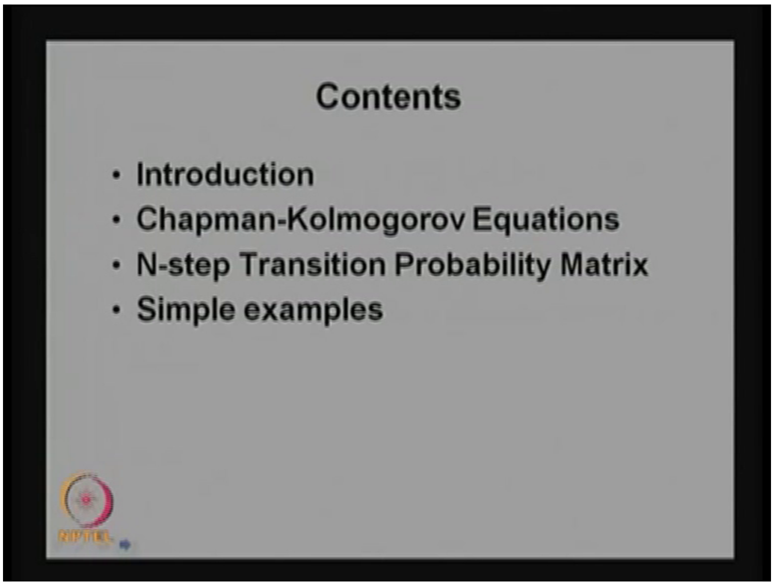
Introduction to Probability Theory and Stochastic Processes
Prof. S. Dharmaraja
Department of Mathematics
Indian Institute of Technology, Delhi

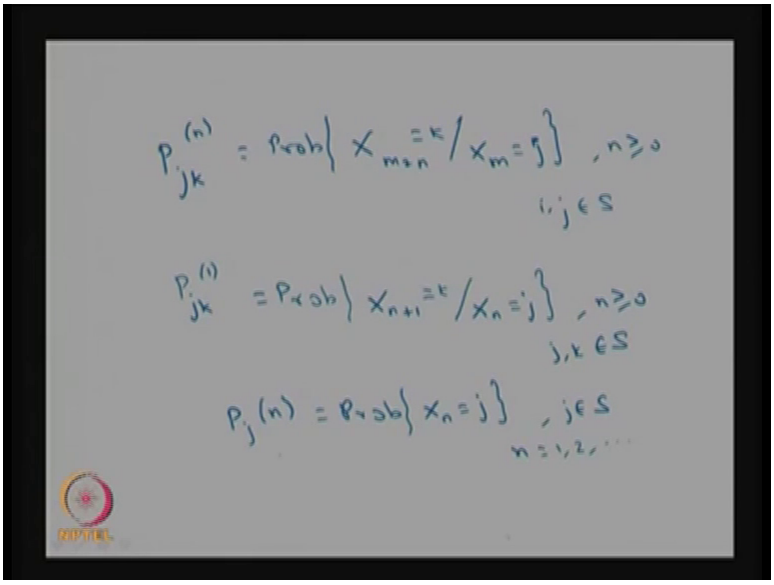
Lecture – 64

(Refer Slide Time: 00:01)



This is a discrete time Markov chain, and this is the lecture 2. And in this lecture we are going to discuss about the Chapman-Kolmogorov equations. Then we are going to discuss N-step transition probability matrix. And we are going to discuss a few more examples in this lecture.

(Refer Slide Time: 00:32)



In the last class, we have discussed the transition probability of a j to k in n steps, as the probability that the X m plus n takes a value k given that X m was j; for n is greater than or equal to 0, and j belonging to S.

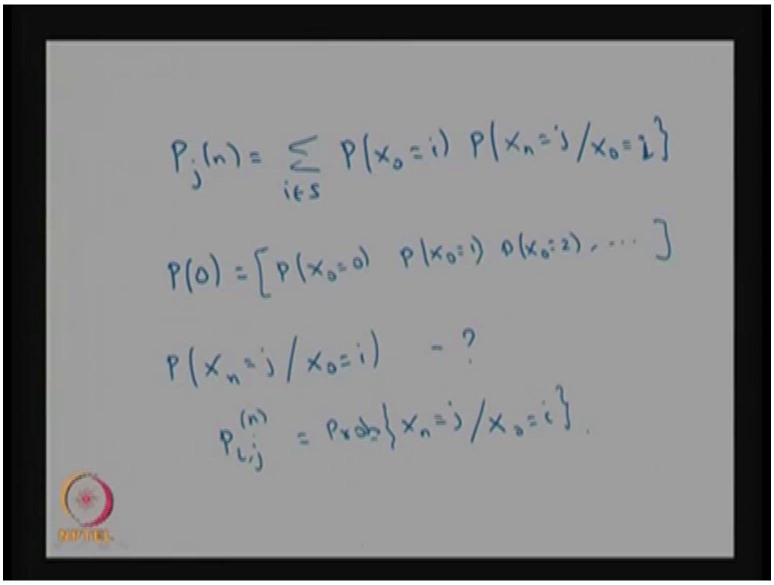
Since the underlying DTMCs is a time homogeneous, this is the N-step transition probability of system is moving from the state j to k in N-steps. So, this we denote it as a conditional probability of P j comma k in N-step transition probability; where i comma j is belonging to S, where S is a state space, and n can take the value greater than or equal to 0. Also we have discussed in the last class what is the one step transition probability of P j comma k?

It is we can write it within the bracket 1, or we can remove the bracket 1 in the superscript also. That is nothing but what is the probability that the system will be in the state k in n plus 1 th step, given that it was in the state j in the n th step. Here also j comma k belonging to capital S. So, this is a one-step transition probability.

So, our interest is to find out what is the distribution of X n, whenever the sequence of random variable X n is a time homogeneous a DTMC our interest is to find out the distribution of X n. So, it has the probability mass function the P j of n; that is nothing but what is the probability that, the system will be in the state j at the n th step. So, the j is belonging to S and a n can be 1 or 2 and so on, because you know the distribution of a n is equal to 0; that means, you know the initial probability vector of a X 0.

. So, our interest is to find out what is the distribution of X n for n is equal to 1, 2, 3 and so on. So, how we are going to find out this distribution?

(Refer Slide Time: 03:13)



So, this distribution can be written using the P j of n is nothing but the summation over i belonging to s such that the system was in the state i at 0 th step. And multiplied by what is the probability that it will be in the state j given that it was in the state i at n th at a 0 th step.

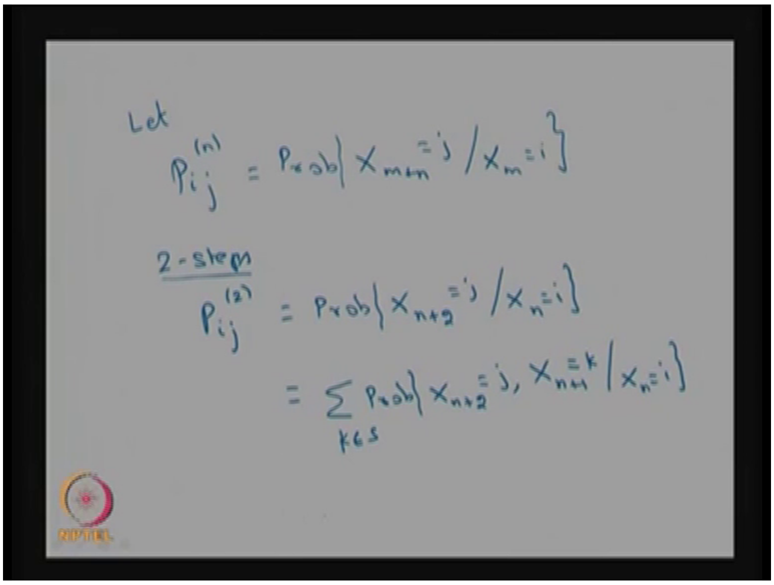
So, this is nothing but what is a probability that the system will be in the state j in the n th step, that is same, as what are all the possible ways the system would have been moved from the state i from the 0 th step to the state j in the n th step. So, this is the product of one marginal distribution, and a one conditional distribution for all possible values of a i, that gives the distribution of a X n in the n th step. So, for that you need to compute this distribution of X n, you need a n step transition probability, as well as the initial distribution vector or initial probability vector or the distribution of X naught.

So, the distribution of X naught can be given as the vector P of 0, this is a vector P of 0 it consists of the element, what is the probability that X naught takes the value 0? What is the probability that X naught takes the value 1? What is the probability that X naught takes the value 2 and so on? So, this is the initial probability vector. Why we have taken the state 0, 1, 2 and so on? Unless, otherwise I have mentioned in the set of the state space that is going to be the possible values of 0, 1, 2 and so on this is a unless, otherwise it is assume it you can take always this values.

So, this is the initial probability vector or initial distribution vector. So, what we need? What is the N-step transition probability of the system will be in the state j, given that it was in the state i at the 0 th step? This is what do you want to find out, what is the conditional probability mass function of n step transition probability vector? So, that we can write it in the form of P i comma j of superscript n, that is nothing but the probability of the system will be in the state j given that the system was in the state i at the 0 th step. That is a we need to compute the n step transition probabilities that is a P ij of n.

So, this can be computed by using the method called a Chapman Kolmogorov equations. So, this Chapman Kolmogorov equation provide a method for computing this N-step transition probabilities. So, how we are going to derive this Chapman Kolmogorov equation, that I am going to do it in the, do it now. So, we are going to derive the Chapman Kolmogorov equations for the time homogeneous discrete time Markov chain.

(Refer Slide Time: 06:47)

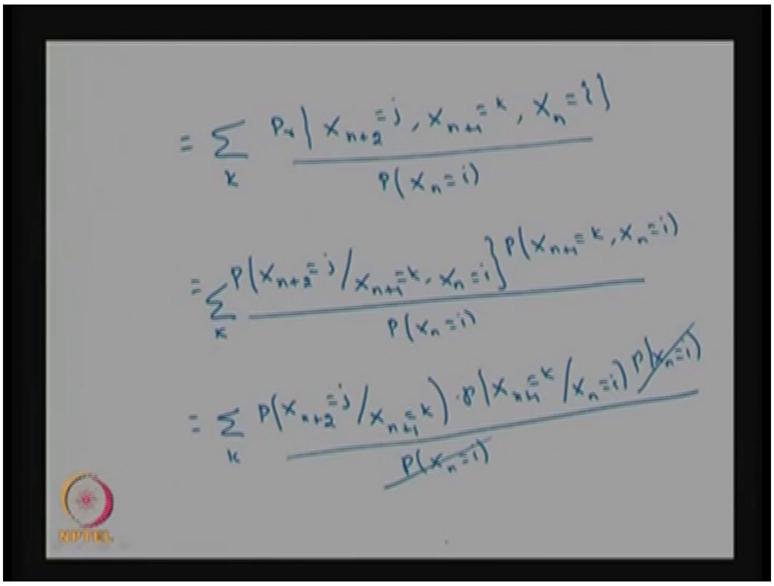


So, let the P i j of superscript n, that is nothing but what is the probability that the X m plus n takes a value j, given that X m was i. Since the discrete time Markov chain is the time homogeneous.

So, this is the transition probability of system moving from the state i to j from the m th step to m plus n th step. Therefore, this transition is the n step transition probability matrix for the time homogeneous discrete time Markov chain. Let us start with the 2-step. The 2 step is nothing but what is the probability that system is moving from the state i to j in 2 steps. So, n plus 2 takes a value j given that X n was i. This is for all n it is true because the DTMC is disc time homogeneous.

So, this probability you can write it as this 2 step transition probability of system moving from i to j the state i to the state j in 2 steps, that you can write it as, what are all the possible ways the system is moving from the state i to j by including one more state in the first step in the state is k, given that the system was in the state i in the n th step for all possible values of k belonging to S. I can write this a conditional 2 step conditional probability mass function from the n th step to n plus second step, that is same as I can include a one more possible state of k in the n plus 1 th step.

(Refer Slide Time: 09:09)



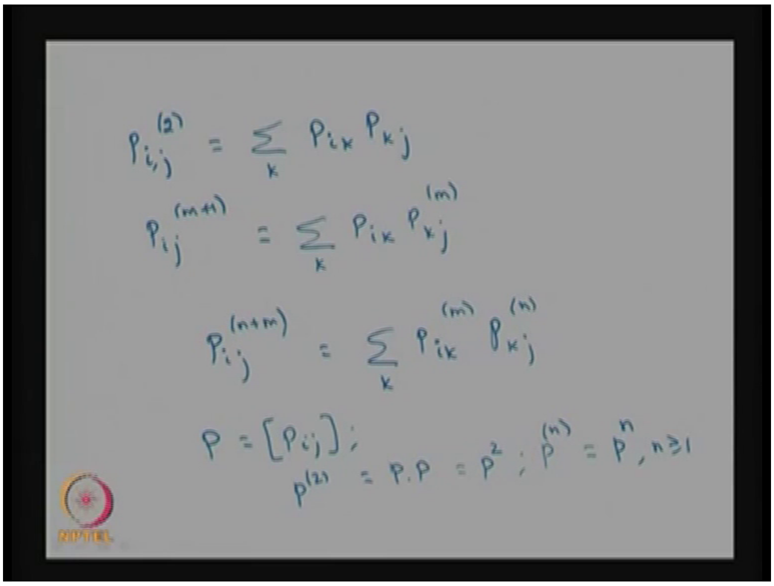
Now, I can expand these as that is same as for possible values of k, and what is the probability that the system was in the state j in the n plus second step. And the system was in the state k in the n plus 1 th step, the system was in the state i in the n th step. Divided by what is the probability that in the n th step it is in the state i. The numerator joint distribution of this 3 state, this 3 random variable, that I can write it as in the form of conditional, what is the conditional probability that the X n plus 2 takes a value j given that X n plus 1 takes a value k and a X n takes a value i product of X n plus 1 takes a value k comma X n takes a value i, divided by what is the probability that X n takes a value i and here the summation is over the k.

So, basically I am writing the numerator, joint distribution of these 3 random variable as the product of our conditional distribution with the marginal distribution of those 2 random variable. Since, the X is are the time homogeneous Markov chain, this conditional distribution by using the Markov property is same as the conditional distribution of X n plus 2 takes a value j given that only the latest value is important, the latest value is needed not the previous history. Therefore, the because of the memoryless property X n takes a value i is remote.

Therefore, this conditional distribution is the conditional distribution with only X n plus 1 with the X n plus 2. And similarly I can apply the joint distribution of these 2 random variable, X n plus 1 and X n I can again write it as the probability of X n plus 1 takes a value k given that X n takes a value i and the probability of X n takes a value i, whole divided by probability of X n and takes the value i.

So, this and this get cancelled so, this is nothing but the conditional probability; this is nothing but the one step transition probability of system moving from k to j, and the second time is a one-step transition probability of system is moving from i to k.

(Refer Slide Time: 12:11)



Therefore the left hand side we have what is the 2 step transition probability of i to j is same as all possible values of k, what is the one step transition probability of system is moving from i to k and a one-step transition probability of k to j. So, this product will give the 2 step transition probability of system is moving from the state i to j, that is same as what is the possible values of k the system is moving from the state i to k and k to j.

So, this is for the 2 step, similarly by using the induction method one can prove i to j of m plus 1 steps, that is same as what is the possible values of k the system is moving from one step from i to k, and the m steps from k to j. This is a 2 step so, this is one step from i to k and one step from k to j, by induction I can prove the m plus 1 step will be i to k and the k to j in n step. Similarly, I can make it in the other way round also, it is i to k in m steps and k to j in one step also. That combination also land up the m plus 1 step the system is moving from i to j.

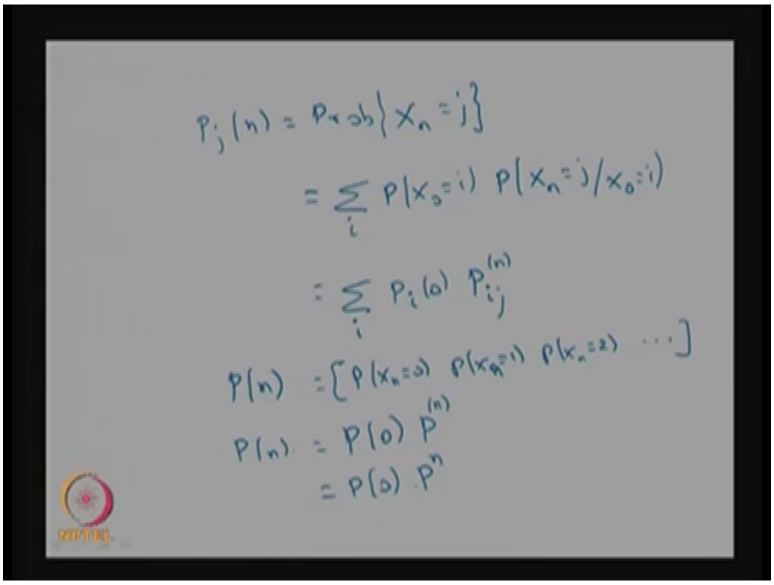
In general, we can make the conclusion, the system is moving from i to j in n plus m steps, that is same as the possible values of k of probability of system is moving from i to k in m steps, and the n step the system is moving from k to j. That will give for all possible values of k that will give the possibility of system is moving from i to j in n plus m steps. So, this equation is known as Chapman Kolmogorov equation for the time homogeneous discrete time Markov chain.

So, whenever you have a stochastic process is time homogeneous a discrete time Markov chain, then that satisfies this equation, and this equation is known as the Chapman Kolmogorov equations. In the matrix form you can write that capital P is the matrix which consists of the element of one step transition probability, one step transition probabilities. In that case, if you make a m is equal to 1 and n is equal to 1 then, the matrix of P of a superscript to 2, that is the in matrix form of a 2 step transition probability, that is nothing but if you put n is equal to 1 and m is equal to 1, you will get a P into P and that is going to be P square.

So, the right hand side piece of superscript within bracket 2 means it is a 2 step transition probability matrix, and the right hand side a P square, that is the square of the P matrix, where P is the one step transition probability matrix. So, in this form in general you can make a the n step transition probability matrix is nothing but P of n for n is greater than or equal to 1. For n is equal to 1 it is obvious for n is equal to 2 onwards the P power n, that is same as the n step transition probability matrix.

Hence, so now, we got the n step transition probability is nothing but the P power n, where P is the one step transition probability matrix.

(Refer Slide Time: 16:25)



Therefore, in matrix form I can give the P of n the P of n is nothing but in the matrix form of the distribution of a X n, or this is nothing but the vector which consists of the n th step where the system will be. So, this is nothing but what is the probability that in the n th step the system will be in the state 0, or in the n th step the system will be in the state 1, and in the n th step the system will be in the 2 and so on, this is the vector.

So, the P of n you can find out in the matrix form by using the above equation, it is going to be P of 0 that is also vector initial probability vector; multiplied by P power P of a within bracket n that is a n step transition probability matrix, but at the n step transition probability matrix is nothing but the P power n. Therefore, this is same as the P of 0 into P power n. In the last slide, we got P of a super script a within bracket n that is a n step transition probability matrix is same as the one step transition probability with the power n. Therefore, this is going to be the distribution of a X n in the vector form; that is same as the P 0 multiplied by a P power n.

Where the P is nothing but the one step transition probability matrix; that means, if you want to find out the distribution of X n for any n, you need only the initial probability vector and a one-step transition probability matrix. Because the discrete time Markov chain is a time homogeneous, we need only the one step transition probability matrix, and the initial probability vector that gives to find out the distribution of X n for any n.

So, with the help of one step transition probability matrix and the initial probability vector you can find the distribution of X n for any n.