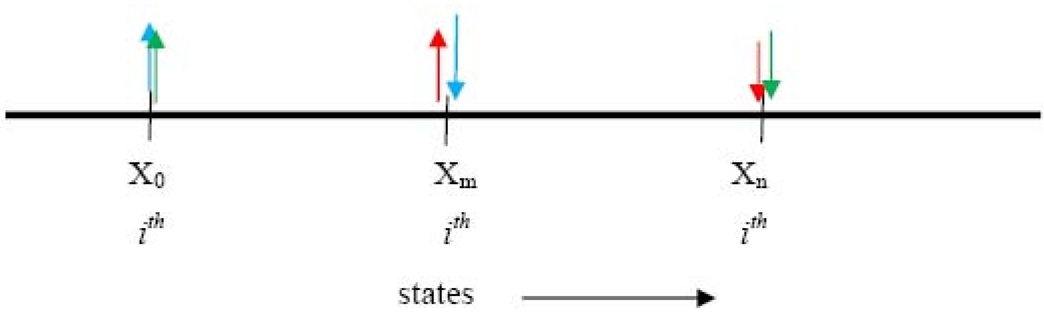

Figure 1.21: Illustration of return of the stochastic process to i^{th} state exactly at end of n time

The colour scheme in Figure 1.21 is quite easy to understand if we concentrate on the fact that it can be green line starting at i^{th} position at $n=0$ and returning to i^{th} position at $n=n$ for the first time. While the blue would denote that reaches i^{th} position once at $n=m$ time and then again the stochastic process continues till it reaches i^{th} position at $n=n$ for the second time. Continuing with the same logic we can have such visits to i^{th} position many number of times but remembering that i^{th} position is reached at $n=n$, i.e., $X_n=i$.

Then we have

$p_{ii}^{(n)}=\sum_{m=1}^{n} f_i^{(m)} \times p_{ii}^{(n-m)}$, from which it is easy to note that

$p_{ii}^{(1)}=f_i^{(1)}=p_{ii}^{(1)}$ for $n=1$

$p_{ii}^{(2)}=f_i^{(1)}\times p_{ii}^{(1)}+f_i^{(2)}$ for $n=2$.

Utilizing this two formulae we easily get

$$f_i^{(2)}=p_{ii}^{(2)}-f_i^{(1)}\times p_{ii}^{(1)}$$

Furthermore through simple induction we can show that

$$f_i^{(n)}=P_{ii}^{(n)}-f_i^{(1)}\times P_{ii}^{(n-1)}-f_i^{(2)}\times P_{ii}^{(n-2)}-f_i^{(3)}\times P_{ii}^{(n-3)}-\ldots-f_i^{(n-1)}P_{ii}^{(1)}$$

So the probability that the systems ever returns to its original state i is given $f_i=\sum_{n=1}^{\infty} f_i^{(n)}$

Hence

- $f_i = 1$ Þ that the system returns to state i in a certain number of steps
- $f_i = 0$ Þ that the system never returns to state i
- $f_i < 1$ Þ that the system may or may not returns to state ***i*** in a certain number of steps

One can understand that these transitional probabilities are dependent on (i) initial and final states and (ii) time of transition from the initial to the final states, i.e., $P_{i,j}^{n,n+1} = f(i, j, \Delta n)$, where $f(\cdot)$ is of some function form of i, j as well as difference in time periods.

Wald's Equations

Consider Z_i's be the sum of random numbers random variables (r.v's), i.e., $Z_i = \sum_i^Y X_i$ and remember that X_i's are i.i.d random variables (r.v's), then $E(Z_i) = E(Y) \times E(X_i)$. The basic concepts of Wald are the founding stones based on which the rich branch of Sequential Analysis (Sequential Estimation, Sequential Interval Estimation and Sequential Hypotheses) have developed.
Now if we have $E(X_0 = 1) = 1$, then $E(X_n) = \mu^n$. Without repetition we would like to mention that for the interested readers we would advice that rather than panic they should wait till we cover the simple proof of Wald's equations. Remember the importance of this equation stems from the fact that one can also find variance which is given by $V(X_n) = E[V(X_n|X_{n-1})] + V[E(X_n|X_{n-1})]$.

Stationary transition probability

When the one state transition probabilities are independent of time, i.e., n, we say we have a Markov process which has stationary transition probabilities.

Thus if $p_{ij}^{n,n+1} = p_{ij}$, then we all know the transition probability matrix is denoted as $P = \begin{pmatrix} p_{11} & p_{12} & \cdots \\ p_{21} & p_{22} & \cdots \\ \vdots & \vdots & \cdots \end{pmatrix}$, i.e., $P = \|p_{ij}\|$ and this is the Markov matrix or the transition probability matrix of the process. So generally the $(i+1)^{th}$ row of P denotes the distribution or the collection of realized values of X_{n+1} under the condition that $X_n = i$. If the number of states are finite then the transition probability matrix, $P = \|P_{i,j}\|$, is a square matrix. Also it would be true that we have

(i) $p_{ij} \geq 0, \ \forall \ i, j = 0,1,2,K$

(ii) $\sum_{j=0}^{\infty} p_{ij} = 1, \ i = 0,1,2,K$

Now the whole process, i.e., $p_{ij}^{n,n+1} = p_{ij} = Pr\{X_{n+1} = j | X_n = i\}$ is known if we know the value of X_0, or more generally the probability distribution of the process.

Assume $P[X_0 = i] = p_i$, given which we want to calculate $P[X_0 = i_0, X_1 = i_1, K, X_n = i_n]$ which is any probability involving $X_{j_1}, X_{j_2}, \ldots, X_{j_{2-1}}, X_{j_2}$,such that using the axioms of probability.

Thus

$$\begin{aligned} P[X_0 = i_0, X_1 = i_1, K, X_n = i_n] = {} & P[X_n = i_n | X_0 = i_0, X_1 = i_1, K, X_{n-1} = i_{n-1}] \\ & \times P[X_0 = i_0, X_1 = i_1, K, X_{n-1} = i_{n-1}] \end{aligned} \quad (1.1)$$

By definition of Markov process we also have

$$P[X_n = i_n | X_0 = i_0, X_1 = i_1, K, X_{n-1} = i_{n-1}] = P[X_n = i_n | X_{n-1} = i_{n-1}] \quad (1.2)$$

Substituting (1.2) in (1.1) we have

$$Pr\{X_0 = i_0, X_1 = i_1, \dots\dots, X_n = i_n\} = p_{i_{n-1},i_n} \times Pr\{X_0 = i_0, X_1 = i_1, \dots\dots, X_{n-1} = i_{n-1}\}$$

Further more in the next step we have

$$Pr\{X_0 = i_0, X_1 = i_1, \dots\dots, X_n = i_n\} = p_{i_{n-1},i_n} \times p_{i_{n-2},i_{n-1}} \times Pr\{X_0 = i_0, X_1 = i_1, \dots\dots, X_{n-2} = i_{n-2}\}$$

Using induction we finally have

$$Pr\{X_0 = i_0, X_1 = i_1, \dots\dots, X_n = i_n\} = p_{i_{n-1},i_n} \times p_{i_{n-2},i_{n-1}} \times \dots.. \times p_{i_0,i_1} \times p_i$$

To make things better for understanding and also to bring into light the usefulness
of Markov chains we give few more examples which we are sure would be appreciated by the readers.

Previous Next

Examples 1.13

Let ξ be discrete random variables such that the realized values of ξ_1, ξ_2, K , are non-negative integer values, such that $P[\xi = i] = a_i \geq 0$ and $\sum_{i=1}^{\infty} a_i = 1$ We must remember that the observations are independent. From this ξ we define two Markov processes which are.

Case 1: Consider X_n, $n = 0,1,2,K$ such that $X_n = \xi_n$ and assume (which is not at all difficult to do so considering that is the initial conditions for any process which is known beforehand) $X_0 = \xi_0$, then the Markov matrix is given by $P = \begin{Vmatrix} a_0 & a_1 & \cdots \\ a_0 & a_1 & \cdots \\ \vdots & \vdots & \cdots \end{Vmatrix}$, where the fact that each row is exactly equal due to the simple fact that the random variable X_{n+1} is independent of X_n

Case 2: Another class of important Markov chains is seen when we consider the successive partial sums, η_n of ξ_i's , $i = 1,2,K\ ,n$, i.e., $\eta_n = \xi_1 + \xi_2 + K + \xi_n$. By definition we have $\eta_0 = 0$. Then the process $X_n = \eta_n$ is a Markov chain and we can easily calculate the transition probability matrix as

$$P[X_{n+1} = j | X_n = i] = P[\xi_1 + K + \xi_{n+1} = j | \xi_1 + K + \xi_n = i] = P[\xi_{n+1} = j - i]$$
$$= \begin{cases} a_{j-i} & for\ j \geq i \\ 0 & for\ j < i \end{cases}$$

Here we use the independence of ξ_i

Thus writing the transition probability matrix we have it of the following form, which is

$$P = \begin{Vmatrix} a_0 & a_1 & a_2 & a_3 & \cdots \\ 0 & a_0 & a_1 & a_2 & \cdots \\ 0 & 0 & a_0 & a_1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots \end{Vmatrix}$$

If the possible values of the random variables are permitted to be both positive as well as negative integers, then instead of labeling the states by non-negative integers, as we usually do, we may denote them with the totality of the state space, which will make the transition matrix look more symmetric in nature, such that we denote it like

,

$$P=\begin{Vmatrix} & \vdots & \vdots & \vdots & \vdots & \vdots & \\ \cdots & a_0 & a_1 & a_2 & a_3 & a_4 & \cdots \\ \cdots & a_{-1} & a_0 & a_1 & a_2 & a_3 & \cdots \\ \cdots & a_{-2} & a_{-1} & a_0 & a_1 & a_2 & \cdots \\ \cdots & a_{-3} & a_{-2} & a_{-1} & a_0 & a_1 & \cdots \\ \cdots & a_{-4} & a_{-3} & a_{-2} & a_{-1} & a_0 & \cdots \\ & \vdots & \vdots & \vdots & \vdots & \vdots & \end{Vmatrix}$$

where $P[\xi = k] = a_k \geq 0$, $k = 0, \pm 1, \pm 2, \mathrm{K}$ and $\sum_{k=-\infty}^{\infty} a_k = 1$

Previous Next

Examples 1.14

Consider a one dimensional random walks, whose state space is a finite or infinite subset $a, a+1, \ldots, b$ of integers, such that a particle at state i can in a single transition move to $i-1$ or $i+1$ state or it can remain at the same state, i. Now all these *three* movements has some probability such that we $p_i, q_i, r_i \geq 0$, $\forall\, i = 1,2,\ldots$, and $p_i + q_i + r_i = 1$. It is also true that $p_0 + r_0 = 1$. So that we have (i) $P[X_{n+1} = i-1] = q_i$, (ii) $P[X_{n+1} = i] = r_i$ and $P[X_{n+1} = i+1] = p_i$. It is obvious that $p_0 \geq 0$, $r_0 \geq 0$ and $p_0 + r_0 = 1$, hence given these set of information one can write the transition matrix as below

$$P = \begin{Vmatrix} r_0 & p_0 & 0 & 0 & 0 & 0 & \cdots \\ q_1 & r_1 & p_1 & 0 & 0 & 0 & \cdots \\ 0 & q_2 & r_2 & p_2 & 0 & 0 & \cdots \\ 0 & 0 & q_3 & r_3 & p_3 & 0 & \cdots \\ 0 & 0 & 0 & q_4 & r_4 & p_4 & \cdots \\ 0 & 0 & 0 & 0 & q_5 & r_5 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots \end{Vmatrix}$$

Let us consider a gamble as an example of simple random walk. Suppose that we have two persons, say Ram and Shyam with initial amounts of A and B INR with them respectively. Consider the probability of Ram winning one unit from Shyam is p_k and the corresponding of losing one unit is q_k, where $k \geq 1$ and $r_0 = 0$. So in case if we denote X_n the fortune of Ram after *n* such change in position, then $\{X_n\}$ clearly denotes a random walk. It is very easy to see that once the state reaches either 0 or $(A+B)$, the process remains in that state. This process is known as the gambler's ruin, and in that case the transition probability matrix is given as shown below

$$P = \begin{Vmatrix} 1 & 0 & 0 & 0 & 0 & \cdots & 0 \\ q_1 & r_1 & p_1 & 0 & 0 & \cdots & 0 \\ 0 & q_2 & r_2 & p_2 & 0 & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & q_{A+B-1} & r_{A+B-1} & p_{A+B-1} \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{Vmatrix}$$

Different variants of this game can be constructed so that we have different examples of random walk, some of which are briefly discussed below in order to motivate the reader in the application aspect of random walks.

Previous Next

Case 1: Suppose that Ram has A amount of money and Shyam has *infinite* amount of money and they play this gamble using an *unbiased* coin. In that case the transition probability matrix looks as given below

$$P=\begin{Vmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & 0 & \frac{1}{2} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots \end{Vmatrix}$$

Case 2: Suppose that Ram has A amount of money and Shyam has B amount of money and they play this gamble using an *biased* coin, such that the probability of Ram winning one unit of money is p and losing one unit of money is q . In that case the transition probability matrix looks as given below

$$P=\begin{Vmatrix} 1 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & q & 0 & p & 0 & 0 & \cdots & 0 \\ 0 & 0 & q & 0 & p & 0 & \cdots & 0 \\ 0 & 0 & 0 & q & 0 & p & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 1 \end{Vmatrix}_{(A+B+1)\times(A+B+1)}$$

Case 3: Suppose that Ram has A amount of money and Shyam has B amount of money and they play this gamble using an *biased* die, such that the probability of Ram winning one unit of money is when numbers 1 or 2 come, losing one unit of money is when numbers 5 or 6 come, and the outcome of the game being a draw when numbers 3 and 4 come. In that case the transition probability matrix looks as given below

$$P=\left\|\begin{array}{cccccccc} 1 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & 0 & \cdots & 0 \\ 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & 0 & \cdots & 0 \\ 0 & 0 & 0 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 1 \end{array}\right\|$$

Examples 1.15

The significance of random walks is not only apparent in different examples of gambling, but also is also evident (due to the reasonable and good discrete approximations) to many physical process which are to do with diffusion of particles, or particles which are continuously colliding and being randomly bombarded, say for example gas particles in a gas chamber which we can for theoretical situation considering as adiabatic (i.e., no heat is being transferred into or out of the system) process, such that when a particle collides with another particle or with the wall of the container it rebounds with the same level of total energy. Thus as the particles are subjected to collisions and random impulses then its (any single particle) position fluctuates randomly but it does describe a continuous path. For simplicity if we consider the future position to be dependent on the present position, then the process denoted by X_t is such that $\{X_t\}$ is Markovian, where t is the time. A discrete approximation to such a continuous motion corresponds to a random walk. A classical example is the *symmetric random walk*, where the state space is denoted is the totality of all integers and if the general transition probability matrix values or elements is given by $P = \begin{cases} p & if\ j = i-1 \\ p & if\ j = i+1 \\ r & if\ j = i \\ 0 & otherwise \end{cases}$, $p > 0$ where $p > 0$, $r \geq 0$, $2p + r = 1$, and for the *symmetric random walk* we have $r = 0$, $p = \dfrac{1}{2}$

Now the question which is apparent is the fact that why are random walks so useful? Apart from the above examples which are all related to gambling, can we find some interesting examples of random walks? The answer is yes. Apart from gambling (which by itself is very interesting and exciting), random walks are frequently used as approximations to describe a variety of physical process, e.g., diffusion of particles. Let us give an example where we consider a gas particle in a box as shown in Figure 1.21, whose initial position is A (x_A, y_A, z_A) which are given by the Cartesian coordinate system X, Y and Z. Now consider after some time the gas particle is at position B (x_B, y_B, z_B).

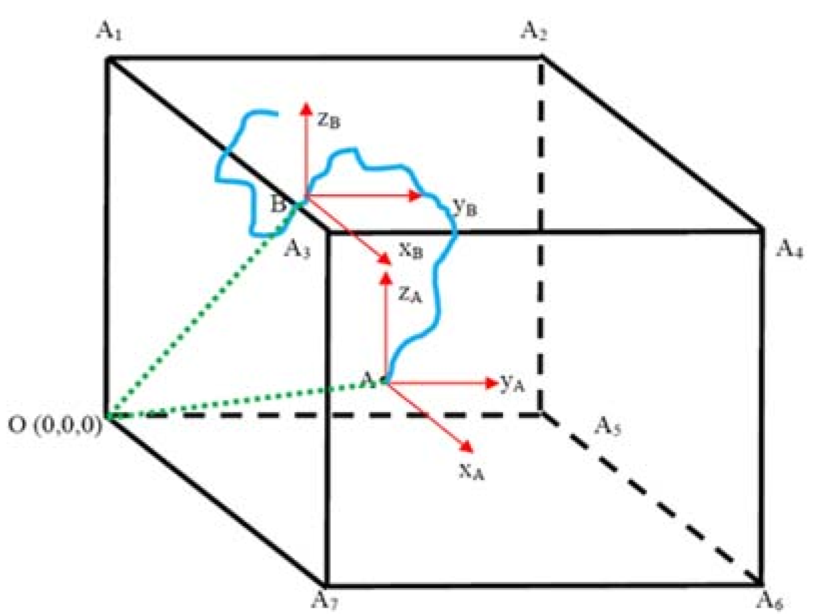

Figure 1.21: A gas molecule moving randomly inside the chamber

Previous Next

As this particle is subjected to random collision and impulses, which happens due to the particle being bombarded by all other gas particles, as well as the wall of the container, hence its position changes randomly, though the particle as such can be described by its continuous path of movement. With the important assumption that the future position of the particle, i.e., say B depends only on its present position, say A, then the process denoted by X_t is such that $\{X_t\}$ is Markovian, and t here is the time. A discrete approximation of this continuous process is provided by the random walk, and a classical discrete version of the Brownian motion is provided by the *symmetric* random walk. A symmetric random walk is a Markovian chain where the state space is the continuous real line (we consider a simple example here on), and the transition probability is given as below, i.e.,

$P = \begin{cases} p & if\ j = i-1 \\ p & if\ j = i+1 \\ r & if\ j = i \\ 0 & otherwise \end{cases}$, $i, j = 0, \pm 1, \pm 2, \ldots$ where $p > 0$, $r \geq 0$, $2p + r = 1$, and for the *symmetric random walk* we have $r = 0$, $p = \frac{1}{2}$

What is important to note about the stochastic process is the *initial condition*, based on which we can find all the characteristics of the process. Now consider the transition probability matrix given as

$$P = \begin{Vmatrix} r_0 & p_0 & 0 & 0 & 0 & 0 & \cdots \\ q_1 & r_1 & p_1 & 0 & 0 & 0 & \cdots \\ 0 & q_2 & r_2 & p_2 & 0 & 0 & \cdots \\ 0 & 0 & q_3 & r_3 & p_3 & 0 & \cdots \\ 0 & 0 & 0 & q_4 & r_4 & p_4 & \cdots \\ 0 & 0 & 0 & 0 & q_5 & r_5 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots \end{Vmatrix}$$,such that in case

$r_0 = 1$, and $p_0 = 0$, then state 0 is the absorbing state, such that once the process reaches state 0 it continues staying there. Now when $r_0 = 1, and\ p_0 = 0,$ then state 0 acts as a reflecting state, like a molecule rebounding after hitting the wall (Figure 1.22).

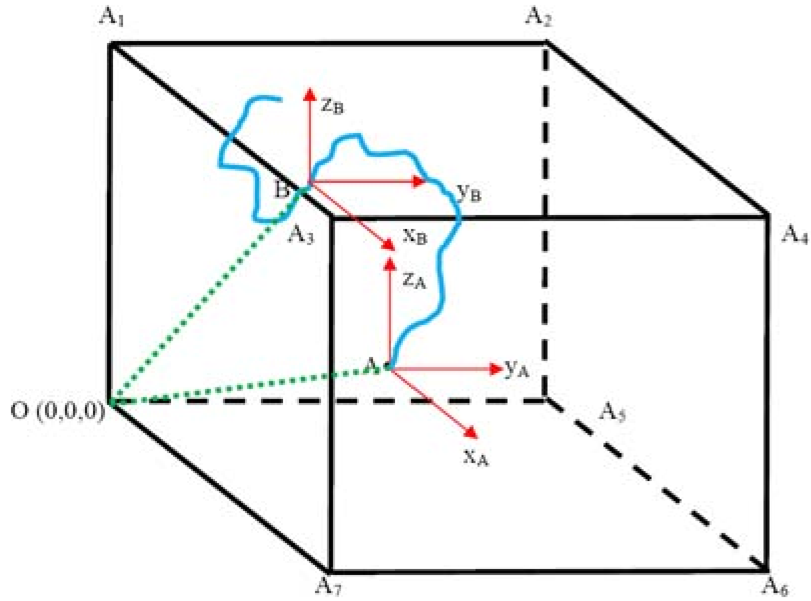

Figure 1.22: A gas molecule moving randomly inside the chamber but rebounding from the walls

Transition probability matrices of a Markov chain

A Markov is completely defined by one step transition matrix and the specifications of a probability distribution on the state of the process at time t=0. Given this, what is of main concern is to calculate the *n*-step transition probability matrix, i.e., $P^{(n)} = \left\|P_{ij}^{n}\right\|$, where p_{ij}^{n} is the probability that the process goes from state i to state j in n transitions, thus we have $P_{ij}^{n} = P\left[X_{n+m} = j \middle| X_m = i\right]$and remember this is a stationary transition probability.

Theorem 1.1

If one step probability matrix of a Markov chain is $P = \left\|P_{ij}\right\|$, then $P_{ij}^{(n)} = \sum_{k=0}^{\infty} P_{ik}^{r} P_{kj}^{s}$ for any fixed pair of nonnegative integers r and s, satisfying $r + s = n$, where we define

$$P_{ij}^{0} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

Proof of Theorem 1.1

The proof is very simple. Consider the case when, $n = 2$, i.e., the event when we move from state i to j in two transitions in mutually exclusive ways such that the first transition takes place from i to k state and then from k to j state. Here $k = 0,1,2,\Lambda$. Now because of Markovian assumption the probability of first transition from state i to state k is P_{ik}, and that of moving from state k to j is P_{kj}. If the probability of the process initially being at state j is p_j, then the probability of the process being at state k at time n is given by $p_k^{n} = \sum_{j=0}^{\infty} p_j P_{jk}^{n} = P\left[X_n = k\right]$. What is of main interest to us is to find $\lim_{n \to \infty} P_{jk}^{n}$, and for doing that we need to describe few important properties of Markov chain, which we now again do in order to have a better understanding of this process.

Properties

- State j is ***accessible*** from state i if for some integers $n \geq 0$, $P_{ij}^{n} > 0$, which means that state j is accessible from state i if there is positive probability that in a finite transitions state j can be reached starting from state i.
- Two states, i and j, are each ***accessible*** to each other then the two states are said to ***communicate*** and the notation is as follows, i.e., $i \leftrightarrow j$. In case two states do not communicate then either (i) $P_{ij}^{n} = 0$ for all $n \geq 0$ or (ii) $P_{ji}^{n} = 0$ for all $n \geq 0$ or both are true.
- Property of ***communicative*** (one should remember that communicative property is an ***equivalence*** relationship)
    - ***Reflexivity*** : , i.e., $i \leftrightarrow i$

$$P_{ij}^{0} = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

- ***Symmetry*** : In case $i \leftrightarrow j$ then $j \leftrightarrow i$
- ***Transitivity*** : In case $i \leftrightarrow j$ and $j \leftrightarrow k$, then $i \leftrightarrow k$. Now as $i \leftrightarrow j$ is true hence there exists an integer *n*, such that $P_{ij}^{n} > 0$. Also $i \leftrightarrow j$ being true, there exists an integer *m*, such that $P_{jk}^{m} > 0$. Consequently we have $P_{ik}^{m+n} = \sum_{r=0}^{\infty} P_{ir}^{n} \times P_{rk}^{m} \geq P_{ij}^{n} \times P_{jk}^{m} > 0$

Previous Next

Examples 1.16

Let us consider the example given below

	0	1	2	3	4	5
0	1	0	0	0	0	0
1	½	0	½	0	0	0
2	0	½	0	½	0	0
3	0	0	½	0	½	0
4	0	0	0	½	0	½
5	0	0	0	0	0	1

From the matrix above we can easily find out the ***equivalence classes*** as: $\{0\},\{1,2,3,4\},\{5\}$.

Examples 1.17

In case the matrix is of the following form as shown below:

	0	1	2	3	4	5
0	0	1	0	0	0	0
1	½	0	½	0	0	0
2	0	½	0	½	0	0
3	0	0	½	0	½	0
4	0	0	0	½	0	½
5	0	0	0	0	1	0

Then the ***equivalence class*** is only one and it is $\{0,1,2,3,4,5\}$

Equivalence class and Irreducible Markov chain

Let us now partition the totality of the *states* into ***equivalence*** classes, such that the states in equivalence class are those that ***communicate*** with each other. A Markov chain is ***irreducible*** if the ***equivalence property*** or relation induces only one class, i.e., ***all states communicate***. Thus in irreducible Markov chain all the states communicate with each other. For the example given below,

i.e.,

$$P=\left\|\begin{matrix} 1 & 0 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & q & 0 & p & 0 & 0 & \cdots & 0 \\ 0 & 0 & q & 0 & p & 0 & \cdots & 0 \\ 0 & 0 & 0 & q & 0 & p & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \cdots & 1 \end{matrix}\right\|$$

We have three equivalence classes, which are $\{0\}$, $\{1,2,K,(A+B-1)\}$, $\{(A+B)\}$, such that

1) $\{1,2,\ldots,(A+B-1)\} \rightarrow \{0\}$ **true**

2) $\{1,2,\ldots,(A+B-1)\} \rightarrow \{(A+B)\}$ **true**

3) $\{0\} \rightarrow \{1,2,\ldots,(A+B-1)\}$ **false**

4) $\{(A+B)\} \rightarrow \{1,2,\ldots,(A+B-1)\}$ **false**

5) $\{1,2,\ldots,(A+B-1)\} \rightarrow \{1,2,\ldots,(A+B-1)\}$ **true**

Previous Next

Periodicity

If we denote the ***period of state*** i as $d(i)$, then it denotes the greatest common divisor (also known as HCF) of all integers $n \geq 1$ for which $P_{ii}^{n} > 0$. In case $P_{ii}^{n} = 0$, $\forall\ n \geq 1$, then we define $d(i) = 0$. For the case when the transition probability matrix is denoted by

$$P = \begin{Vmatrix} r_0 & p_0 & 0 & 0 & 0 & 0 & \cdots \\ q_1 & r_1 & p_1 & 0 & 0 & 0 & \cdots \\ 0 & q_2 & r_2 & p_2 & 0 & 0 & \cdots \\ 0 & 0 & q_3 & r_3 & p_3 & 0 & \cdots \\ 0 & 0 & 0 & q_4 & r_4 & p_4 & \cdots \\ \cdots & 0 & 0 & 0 & q_5 & r_5 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots \end{Vmatrix}$$

, such that $r_i = 0$, then periodicity of each state is two (2).

Now in case if due to some reason $r_i = 0$ for some $i = i^{*}$, then periodicity of every state, i, is one (1)

The periodicity of the following case, where the transition probability matrix is as given

$$P = \begin{Vmatrix} \underbrace{\begin{matrix} 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 1 \\ 1 & 0 & 0 & \cdots & 0 & 0 \end{matrix}}_{n} \end{Vmatrix}$$

has the periodicity of each state as *n*.

Theorem 1.2

1) If $i \leftrightarrow j$, then $d(i) = d(j)$

2) If any state i has the value of periodicity as $d(i)$, then there exists an integer N (depending on i), such that for all integers $n \geq N$, $P_{ii}^{n \times d(i)} > 0$

3) If $P_{ji}^{m} > 0$, then $P_{ji}^{m+n \times d(i)} > 0$ for all *n* (positive integer) sufficiently large

A Markov chain for which ***each*** state has periodicity of one (1) is called ***aperiodic***.

Generating functions

Let us first define the concept of general generating function. In case we have a sequence, $\{P_{ij}^{n}\}$, then the generating function, $P_{ij}(s)$, is defined as $P_{ij}(s) = \sum_{n=0}^{\infty} P_{ij}^{n} \times s^{n}$ for the case when $|s| < 1$. Hence in a similar manner the generating function for the sequence $\{f_{ij}^{n}\}$, $i \neq j$, is given by $F_{ij}(s) = \sum_{n=0}^{\infty} f_{ij}^{n} \times s^{n}$, $|s| < 1$

We know that if we have (i) $A(s) = \sum_{k=0}^{\infty} a_k \times s^k$, $|s| < 1$ and (ii) $B(s) = \sum_{l=0}^{\infty} b_l \times s^l$, $|s| < 1$, then we can write the product of $A(s) \times B(s)$ as given below, i.e.,

$$A(s) \times B(s) = \left(\sum_{k=0}^{\infty} a_k \times s^k\right) \times \left(\sum_{l=0}^{\infty} b_l \times s^l\right)$$

$$= a_0 b_0 + (a_1 b_0 + b_1 a_0)s + (a_2 b_0 + a_1 b_1 + a_0 b_2)s^2 + \text{K}$$

$$= \sum_{k=0}^{\infty} s^k \times \left(\sum_{j=0}^{\infty} a_j b_{k-j}\right)$$

$$= \sum_{k=0}^{\infty} c_k s^k \text{, where } c_r = a_0 b_r + a_1 b_{r-1} + \text{K} + a_r b_0$$

Let us identify a_k's with the f_{ii}^{k}'s and the b_l's with the P_{ii}^{l}'s and if we compare $P_{ii}^{n} = \sum_{k=0}^{n} f_{ii}^{k} \times P_{ii}^{n-k}$ for $n \geq 1$ with $c_r = a_0 b_r + a_1 b_{r-1} + \text{K} + a_r b_0$ we immediately obtain $F_{ii}(s) \times P_{ii}(s) = P_{ii}(s) - 1$ for $|s| < 1$ or $P_{ii}(s) = \frac{1}{[1 - F_{ii}(s)]}$ for $|s| < 1$

One should remember that $P_{ii}^{n} = \sum_{k=0}^{n} f_{ii}^{k} \times P_{ii}^{n-k}$ for $n \geq 1$ is not valid for n = 0 and using simple arguments we have

$$P_{ij}^{n} = \sum_{k=0}^{n} f_{ij}^{k} \times P_{jj}^{n-k} \quad \text{for } i \neq j \quad n \geq 0$$

where f_{ij}^{k} is the probability that the first passage from state i to state j occurs at the k^{th} transition. Again $f_{ij}^{0} = 0 \ \forall\, i, j$, and utilizing $A(s) \times B(s) = \left(\sum_{k=0}^{\infty} a_k \times s^k\right) \times \left(\sum_{l=0}^{\infty} b_l \times s^l\right)$, we can easily show that $P_{ij}^{n} = \sum_{k=0}^{n} f_{ij}^{k} \times P_{jj}^{n-k}$ for $i \neq j \quad n \geq 0$ can be written as $P_{ij}(s) = F_{ij}(s) \times P_{jj}(s)$, $|s| < 1$

We say a state i is ***recurrent*** iff $\sum_{n=1}^{\infty} f_{ii}^{n} = 1$, i.e., a state is recurrent iff the probability is 1 so that starting at state i it will return to state i after a finite number of transitions. A ***non-recurrent*** state is said to be ***transient***.

Previous Next

Before we proceed with some proves relevant for our stochastic processes we need to see and understand two important proves, which are stated as Lemmas (Lemma 1.1 (a) and Lemma 1.1 (b)) below.

Lemma 1.1 (a)

If $\sum_{k=0}^{\infty} a_k$ *converges* (i.e., $< \infty$) then it implies (i.e., $\Rightarrow$) $\lim_{s\to 1-}\left(\sum_{k=0}^{\infty} a_k \times s^k\right) = \sum_{k=0}^{\infty} a_k = a$

Proof of Lemma 1.1 (a)

What we will prove here is $\lim_{s\to 1-}\left|\sum_{k=0}^{\infty} a_k \times \left(s^k - 1\right)\right| = 0$ (see carefully this is what we have to prove as written above)

Since $\sum_{k=0}^{\infty} a_k$ *converges* then for any $\varepsilon > 0$ one can find $N(\varepsilon)$ such that $\left|\sum_{k=N}^{N'} a_k\right| < \left(\frac{\varepsilon}{4}\right)$ holds true for all values of $N' \geq N$, then choose that N and write

$$\begin{aligned}\left|\sum_{k=0}^{\infty} a_k \times \left(s^k - 1\right)\right| &= \left|\sum_{k=0}^{N} a_k \times \left(s^k - 1\right) + \sum_{k=N+1}^{\infty} a_k \times \left(s^k - 1\right)\right| \\ &\leq \left|\sum_{k=0}^{N} a_k \times \left(s^k - 1\right)\right| + \left|\sum_{k=N+1}^{\infty} a_k \times \left(s^k - 1\right)\right|\end{aligned}$$

Now for $0 < s < 1$ we have

$$\begin{aligned}\left|\sum_{k=0}^{N} a_k \times \left(s^k - 1\right)\right| &\leq \sum_{k=0}^{N}\left|a_k \times \left(s^k - 1\right)\right| \\ &\leq \sum_{k=0}^{N}\left|a_k\right| \times \left|s^k - 1\right| \\ &\leq M \times N \times \left|s^N - 1\right|, \text{ where } M = \max_{0\leq k\leq N}\left|a_k\right| < \infty\end{aligned}$$

s being sufficiently close to 1 we have

$\left|\sum_{k=0}^{N} a_k \times \left(s^k - 1\right)\right| < \left(\frac{\varepsilon}{2}\right)$ [this can be arbitrarily done considering any combination of N and N', such that $\left|\sum_{k=0}^{N} a_k \times \left(s^k - 1\right)\right|$ can be made lesser and lesser to $\left(\frac{\varepsilon}{2}\right)$].

Now we need to find the *second* term which is $\sum_{k=N+1}^{\infty} a_k \times \left(s^k - 1\right)$, hence we have

$$\left|\sum_{k=N+1}^{\infty} a_k \times \left(s^k - 1\right)\right| = \left|\sum_{k=N+1}^{\infty}\left(\sum_{r=k}^{\infty} a_r - \sum_{r=k+1}^{\infty} a_r\right) \times \left(s^k - 1\right)\right|$$

$$= \left|\sum_{r=N+1}^{\infty} a_r\left(s^{N+1} - 1\right) + \sum_{k=N+2}^{\infty}\sum_{r=k}^{\infty} a_r\left(s^k - s^{k-1}\right)\right|$$

Look carefully and we immediately note that the *first* term is bounded by $\left(\frac{\varepsilon}{4}\right) \times \left|s^{N+1} - 1\right|$, while the *second* term is bounded by $\left(\frac{\varepsilon}{4}\right) \times s^{N+1}$. Hence we have $\left|\sum_{k=N+1}^{\infty} a_k \times \left(s^k - 1\right)\right| \le \left(\frac{\varepsilon}{2}\right)$

Combining these two we have $\left|\sum_{k=N+1}^{\infty} a_k \times \left(s^k - 1\right)\right| < \varepsilon$, provided s being sufficiently close to 1.

Previous Next

Lemma 1.1 (b)

(b) If $a_k \geq 0$ and $\lim_{s \to 1-}\left(\sum_{k=0}^{\infty} a_k \times s^k\right) = a \leq \infty$, then $\sum_{k=0}^{\infty} a_k = \lim_{N \to \infty}\left(\sum_{k=0}^{N} a_k\right)$

Proof of Lemma 1.1 (b)

Since $\sum_{k=0}^{\infty} a_k \times s^k \leq \sum_{k=0}^{\infty} a_k$ for $0 < s < 1$, hence the case of $a = \infty$ is obvious. In case $a = \infty$, then by our hypothesis $\sum_{k=0}^{\infty} a_k \times s^k < a < \infty$ for $0 < s < 1$, hence $\sum_{k=0}^{n} a_k \leq a$ for $0 < s < 1$. Now as $\sum_{k=0}^{n} a_k$ is a monotone increasing function in n, hence it has a finite limit which will be equal to a. Here remember we utilize the result proved from (a).

Previous Next