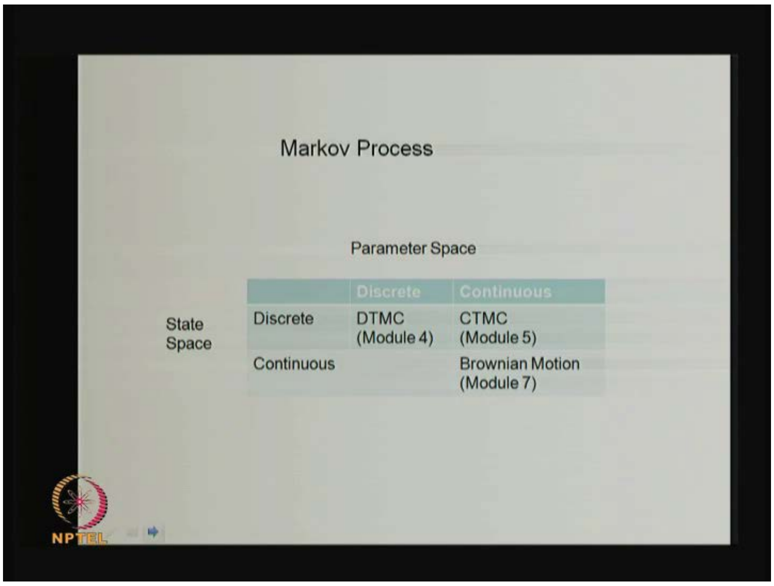
Stochastic Processes
Prof. Dr. S. Dharmaraja
Department of Mathematics
Indian Institute of Technology, Delhi

Module - 5
Continuous-time Markov Chain
Lecture - 1
Definition, Kolmogrov Differential Equations and Infinitesimal Generator Matrix

Good morning; this is stochastic processes module 5 continuous time Markov chain. I am planning for 6 to 8 lectures in this module, and I am going to start the lecture one, to the definition of continuous time Markov chain, then the derivation of kolmogrov differential equations. And I am going to give some simple examples for the continuous time Markov chain, and also I am trying to give the stationary and the limiting distributions of continuous time Markov chain in this lecture.

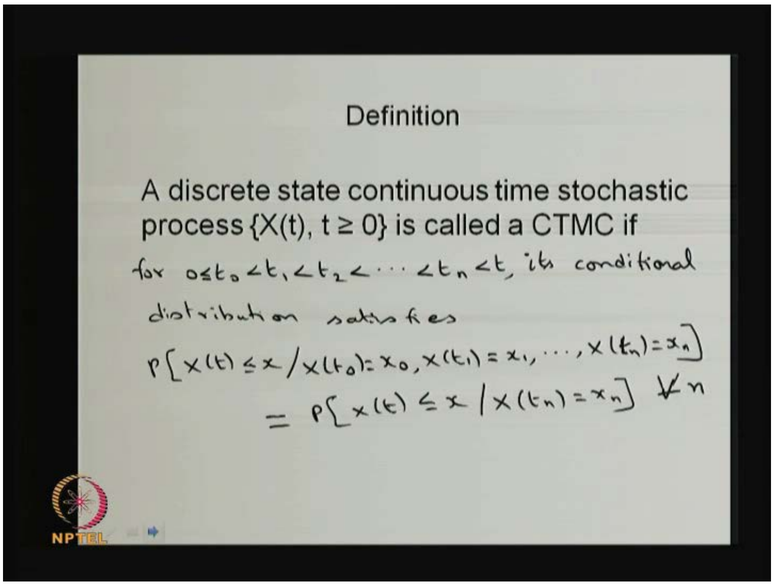
(Refer Slide Time: 01:17)



Let me start with the introduction of continuous time Markov chain. The continuous time Markov chain is a special case of stochastic process. This is the stochastic process, in which the Markov properties satisfied; therefore, it is called Markov process. Based on the classification of a state space, and the parameter space, whether it is a discrete or continuous, we can classify the Markov process. Suppose the state space is a discrete, then we say that Markov process is a Markov chain. Along with state space is a discrete, if the parameter space is also discrete, then we say discrete time Markov chain; that means, a stochastic process satisfying the Markov property, state spaces discrete, and the parameter space is also discrete. This we have discussed in the module 4.

A stochastic process satisfying the Markov property, and state spaces discrete, and the parameter space is continuous, then that stochastic process is called the continuous time Markov chain, that we are going to discuss in the module 5. There are other types of Markov process also, which has the state spaces continuous, and the parameter space is also continuous; that is called the Brownian motion or wiener process, that we are going to discuss in the module 7. Now in this lecture we are going to discuss the continuous time Markov chain, under module 5.

(Refer Slide Time: 03:07)



Let me start with the definition; definition of continuous time Markov chain. A discrete state, continuous time; that means, the state space is discrete; that means, the possible values of the random variable going to take the value, for possible values of parameter space; that is going to be finite, or countably infinite; therefore, the state space is going to be call it as a discrete. Continuous time means, the parameter space, or the possible values of the t, that collection is a uncountably infinite; therefore, it is called a continuous time, that means a parameter space is continuous. So, a discrete state continuous time, stochastic process the X of t, for t greater than or equal to 0, need not to be t greater than or equal to 0 also, but here I am making the very simplest one.

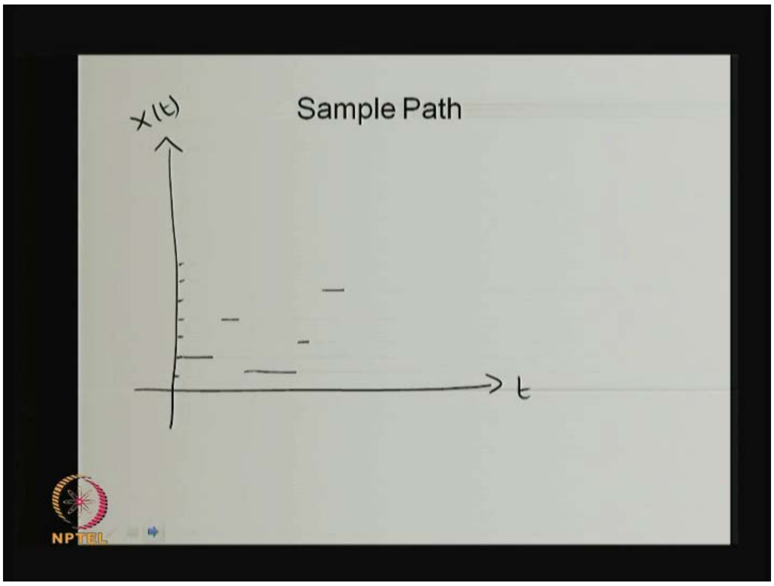
So, the x of t, for fixed t it is a random variable, for every t that collection, that is going to be a stochastic process, and the state space is discrete, and parameter space is continuous, and that stochastic process is going to be call it as a continuous time Markov chain, if it satisfies the following condition. If you take n time points, arbitrary time points, n plus 1 time points; that is t naught to t n, we can said t naught can be 0 also. and with this in equality t naught less than t 1, less than t 2 and so on t n. and you take the any arbitrary t; that is t n less than t, if this in equality. For fixed t that x of t is going to be a random variable. therefore, now we are going to find out the conditional distribution for this n plus 1 random variable, with the random variable x of t; that means, at t naught you have a x of t naught; that is a random variable at a t 1, x of t one is a random variable.

Similarly, at t n, x of t n is a random variable. you have n plus 1 random variable, with this n random variable given; that means, it takes already some values, with x naught x 1 x n so on respectively, and you are finding the conditional CDF for the random variable x of t, so that means, you have n plus 2 random variables, taken at the arbitrary time points t naught to t n, as well as small t. you are finding the conditional CDF of the random variable x of t, given that already the other n plus 1 random variables, taken at those arbitrary time points, it taken the value x naught x 1 and so on till x n, it is taken already this values, that conditional distribution, conditional CDF. If that is same as again it is a conditional CDF of x of t, given the last random variable x of t n is equal to x.

So, this n plus 1 time points are arbitrary time points, so if it is satisfies for all n, for every n; that means, the conditional distribution of n plus 1 random variable is same as the conditional distribution of the last random variable, if this properties satisfied by the discrete state continuous time stochastic process, for arbitrary time points, then that stochastic process is called continuous time Markov chain. this is very important concept this is called Markov property; that means, the t is sort of future. So, what is the probability that the random variable be in some state at the future time point t, given that you know the present state; that is where the system is in time point t n; that is small x n, and I know the past information, starting from x of t naught till x of t n minus 1, I know the information; that means, what is the probability that future the random variable x of t will be in some state, given that it was in the states x naught at time point t naught, it was in the state x 1 at the time point t 1 and so on.

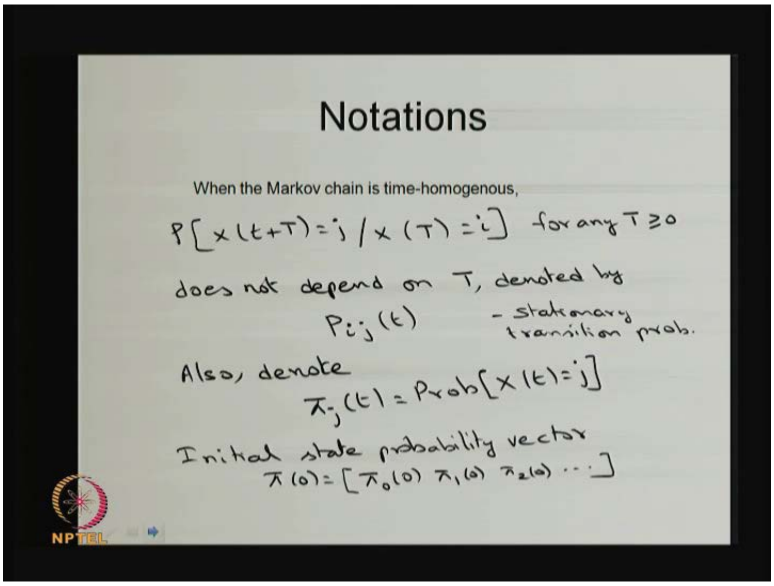
Latest at the time point t n, the system was in the state x n, that is same as what is the probability that the future the random variable will be in some state at time point t, given that it is now in the state x n at the time point t n; that means, future given present as well as the past information is same as, future given only the present, which is and independent of the past information; that is called the memory less property or Markov property. So, since this properties satisfied by the stochastic process, which has the state space is discrete, and parameter space is continuous, then that stochastic process is called continuous time Markov chain, so this is the definition. now we are going to give some more properties over the continuous time Markov chain, and some simple examples, as well as the I am going to explain the limiting distribution and the stationary distribution for continuous time Markov chain in this lecture.

(Refer Slide Time: 09:23)



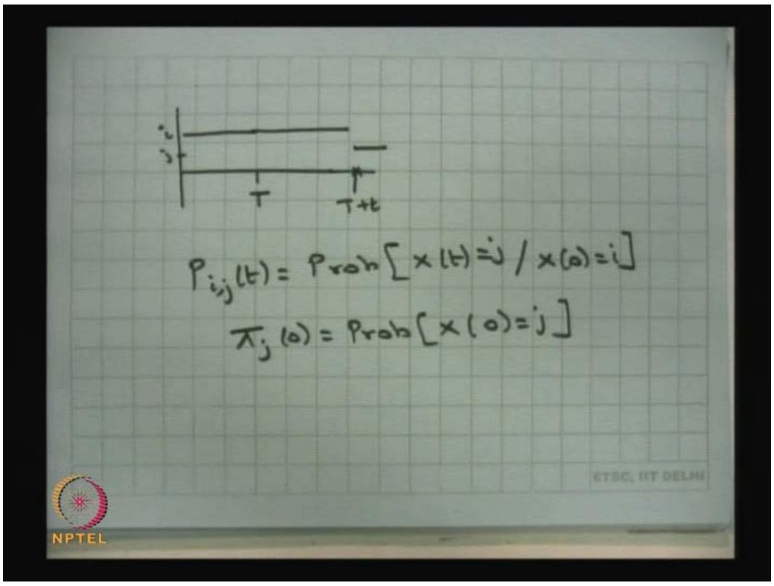
Let me show the sample path, over the time t; that is x axis, the y axis is X of t. So, the system was in some state at time point zero. It was in the same state for some time, then it moved into the some other state, then it was there in that state for some time, then it moved into some other state and so on. If you see the sample path, the following observation; the system can stay in some state for some amount of time, after that it will move to the some state. So, there is no equal interval of a system going to be in some state also. It can be some positive amount of time the system can be in the some discrete states. So, here the observations are the; state spaces discrete, whereas the parameter space is continuous, and the time spending each state; that is going to be a some positive amount of time, before moving into any other states. So, this is the observation in the sample path which I have drawn.

(Refer Slide Time: 10:43)



Now, I am going for few notations to study, or to study the behavior of a continuous time Markov chain. Whenever, the Markov chain; that means, here it is a continuous Markov chain, it is a time homogeneous. then the conditional probability of system being in the state j, at time point t plus capital T given that, the capital T it was in the state i; that does not depend on capital T. here we assume that the state changes from i to j at a future time point t plus capital T. this transition probability says, the system was in the state i at the time point t.

(Refer Slide Time: 11:41)



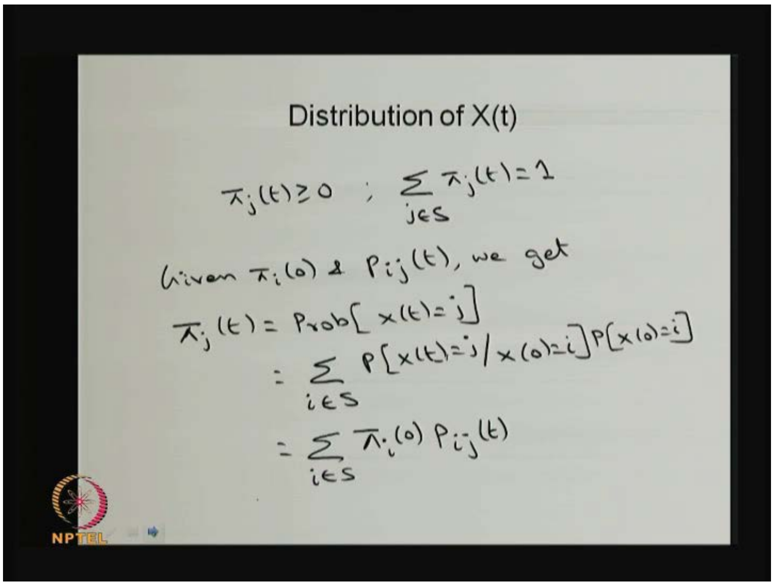
Let me draw the simple diagram; the system was in the state i at the capital T, then what is the probability that, the system will be in the state j, what is the probability that the system will be in the state j at the time point T plus t. It is independent of a capital T, whenever the Markov change is going to be a time homogeneous. For any t greater than or equal to 0; that means, the actual time does not matter, only the length matters, the length of the transition time. That means the small t is matters not the capital T, whenever it is a time homogeneous; that we can denote it as a p i j of t, because it depends on only the interval, not the actual time; therefore it is a function of small t, p i j of t; that means, that is the transition probability the systems. So, the same thing can be written as, the pi j of t, this is a notation. What is the transition probability that the system was, what is the probability that the system will be in the state j, given that it was the state i at time 0.

Since it is valid for any interval of T to T plus t, it is independent of capital T. Therefore I can represent these transition probability as, probability that the system in the state j at time t, given that it was in the state i at time 0. this denoted by p i j of t. So, this notation you should remember, it’s a transition probability, with the suffix two letters i, j of t. This also call it as a stationary transition probability, stationary means it is a time invariant, only the length of the interval is matters. Similarly, I am denoting the next notation i j of t. the p j of t is a conditional probability, whereas the pi j of t is that is a unconditional one. what is the probability that, the system will be in the state j at time t.

there is a possibility system would have been coming to the state j before time t, for at time 0 itself, or it would have come before just before t, whatever it is, this probability will give the interpretation, what is the probability that.

The system will be in the state j at time t, only it gives the information at the time t, this is a unconditional probability. I need another notation for a initial state probability vector also; that is pi naught, pi naught is a vector, which consists of entities. What is a probability that, the system was in the state 0 at time 0; therefore, this I can write it as pi j of 0; that is nothing, but what is the probability that, the system was in state j at time 0. So, this is a meaning of pi j of 0, what is the probability that, the system will be in the state the system was in the state j at time 0; that is pi j of 0, like with these entities we are framing the vector; that is pi naught. So, in this we are giving a three notations; one is the transition probability p i j of t, that is a conditional probability. The other one is unconditional probability; that is pi j of t, and initial state probability vector pi naught.

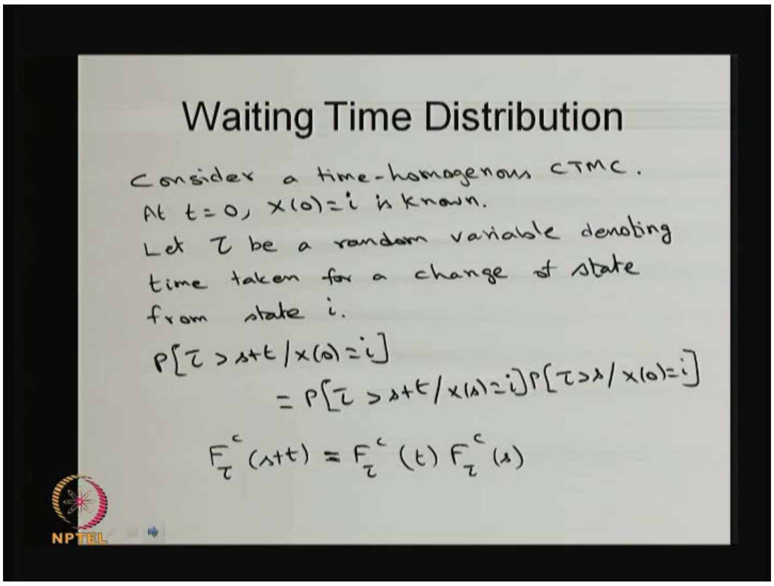
(Refer Slide Time: 15:56)



Using these, I am trying to find out what is the distribution of x of t, for any time t. for any time t x of t will make a stochastic process; here it is a continuous Markov chain. the default one is a time homogeneous continuous time Markov chain, and our interest is to find out what is the distribution of the random variable x of t. it has the probability mass function; that is pi j of t, and if you make a summation over s, where s is a state space; that summation is going to be one. If I know the initial state probability vector, with entities pi i of 0, as well as if i know the transition probability of system moving from the state i to j, from 0 to small t. I can able to find out, what is the probability mass function of system being in the state j at time t; that is pi j of t, that is same as probability that x of t is equal to j, that is same as, I can make a summation, I can make a conditional. what is a probability that the system will be in the state j at time t, given that it was in the state i multiplied by what is a probability that system was in the state i at time 0, for all possible values of i, where s is by s is nothing but the state space. I know that pi I know that the probability of x of 0 is equal to i; that is same as pi i of 0.

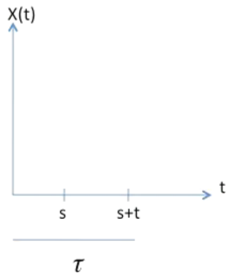
And this transition probability, since the Markov chain is a time homogeneous, so 0 to t; that is nothing, but 0 to, 0 is the time point, and t is any time point, and i is the state in which the system was in the state in the at time 0. So, p i j of t, if i multiply pi i of 0 p i j of t, for all possible values of i, I will get the probability that the system will be in the state j at time t. That means if you want to find out the distribution of x of t for any time t, I need initial state probability vector, as well as the transition probability of system moving from one state to other states, this is given. Usually the initial probability initial state probability vector is given. So, what do you want to find out is p i j of t. So, how to find the p i j of t; that derivation I am going to do it in the another two three slides.

(Refer Slide Time: 19:00)



So, before going to the p i j, you see the sample path of, the sample path of a time homogeneous continuous time Markov chain. As I said the system is staying for some positive amount of time in any state, before moving into any other states. Our interest is, what is the distribution, or what is the waiting time distribution of system being in any state, before moving into any other states; that is our interest to find out. So, how we are going to find out that I am going to explain; that is called the waiting time distribution; that means, what is the distribution of a time spending in any state, for a time homogeneous continuous time Markov chain, before moving into any other states. I assume that at time 0, the system was in the state i; that mean, x of 0 is equal to i that is known, or the probability of x of 0 is equal to i that probability is 1. Let me make out the random variable tau; that is a random variable denoting, the time taken for a change of state, from the state i. change of states means, it does not matter which state it goes, my interest is to find out, what is the waiting time distribution for the state i, the time spend in the state i, for that let me make a simple graph.

(Refer Slide Time: 20:54)



$$P[\tau > s + t \mid X(0) = i] = P[\tau > s + t \mid X(0) = i, \tau > i]$$

$$\times P[\tau > s \mid X(0) = i]$$

$$= P[\tau > s + t \mid X(s) = i]$$

$$\times P[\tau > s \mid X(0) = i]$$

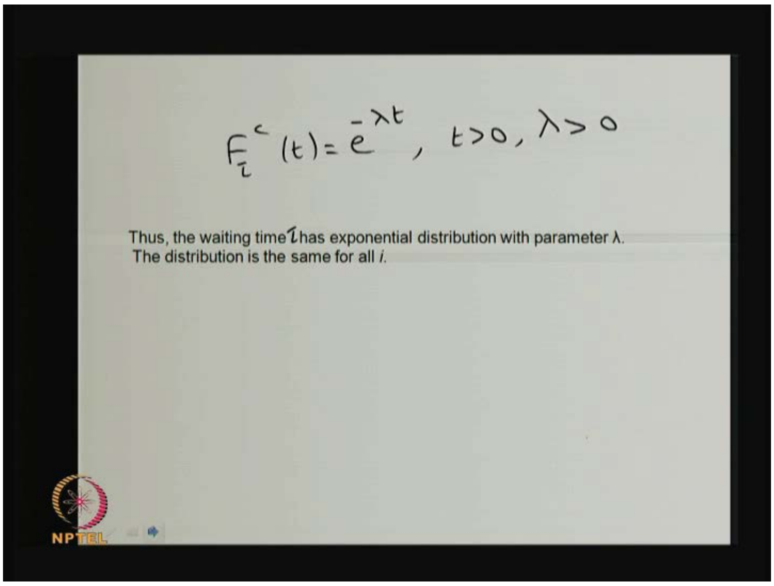
$$F_{\tau}^{c}(s + t) = F_{\tau}^{c}(t) F_{\tau}^{c}(s)$$

So, this is t, and this is X of t. Suppose you assume that the system was in the state i, at the time point 0, after some time it moved into some other state, at the time point s it was in the state i, at the time point t also it moved into some other state. So, the tau here is nothing, but the time spent in the state i from here to here; so that is a random variable. So, what I am going to do, I am going to find out what is the complement CDF for the random variable tau; that is what is the probability that, the tau greater than s plus t given that x of 0 is equal to i; that is same as, the probability of the tau is greater than s plus t given that. I can introduce one more condition, tau is greater than s, then I can multiply by using the total theorem of probability, tau greater than s given that x of 0 is equal to i; that is same as. The first one I can rewrite as a probability of a tau greater than s plus t given that x of s is equal to i.

Because x of 0 is equal to i, as well as the tau is greater than s, where tau is a time spending the state i; therefore, I can make out x of s is equal to i, by combining these two concept multiplied by the probability of tau greater than s given that x of 0 is equal to i; that is same expression here. Now the probability of tau greater than s plus t given that x of s is equal to i, that I can rewrite, because this Markov chain is a time homogeneous Markov chain. So, the s to s plus t that is same as the compliment CDF of the random variable tau, for the time t, because it is s to s plus t. since is Markov chain is time homogeneous, only the length is matters; that is the interval of length t, therefore this is nothing, but the compliment CDF for the random variable tau, with the time point t multiplied by, this is nothing but 0 to s, so this is the compliment CDF of the random variable tau the time point s. Whereas, the left hand side is the complement CDF for the random variable tau, for the time point s plus t.

So, what we got the result is, the compliment CDF of the unknown random variable tau at the time point s plus t; that is same as the product of complement CDF at the time point s and t. So, this is valid for all s and t, greater than 0. So, we have to find out, what is the random variable, or what is the distribution going to satisfies this compliment CDF at the time point s plus t, same as the product of compliment CDF at the time point s and t. If any distribution satisfies this compliment CDF property, then we can conclude the random variable, then we can find out the distribution for the random variable tau. So, in this derivation we have use a time homogeneous property, as well as the total probability rule, as well as we have use the Markov property; therefore, it land up the compliment CDF satisfying the equation.

(Refer Slide Time: 25:36)



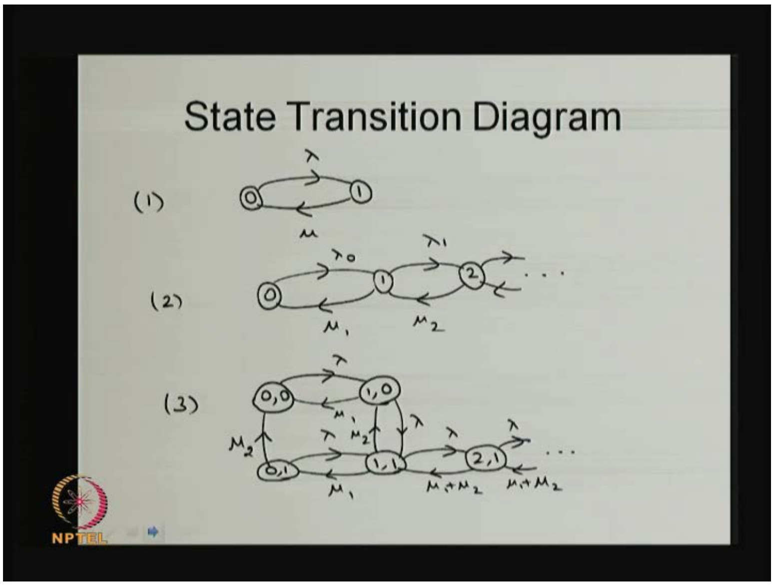
Now, we have to find out, what is the distribution going to satisfy this property. So, if you substitute any function with e power any parameter lambda, with the exponential of e power minus lambda t; the previous equation is going to be satisfied, the previous equation is going to be satisfied this equation is going to be satisfied as long as the function is of the form e power minus lambda t. for lambda is greater than 0, and t is greater than 0. Since the compliment CDF is e power minus lambda t; therefore, the CDF of the unknown random variable tau; that is 1 minus e power lambda t, for t greater than 0 for some lambda. And you know that if the CDF of the random variable is one minus e power minus lambda t, for t greater than 0 and lambda greater than 0, then that random variable is exponentially distributed random variable.

So, we can conclude the amount of time, or the time taken by the system staying in the state i, that time is exponentially distributed; that is a continuous random variable, whose distribution is exponential distribution, with the parameter lambda. even we can specify lambda suffix i; that means, it is going to be a function of a, it depends on the i; that means, the if the random variable is going to spent in some state, and that is always exponential distribution with some parameter lambda, and that parameter lambda may depend on the state i. That means if I go back to the sample path, I can say that, the time the system spending in this particular state; that is exponential distributed with some parameter, then it moved into some other state. The time spending in this state, that is also exponential distributed with some other, it could be some other parameter, it depends on that particular state.

Then it moved into the some other state, and time spending in this state; that is also exponentially distributed. And later we can conclude all these, the time spending in each state, because of it is a Markov property satisfied, the time spending this state, the time spending this state, all are exponentially distributed, which is independent of the other. So, all are going to be mutually independent random variables, then only the Markov properties going to be satisfied. That means, whenever the system is moving from one state to another state, you will have an exponentially distributed time spending in each state, and they are form a mutually independent. And since the exponential distribution has the memory less property, the system spending in this state.

If you just observe at any time t and what is probability that the system will be, for some more time in the same state, given that it was spending already this much time in this state, then that is also exponential distribution, because of memory less property of exponential distribution, and which is independent of how much time spending in the same state already. Therefore, the Markov property is going to be satisfied throughout the time, whether the system spending in this state or the other state and so on. So, the Markov property is going to be satisfied for all the time points, and the time spending in each state is exponentially distributed, and all the random variable spending in each state, all are going to be mutually independent random variables. Now, we found out what is the time spending in each state, and that is exponentially distribution with some parameter lambda i, and the distribution is same for all i, whereas, the value of the parameter lambda may be depends on the i.

(Refer Slide Time: 30:12)

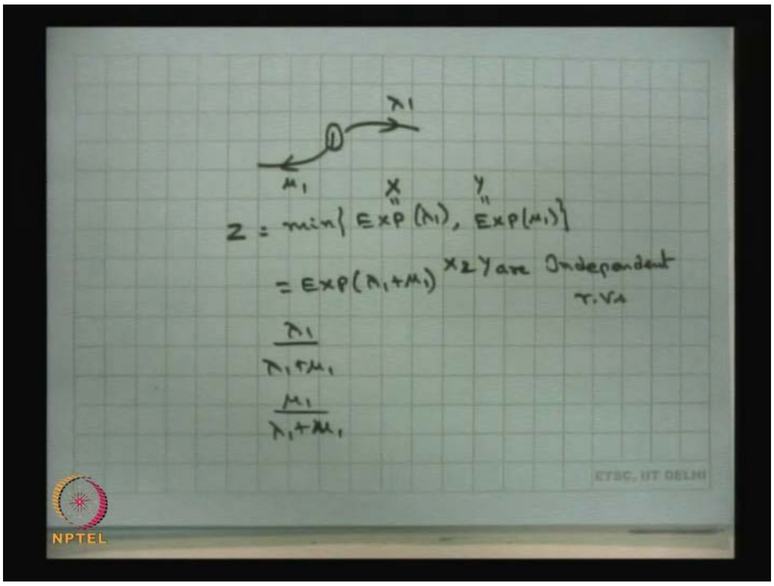


Now I am going to give few state transition diagrams, for the time homogeneous continuous time Markov chain. you see the first example, it has only two states 0 and 1, so the state space s is 0, 1, and the time spending the state 0, before moving into the state 1; that is exponentially distributed with the parameter lambda. Once, the system comes to the state 1 the time spending the state 1 before moving into the state 0; that is exponentially distributed with the parameter mu, lambda is strictly greater than 0, and mu is also strictly greater than 0. That means you know the exponentially distribution has the mean one divided by the parameter. therefore, the average time spending the state 0, before moving into the state 1; that is 1 divided by lambda. the average time spending in the state one before moving into the state 0; that is 1 divided by mu.

Since it is two state, so over the time the system will be in the state 0 or 1, and you can classify the states also, the way we have discussed in the discrete time Markov chain. Since both the states are communicating, both the states are accessible from each other, each other direction; therefore, both the states are communicating each other. Since the state space is 0 and 1, and both the states are communicating each other; therefore, this is irreducible Markov chain. For irreducible Markov chain, all the states are of the same type. For a finite Markov chain, we have at least one positive recurrence state; therefore, both the states are going to be a positive recurrence state, but here, there is no periodicity for the continuous time Markov chain. Therefore, we can conclude the first example, both the states are positive recurrent, and the Markov chain is irreducible Markov chain.

So, the continuous amount of time system spending in state 0 and 1; that is exponentially distributed with the parameter which I discussed earlier. Now I am moving into the second example; in the second example we have a state space is a countably infinite, and the system spending in the state 0, before moving into the state 1; that is exponentially distributed with the parameter lambda naught. Whereas, the state 1, the system can spend exponential amount of time, the amount of time spending in the state 1, before moving into the state 2; that is exponentially distributed with the parameter lambda one. And similarly, the system spending in the state 1, before moving into the state 0; that is exponential distributed with the parameter mu 1.

(Refer Slide Time: 33:53)

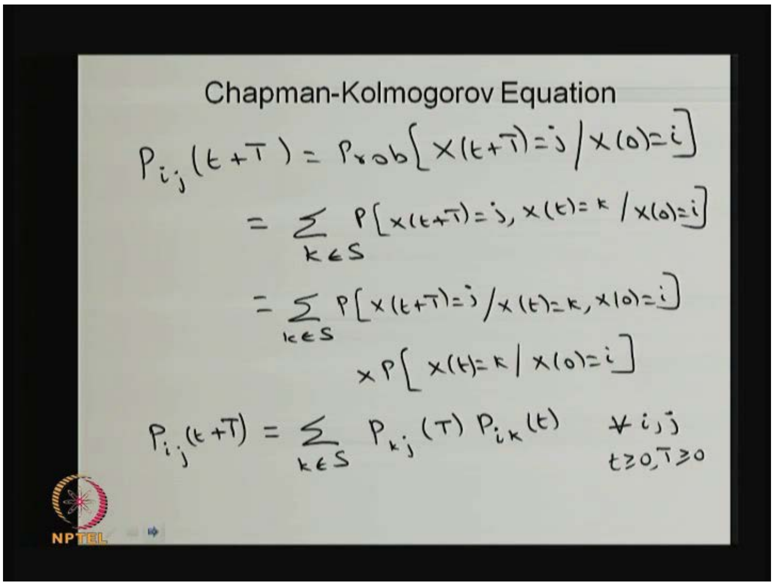


Therefore, this is mu 1, and this is lambda 1; therefore the time spending in the state 1, before moving into any other state that is going to be minimum of, the exponentially distributed with the parameter lambda 1, one random variable, you can call it as x, and the you can call it as another random variable; that is exponentially distributed with the parameter mu 1. Therefore the amount of time spending in the state 1, before moving into any other state; that is just now we have concluded that waiting time distribution is exponentially distributed that will come from here also. So, here these two random variables are independent, x and y are independent random variables, both the random variables are independent. Therefore, the time spending in the state 1, before moving into any other state that is going to be minimum of the random variable with exponentially distributed parameter lambda 1, and the random variable, which follows exponential distribution with parameter mu 1.

You know that the minimum of two exponential as long as both the random variables are independent random variable, then this is also going to be exponential distribution, with the parameter lambda 1 plus mu 1, as long as both the random variables are independent, and both are exponential. You can do it as homework, minimum of two exponential are going to be exponential with the parameter lambda 1 plus mu 1. Therefore the time spending in the state 1; that is exponential distribution with the parameter lambda 1 and mu 1. Also one can discuss, what is a probability that the system moving into the state 2, before moving into the state 1; that is lambda 1 divided by lambda 1 plus mu 1. Similarly what is the probability that, the system moving into the state 0 before moving into the state 2; that is mu 1 divided by lambda 1 plus mu 1; that also you can one can find out. So, what is the conclusion here is, the time spending in the state 1; that is exponential distribution with the parameter lambda 1 plus mu 1.

Similarly, the time spending in the state 2; that is, suppose if it is lambda 2, then lambda 2 plus mu 2. So, this is one type of a continuous time Markov chain. The third example this is also continuous time Markov chain, this sort of a 2 dimensional Markov chain with the, labeling with 0, 0 1, 0 2, 0 and so on. So, all the labeling, which is parameters for the exponential distribution. So, the change from the discrete time Markov chain state transition diagram, and the state transition diagram of a continuous time Markov chain. Here there is no self-flow, and the labels are the parameters for exponential distribution, whereas the discrete time Markov chain; it is one-step transition probability, going from one state to other states. Here the labels, the rho gives the, the time spending in the state exponential distribution with the parameter lambda naught, and moving into the straight one and so on.

(Refer Slide Time: 37:55)



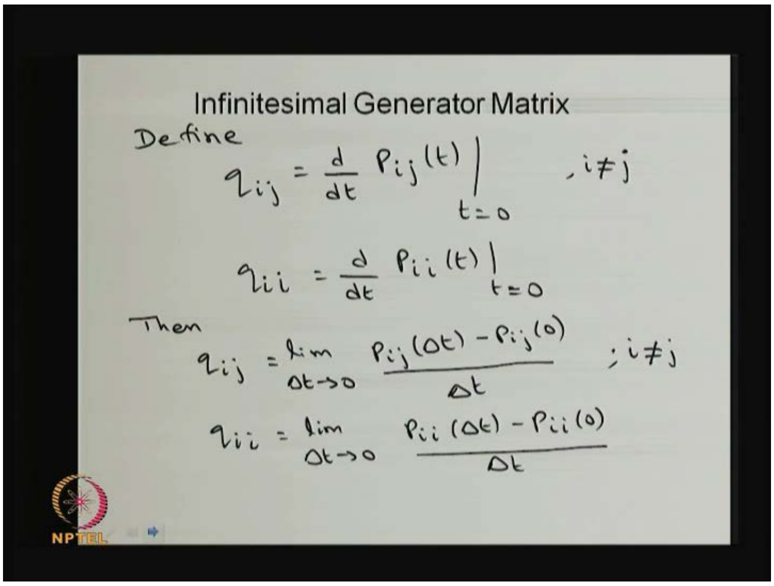
Now, I am going to find out how now I am going to find out the p i j of t, for that I am going to do the derivation, starting with Chapman Kolmogorov equation. We start with what is the transition probability of system is moving from i to j, during the time 0 to t plus capital T; that is nothing, but what is a transition probability, system will be in the state j, at the time point t plus capital T given that it was in the state i at time 0, that is same as. I can in between make some other state, I can make one more state k at time point t, for all possible values of k also, I will get the same result; that is same as, I can make a summation over k, k belonging to s, s is a state space; that is same as what is conditional probability of system will be in the state j at the time point t plus capital T, given that it was in the state i at time 0, as well as it was in the state k at small t also, multiplied by what is the transition probability of system moving from 0 to t, from the state i to k. That is same as the first conditional probability.

You see this is same as the Markov property, which will have discussed in the, definition of a continuous time Markov chain, there I have discussed the CDF cumulative distribution function, here it is the probability mass function, whereas this is the conditional probability mass function. What is the conditional probability mass function of system, will be in the state j at time point small t plus capital T, given that it was in the state i at the time point 0, as well as it was in the state k at the time point t, and you know that 0 less than t less than t plus capital T, because the way we made it is all these values are greater than 0; therefore, by using the Markov property of a continuous time Markov chain. So, this is same as what is a probability that the system was in the state k, at time small t, and move into the state j at the time point t plus capital T.

Again we use the time homogeneous property; first we use the Markov property; therefore, this is a transition probability of a t to T plus t, moving from the state k to j, then use time homogenous property; therefore, only the length matters; therefore, t to capital T; that is 0 to capital T. Therefore, the system is moving from the state k to j, from 0 to capital T; that is p k j of t. The second one it is a transition probability, system is moving from state i to k during the interval 0 to capital T; therefore, this is i to k of t. So, this is valid for all i comma j, with the t greater than or equal to 0, and capital T is also greater than or equal to 0. Therefore, the left hand side is the transition probability of system, is moving from the state i to j, from 0 to t plus capital T; that is same as summation over, I can rewrite in a different way, i to k in the interval 0 to small t, k to j instead of small t to small t plus capital T, because of the time homogenous I am just making 0 to capital T.

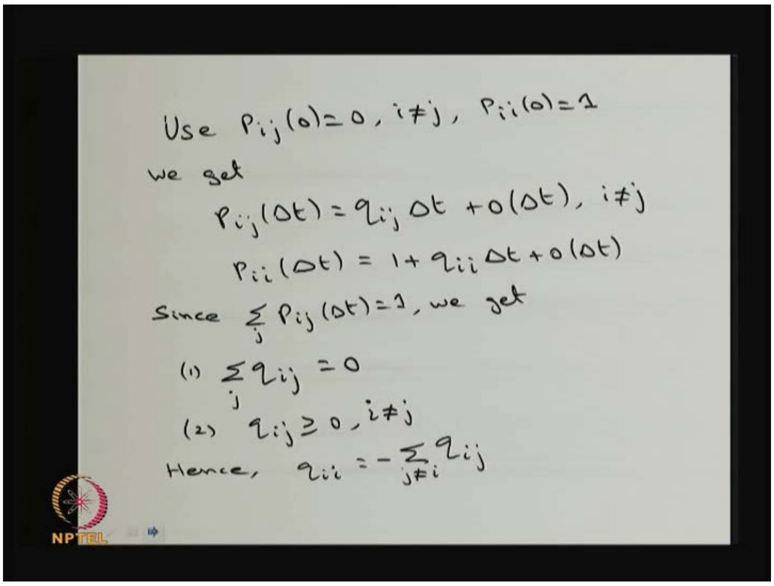
Therefore this is valid for all values of k summation; this equation is called the Chapman Kolmogorov equation for a time homogeneous continuous time Markov chain, because here for this transition probability, we have used Markov property, as well as the time homogenous property also. Therefore, this is a Chapman Kolmogorov equation of the transition probability of system moving from i to j in small t plus capital T can be broken into product of this, for all possible values of t. So, like this you can break it in many more ways with the summation, for different state of k. Using this we are going to find out the transition probability of p i j of t. You remember to find out the distribution of x of t, you need initial state probability vector as well as the transition probability p i j of t. The initial state probability vector is always given, you have to find out that the p i j of t. Once you know the p i j of t you can find out the distribution of x of t for any time t.

(Refer Slide Time: 43:20)



For that I am going to define, the quantity called q i j, and later this is going to form a matrix; that is going to be call it as infinitesimal generated matrix. So, let me start with the definition q i j; that is nothing, but take a derivative of p i j of t; that is a function of t, you can find out the derivative, it is differential function only. So, you take a derivative, then substitute t equal to 0 for all i not equal to j, then you define q i i that is also in the same way separately, because the q i i the diagonal element is going to be different from all other elements, therefore I am defining separately. You know how to find out the derivative, derivative of p i j of t with respect to t; that is nothing, but the limited delta t tends to 0, the difference divided by the delta t.

(Refer Slide Time: 44:30)

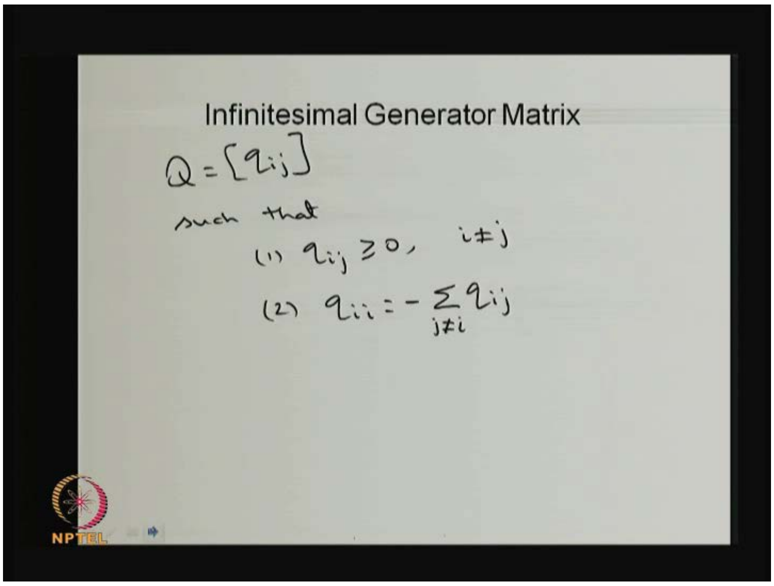


Since p i j of t is a transition probability of system moving from i to j, you can use p i j of 0 equal to 0, for i is not equal to j, for j is equal to i; that is p i of 0 that is equal to 1. That means, what is transition probability of system moving from the state i to i in the interval 0; that is same as one, that probability is 1. So, use this in the previous limit in this p i j of 0 is equal to 0, and the p i of 0 is equal to 1 substitute. Then the delta t tends to 0; therefore, the p i j of delta t, this will go to this side, so q i j times delta t; therefore, this is going to be p i j of delta t is nothing, but the q i j multiplied by delta t, plus small 0 order of delta t; that means, as delta t tends to 0, this whole quantity will tends to 0. Similarly, you substitute p i i of 0 is equal to 1 here, therefore p i i of delta t; that is same as a. This will come to this side, so 1 plus q i i delta t plus order of delta t.

So, this order of delta t; that is also tends to 0 as delta t tends to 0. You know that the summation of p i j, even at the time point delta t, with small negligible time point delta t, at the time also over the i that is equal to 1. Therefore, if you summit up, you can conclude the left hand side is a probability, right hand side for i is not equal to j you have q i j. Whereas that second expression you have one plus q i I; therefore, using the property of summation of p i j is equal to one, you will get the summation of q i j for all i for all j; that is going to be 0, when you add both the equations for all j, you will get the summation over j, the q i j is equal to 0. As well as all the q i j quantities are going to be greater than or equal to 0, from the first one, because the left hand side is a probability, and this is multiplied by the delta t, and delta t is always greater than 0.

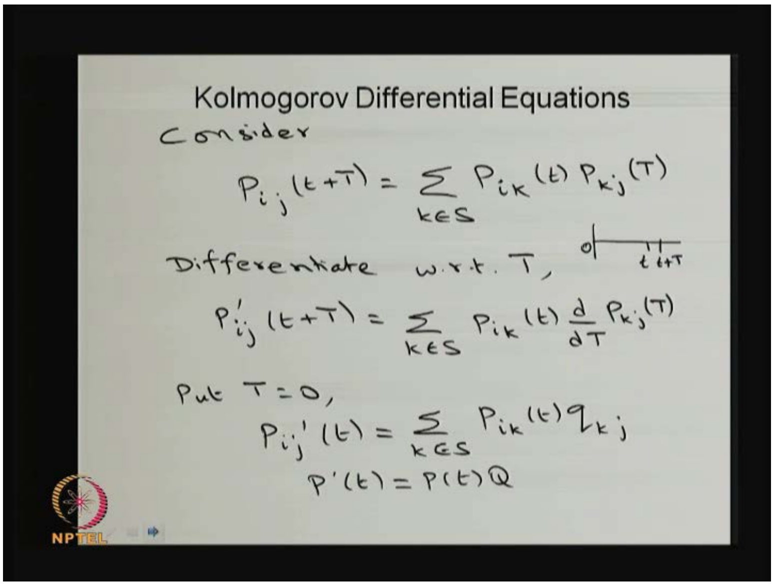
Therefore, the q i j is going to be greater than 0, for all i not equal to j, whereas, if i add over all the j; that is going to be 0. Therefore, you will get the q i i; that is nothing, but you make the summation, for all q i j for r for all j except i, then you make a minus sign, so that is going to be the q i i. That means, the diagonal element is nothing, but make the rho sum, except that the diagonal term, and put the minus sign that is going to be the diagonal term, therefore when you make rho sum is going to be 0. The details of the proof can be found in the reference books. So, the quantity q i j, that has the property, the rho sum is going to be 0, and other than the diagonal elements are greater than or equal to 0. Therefore, the diagonal element is going to be summation of a all the other terms with the minus sign.

(Refer Slide Time: 48:06)



So, using this we can make a matrix; that is going to be Q matrix, with the entities q i j; such that satisfies the property, q i j is always greater than or equal to 0 for i is not equal to j, whereas the diagonal element is minus of summation, therefore it is has the property, the rho sum is going to be 0. So, the difference between this matrix, and the one step transition probability matrix in the D T M C; that is a probability matrix. So, the entries are probability values from 0 to 1, and the summation rho sum is one. Whereas, here because k i j are obtained by differentiating the p i j's. These are all the rates, and these rates are always greater than or equal to 0, other than the diagonal elements, and the diagonal elements are, minus with the summation of all other rho elements. So, this matrix is called the infinitesimal generated matrix. Some books they use the word rate matrix also, and whereas here the rates are placed in the other than the diagonal elements, and sum of the rates could be 0. That means the probability of system moving from that particular state to the that particular state, is not possible; that probability is 0, or when there is small interval of time, there is the transition is not possible. So, whenever the rates are greater than 0; that means, there is a positive probability that the system can, have a transition of system moving from i to j.

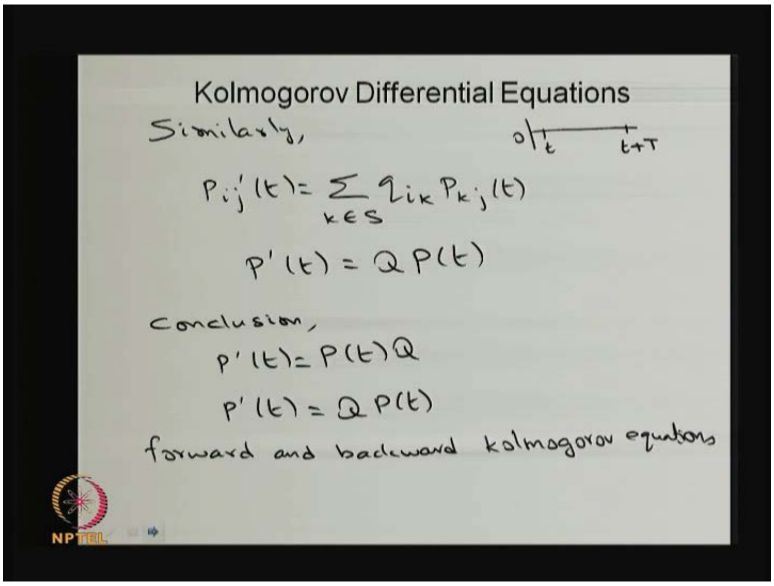
(Refer Slide Time: 49:48)



So, we have defined the q matrix. Now using the q matrix we have find out the p i j of t. So, let me start with the Chapman Kolmogorov equation. Now I am going to differentiate with respect to capital T; that means, I make the interval 0 to small t plus capital T, as a 0 to t, then I make a t to t plus capital T, differentiate with respect to capital T; therefore, the left hand side is going to be, I have written in the dash, so the derivative comes inside the p k j of t, then I am substituting t equal to 0. So, basically I am making a system to move from state 0 to small t, then there is a small interval of time from t to t plus capital T; that is the meaning interpretation of t this.

Then substituting t equal to 0, I get a, the left hand side is going to be p i j of dash t; that is same as the summation over this. Whereas, this is nothing, but the way we have defined the infinite decimal generator matrix entities. So, this is nothing, but the q k j; that is the rate in which the system is moving from the state k to j. In a matrix form, I can make it as p i j of t is going to form a matrix. So, the p dash of t; that is same as p of t times q. So, this is matrix, and the p of t is also matrix, and this is the p dash of t means, each entities are differentiated with respect to time t. So, this is in the matrix form, and this equation is called forward Kolmogorov differential equation, because the derivation goes from 0 to t, then t to t plus T, where considering as a very small interval of time. Therefore, this equation is called a forward Kolmogorov differential equation.

(Refer Slide Time: 52:06)



In the same way, if you do 0 to small t that has a small interval of time, and t to t plus capital T, then I will get the p dash of t is equal to Q times p of t; that is called the backward Kolmogorov differential equation, whether you frame a forward equation or a backward Kolmogorov equation, if you solve that equation, you will get the p i j of t. If you solve p dash of t is equal to p of t into Q; that is a forward equation, p dash of t is equal to Q times p of t; that is a backward equation. If you solve the equation with the initial condition, because it is the differential equation, so you need a initial condition what is a probability, what is the transition probability of system moving from i to j at time 0. If you know the initial condition, by supplying that, solving this equation you will get the p i j of t. Once you know the p i j of t, then you can get the distribution of x of t. So, once you know the p i j of t, the given is pi of 0, and by solving that forward or backward Kolmogorov differential equation you will get the p i j of t, using these two you can get the pi j of t.

So, for a given pi i of 0, and p i j of t; that means, the transition probability, and the initial state probability vector, one can find out the distribution of x of t. So, in this lecture I have started with the Markov process, then I have discussed the definition of a continuous time Markov chain. And also I have given, what is the distribution of time spending in any state before moving into any other state. And also I explain the infinite decimal generator matrix, and using that how to find out the transition probability of p i j t from the Chapman Kolmogorov equation, and we got a forward as well as the backward Kolmogorov differential equations, by solving a forward or backward Kolmogorov differential equation, one can get the p i j of t; that is the transition probability. Using this equation, you can get the pi j of t; that is nothing, but the distribution of x of t. With this let me stop the this lecture, and the next lecture I will go for simple example of a continuous time Markov chain, as well as the stationary limiting distribution, and the steady state distribution in the next lecture.