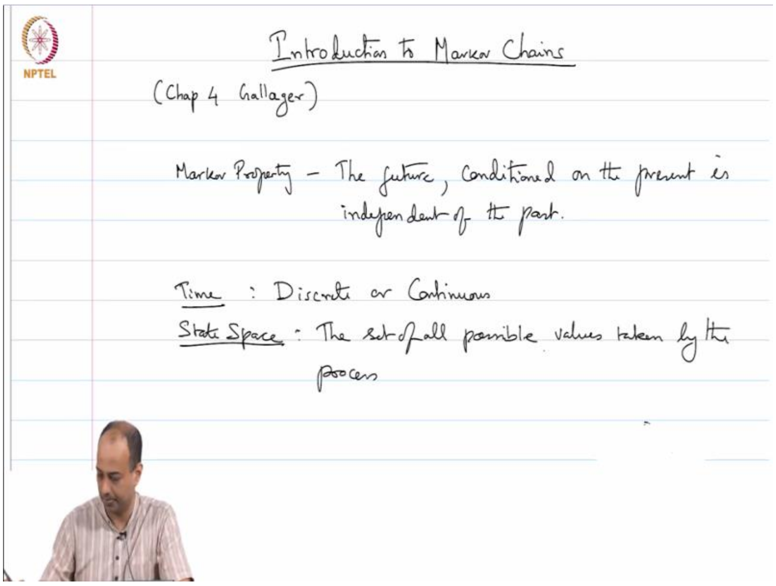
Stochastic Modeling and the Theory of Queues
Prof. Krishna Jagannathan
Department of Electrical Engineering
Indian Institute of Technology – Madras

Lecture –43
Introduction to Finite State Discrete Time Markov Chains

(Refer Slide Time: 00:16)

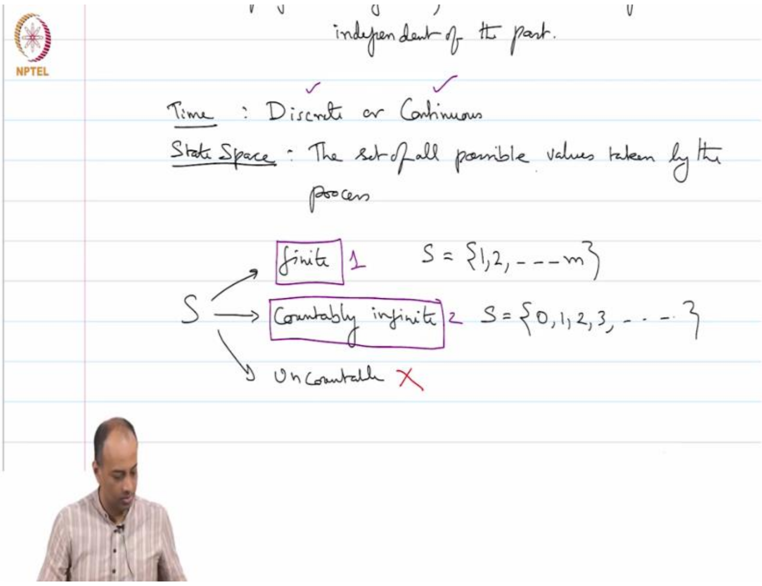


Welcome back. Today, we will start discussing the topic of Markov chains. The remainder of the course is almost entirely about Markov chains. We will discuss Markov chains in both discrete and continuous time. First, let us colloquially define what a Markov chain is? A Markov chain is a stochastic process which has the Markov property. Colloquially, Markov property is a property that the future conditioned on the present is independent of the past.

So, a stochastic process which has the Markov property is a Markov process or a Markov chain. The Markov property is the property that given the present, the future and the past are conditionally independent. So, this is a very colloquial understanding of what a Markov chain is. I will give a more proper definition in a little bit. So, a Markov chain or a Markov process is characterized by two things.

So, one is time; the time could be either discrete or continuous. So, we will study discrete time Markov chain first then study continuous time Markov chains. The other aspect of Markov processes or Markov chains is the so called state space. State space is the set of all possible values taken by the Markov process.

(Refer Slide Time: 02:39)

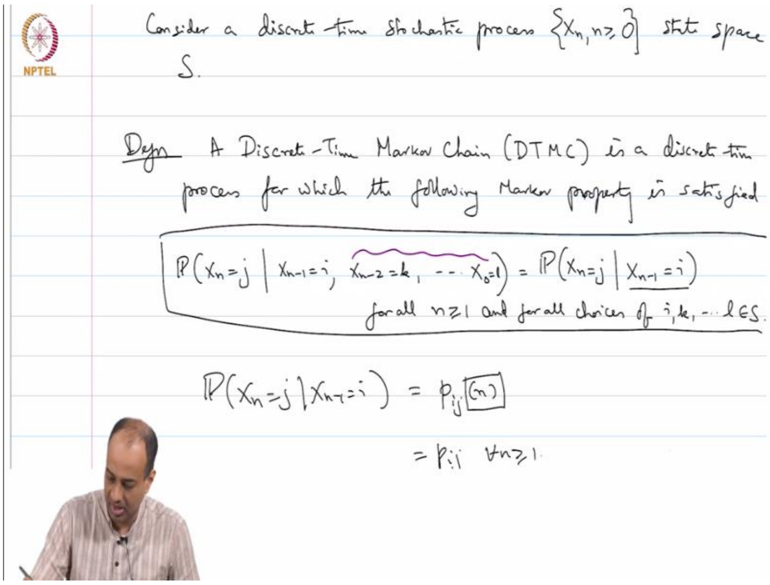


Again here that state space we will subdivide into 3 possibilities. So, let us say state space usually denoted by a set capital S could be a finite state space, it could be a countably infinite state space or it could be a uncountably infinite state space. In this course, we will study both discrete and continuous time or the time axis and for the state space we will consider finite state space first and then proceed to consider countably infinite state space.

This we will do first this we will do second. In this course, we will not consider uncountable state spaces for the Markov chains just because uncountable state space Markov process are just technically much more complex. So, we will only stick to finite and countably infinite state spaces. So, finite state space is canonically some n states or m states some finite m number of states.

So, we can just take it by default as 1 through m and for the countably infinite case state space could be all integers or all whole numbers or something like that. Uncountable state space could be a real line or something like that which we will not study. Now, let us give a definition of Markov chain.

(Refer Slide Time: 04:36)



So, in discrete time so consider discrete time stochastic process X n where n denotes time and we are looking at discrete time and state space S. S is for finite or countably infinite. So, a Markov chain let me define a Markov chain. Definition a discrete time Markov chain or DTMC for short is an integer time. We can just say discrete time process for which following Markov property is satisfied.

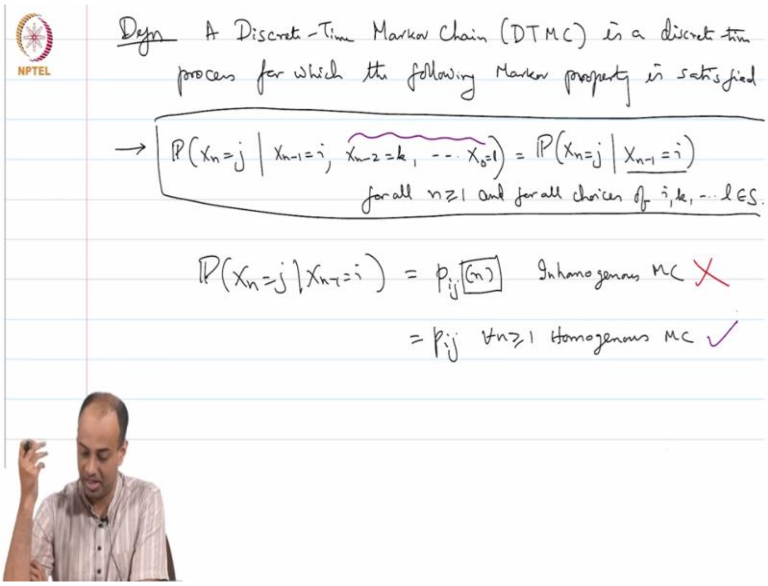
Probability that X n equals j given X n – 1 equals i, X n – 2 equals k dot, dot, dot, dot X 0. So, maybe consider discrete time stochastic process let me say n greater than or equal to 0 and just starting time at 0. Given X 0 equals l is equal to the probability that X n equals j given X n – 1 equals i and this should hold for all n greater than or equal to 1 and for all choices i k dot, dot, dot l in the state space S.

So, this is the Markov property in a more formal way. This is just saying that so let us interpret this X n – 1 equals i as the value of the state at time n – 1. We can just think of n – 1 as the present time. Now, the value of the process at time n the probability that X n takes the value j in the future is only dependent on the value i at the present time. It does not depend on the values in previous times.

So, what we are saying is conditioned on the present value i the probability that the future is equal to j is statistically independent of the state at n – 2, n – 3 and so on and this should be true for all n and all values of the past tense. So, this is the Markov property. So, you are looking at this property so P X n equals j given X n – 1 equals i is all that matters. So, in general this could be let us call this P ij of n.

So, this is the probability that the process goes from state i to state j at time n. Now this dependence on n could exist in general, but if this P ij of n is simply equal to P ij for all n.

(Refer Slide Time: 09:28)

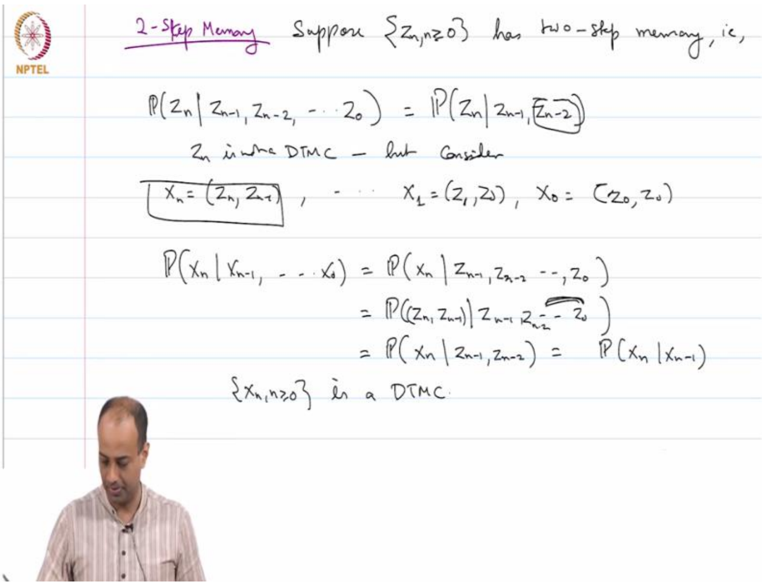


Then such a Markov chain is said to be a homogenous Markov chain. If there is an explicit dependence on n then such a Markov chain is said to be an inhomogeneous Markov chain. So, even inhomogeneous Markov chain satisfy this Markov property it is just that inhomogeneous Markov chain there is not that much you can say in general about their long term behaviour and all that.

So, whenever we talk about a Markov chain without further specification we usually do not refer to a inhomogeneous Markov chain. Whenever we say Markov chain typically we are talking about a homogeneous Markov chain. So, you can think of Markov chain as a process which has one step of memory. So, if I am in state i at this point I will not remember how I got to state i.

The probability that I transition to some other state j is some P i j it does not matter how I got to state i. So, it appears that a Markov chain is a process which has exactly one step of memory. So, recall a Poisson process has no memory at all so you can think of a Markov chain as having one step of memory. Now as it happens you can also use Markov chains to capture two steps of memory or any finite K steps of memory.

(Refer Slide Time: 11:14)



So, let me show how that can be done? Can you accommodate; so let us say two step memory. Suppose, the process Z n has two steps of memory ie probability that so let me just write it like this probability that Z n equals something given Z n – 1, Z n – 2 dot, dot, dot Z 0 is just equal to probability that Z n equals something given Z n – 1, Z n – 2. So, I have used some short hand notation.

So, typically this Markov property is written as shorthand let me just mention this here. Typically, Markov property is written as probability X n given X n – 1, X n – 2 dot, dot, dot X 0 is equal to probability X n given X n – 1. This is just a shorthand notation for the Markov property that we have defined above that big box. This just means that for every possible value of X n X n – 1 etcetera this said property holds.

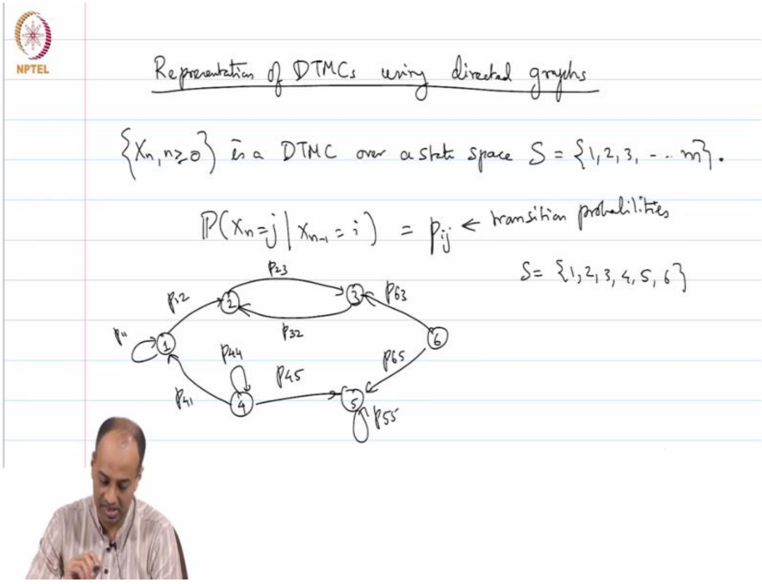
So, in this shorthand notation let us consider a process Z n which has two steps of memory. So, Z n is not a Markov chain because it depends or the value of Z n depends also on Z n – 2. It is not just a function of Z n alone. Now Z n is not a Markov chain, but you consider two tuples in this case. We define let us say X n is equal to Z n, Z n – 1 X 1 is equal to Z 1, Z 0 X 0 is equal to let us say Z 0, Z 0.

So, I am defining these two tuples of these Z as a new process X. For this new X process what you can write is the following. If you consider what is this probably X n given X n – 1 dot, dot, dot X 0 is simply the P X n given so conditioned on X n – 1 to X 0 is same as conditioning on Z n – 1, Z n – 2 so until Z 0. And this is of course X n is just Z n, Z n – 1 given so this is your X n so this is little easy.

Now, remember that this Z process has only two steps of memory. So, maybe I should write Z n – 2 here, but I can forget the rest of the guys here. So, this is simply P X n given Z n – 1, Z n – 2 that is just probability that X n given X n – 1 which means that X n which is defined as that two tuple is a DTMC. So, what this show is that even if you have a process with two steps of memory you can create these two tuples which form a DTMC.

This X process behaves like a Markov chain. Likewise, if you have a process with some K step process K sum finite integer then you can make these K tuples and use these K tuples will become a Markov chain. So, Markov chains bottom line is that as this calculation shows any finite amount of memory can be modeled using a Markov chain that is the point of this brief calculation that I have showed you.

(Refer Slide Time: 16:41)



So, the other thing that I would like to mention is the representation of DTMCs using directed graphs. Suppose, you are given a Markov chain let us say X n greater than equal to 0 is a DTMC over a state space S let us say for simplicity let us just keep a finite state space 1, 2, 3 dot, dot, dot n or m. So, this directed graph representation can be used for Markov chains of finite state space or even countably infinite state space.

The idea is the following. So, you know, as I said this is a homogeneous DTMC I do not have to keep saying this when I do not say explicitly that the Markov chain is homogeneous we are going that it is a homogeneous Markov chain. So, we are given this X n +1 let me just say X n= j given X n = X n – 1 = i is some number P ij. It has no dependence on n it is a homogeneous Markov chain.

So, what we will do is we will draw for each state in the state space S we will draw a nod let us say 1, 2 just for purpose of illustration I have just drawn a 6 date state space and in this let us say so I am going to draw directly edges for these nodes. These nodes are the states. So, there are 6 states. Here so in this example the states 1, 2, 3, 4, 5, 6 and I am going to draw P ij so let us say P 12, P 11.

So, these are all the draw P ij. So, each directed edge I am going to draw the corresponding probability. So, here P 32 is the probability that I go from state 3 to state 2. So, I can draw this Markov chain. So, these are directed edges which are basically the non-zero probabilities draw P ij. These draw P ij are known as transition probabilities because they capture the probability of a transition from state i to state j.

And in this picture I am just drawing out the state space all the states in this case 6 states or I am drawing explicitly directed edges for all those draw P ij which are stochastically positive. For example I may draw P 63 which is the probability of transition going from 6 to 3, but I do not draw an edge going from 3 to 6 that means that P 36 is 0. So, that is the way this notation works.

So, I am just drawing out an example from Gallagher's book whatever this is some example and anything that is not explicitly drawn for example P 66 there is no self loop so P 55 has a self loop that means P 55 is strictly positive, P 66 it is not positive it is 0. Likewise P 63 is positive, but P 36 is 0. So, this is a directed graph representation of a DTMC. So, you can think of; so this picture is perhaps the most easy way to visualize a discrete time Markov chain you know you can think of these states as being lily pads.

And the state as anytime as a frog which keeps jumping on these lily pads this is an analogy professor Gallagher uses. So, you can think of the frog going from state 1 to state 2 and then state 2 to state 3 perhaps and so on. The main thing to remember is that the frog is sitting on state 2 he does not remember how it got to state 2 it jumps to some state 3 with probability P 23 and in more generally when the frog is in state i he jumps to state j with probability P ij without remembering how he got to state i in the first place.

So, you can think of this frog as jumping from these various lily pads. These lily pads are the states and the frog will jump along these directed edges with those corresponding probabilities P ij. So, this is a very intuitive way to visualize these discrete time Markov chains especially in finite states or even countable states. You can draw a similar picture with countable state Markov chains except that you mean you can never really finish the diagram.

It is an infinite number of states so you cannot really finish it in the countable state case. So, this is where I stop for the first module.