Introduction to Stochastic Processes

Professor Manjesh Hanawal
Industrial Engineering & Operations Research
Indian Institute of Technology, Bombay

Lecture 29 - Introduction to Markov Property

So today we are going to start a fresh topic called Markov trends and this is, so this is a continuing our study of random processes. So when we are already defined what is a random process and in that, so far we said that, in general, stochastic processes are complex to describe if you want to describe them in generality because you need to give finite dimensional distributions of this process.

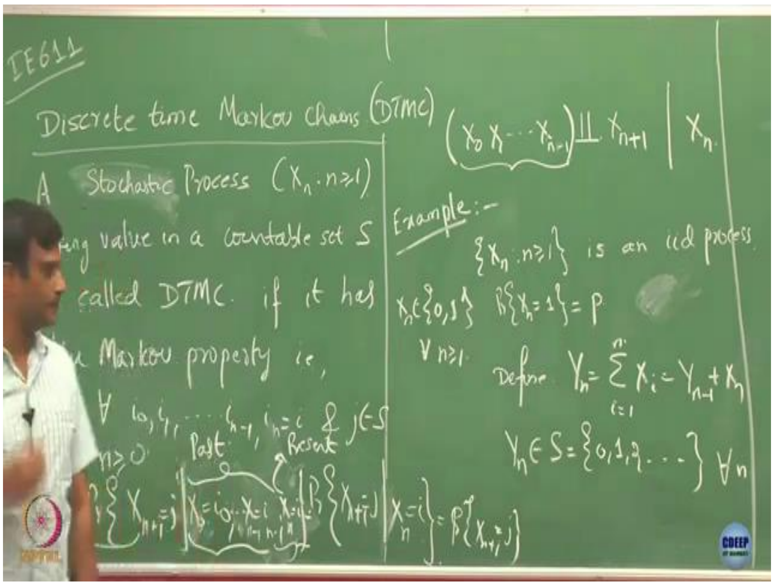
Which involves you to give joint distributions of n number of random variables in that process where n could be any number. So a complete characterization of a stochastic process is complicated but we saw that one simplest process we mostly deal with a IID process, independently and (())(01:20) distributed. And that process was very easy to describe, all you needed to describe is just give one distribution. That is the common distribution that any of the random variables are following. So and from that you could characterize all the finite dimensional distributions.

But look, IID process could be restricted. It is not necessarily like every time whatever you are doing. Like when you have defined something as a process, it is not necessary that every time what happens today is going to be totally independent of what happened the previous day or what is going to happen tomorrow is going to be totally independent of what happened today.

So IID process, basically try to model something which did not have any memory in that. Like I can say that, whatever is going to happen tomorrow is going to be independent of what happened today. And what happened today is actually totally independent of what happened yesterday. That means they are independent in a way also they are not carrying memory. So what happened previously is not going to affect me or the things that is going to happen in the future.

But often in reality, we will see that what happened today can have some impact on tomorrow. Suppose let us say the weather is going to remain, let us say very hot for few days. You will not expect weather to suddenly go and start running the next day. Like some kind effect you expect that what happened today is going to impact tomorrow. So then we want to bring in some aspect of memory in our stochastic process. And to do that, Markov chains come to our help and we are going to describe what we mean by Markov chains today.

(Refer Slide Time: 03:27)



And we will be focusing on discrete time Markov chains. So let us solve. So this is the definition of discrete time Markov chain. So we are going to take a process X n which takes value in a countable set S what does that mean? Each of any variable X n where outcome lie in a countable set. And this countable set, the elements in this countable set we are going to call 1, 2, 3 like that.

We index them. Since this is countable, we can index like that. Now, this process this stochastic process, we are going to call it as discrete time Markov chain. If you are going to take the values. i 0, i 1 all the way. Any values that belong in my set S, and I know that J and it so happens that given that you the X naught random variable with value i naught and up to Xn th random variable to value i that has happened already and now you asked the question in the n plus 1th round, I take the value i.

If this probability is just equal to this, that is, it only depends on what is the current one X n equals i. What is the state you took and the question that you are asking that it goes to stay a value J in the next round is independent of all the value it took previously. So we are going to take the value taken by these random variables as the states of this Markov chain. So every X1, X2, X3 they are all taking values.

And we are saying each one of them are going to take a value from this countable set S and we are going to call this at as set of states and these states are basically we can index them, 1, 2, 3 like that. Continue the count. Now, what we are saying is given any, let us say my process has evolved and it has taken these possible states that is my X naught it took state i naught and X1 took state i 1 and X n minus 1 took state i, n minus 1 and X n state took i.

Let us say this has been observed and now I want to ask the question. What is the probability that in the next state, sorry, in the next round I am going to take the value or I am going to take the state i? If this depends only on the current state you are now currently at and independent of the what has happened so far, then this process is going to be called as DTMC. Now, so this is J, sorry. Now, what we are going to call this quantity?

Let us say I am in round n. I am going to call it as present. Let us say the process is evolved, on some particular day nth day. Let us call that as present. So I am saying, I have been observed that, it has been observed that on day n or today I have taken the state i. So this is like my present and previously till yesterday my states are i 0, on i level up to i n minus 1. This I am going to call it as past. And now, this is my tomorrow and this I am going to call it as my future.

So what this basically this probability is saying this future given all the observation till today only depends on my today, it does not depend on my past. So if this quantity holds, then we are going to call this as a DTMC. So you see that this process has some kind of memory in it. What happens tomorrow in the future? It is going to depend at least on today. It is not like it is even independent of that.

So had it been, if this process has been an IID process, what this would have been? It will be simply probability of X n plus 1 equals to J. It would have been not even dependent on my present. Forget about the past. So when this property holds, at least there is kind of this kind of memory in this. Then we are going to call this as Markov chain. So here discrete time should be obvious because we are indexing time that are discrete valued. I am only counting 1, 2, 3 as like a time index.

This time indexes are not continuous. So most of the courses in this course we will only focus on this kind of discrete time Markov chain. And if time permits, later we will touch upon discrete time, sorry, continuous time Markov chains where my time could be continuous. So instead of this n to be taking only some integer value, it could be taking any value in some interval. So what, basically what this is saying is my process X 0, X 1 all the way up to X n minus 1 is independent of X 1 plus 1 given X n.

So as I said, this is my past, this is my future and this is my current status. So my past is going to be independent. Sorry, my future is going to be independent of the past given my current, so because of this nature, if I want to understand what is going to happening in the future, I only need to know where I am currently now, that is the only thing that is going to impact my future.

So I do not really need to know what happened to me so far. I can ignore that because that is not going to affect, what happens? What is going to happening, how the things are going to happen in the future only depends on where I am now. So X 0, X 1 all the way up to X n minus 1, this is my past. I am going to just call this entire thing as past, this all past is going to be independent of my future given this.

Student: Is it like process is discrete and random variables are also discrete?

Professor: The process is discrete because we are indexing them with discrete time. And my process is discrete valued because I am saying that this is a set taking value in a countable set. That is why this is also going to be discrete here. So, now let us take X n as an IID process. And such that this X n equals to 1 and we are just going to say, this is going to probability P and where X n belongs to 0, 1 for all n.

So let us take the sequence of random variable where each X n is taking value to be either 0 or 1 and it is going to take value 1 with probability P. So you can think of this as a sequence of coin tosses where whenever it comes heads, you are going to take it as 1 and whenever it comes tails, you can take it to 0. And probability of head coming up is P and this is obviously Markov chain because it is, we are telling it to be IID, so it is already going to satisfy my Markov property. Now, let us define another process for all n. Yeah?

Student: Sir, IID….

Professor: X n are IID.

Student: Sir, how is it following the Markov Chain?

Professor: So if my process is IID, I am saying this X n is satisfying this Markov property.

Student: The future is even not dependent on current.

Student: Current and future are independent.

Professor: Yeah. So what? So then we are saying what we want. We want this property to hold. So if my process is IID, X n plus j is anyway this like and this is going to be simply X of n plus 1 equals to j. Whether you are conditional or unconditional does not matter for an IID process.

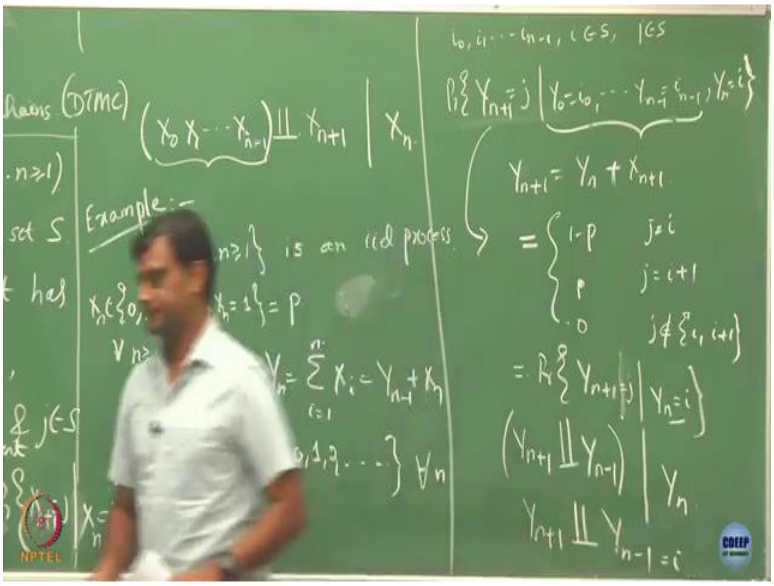
So what I am saying, so for an IID process you need to check this condition holds. For an IID process, probability that X n plus 1 equals to J given this, it is going to be simply X n plus 1 equals to J, but this itself is again this, how? If it is an IID process because X n plus 1 equals to J is independent of X n equals to i, this is a meaning of independence. So that is why this Markov property is trivially satisfied for an IID process. Now let us take this Y n, which is nothing but summation of n, this random variable.

Now we want to ask the question, whether this Y n satisfies a discrete time Markov chain? So first of all, is this Y n is going to be IID? So, let us say, I can write it as, so if I have written it, I could also write it as Y n minus 1 plus X n. I just separate out the first n minus 1. My definition, that sum is Y n minus 1 plus X. So Y n depends on previous Y n minus 1. So this future is, is not independent of my current state.

So this Y n, Y n plus minus 1 and Y n they are no more independent. So this, this sequence is definitely not independent. So then, let us see whether this satisfies Markov property. First, let us say first count what is the value taken by Y n? What is the value taken by Y n? So each of this X i's can take value 0 or 1. So what is the value taken by Y n? It can take value all the way up to 0 to n.

So and for n sufficiently large, it can take all possible values of n. So as I let n go to infinity, I mean each of these Y n’s can take, so I could, that is why I could set Y n belongs to S, I can say where S is simply 0, 1 to all the way up to n. So instead of just saying Y n belongs to 0 to n, like that I just writing 1 value where Y n’s is taking all the values from for all n.

(Refer Slide Time: 20:14)



So now let us try to understand what is this probability? Well, let us say, I want to take some for some i, j, so I want to take this set of, let us say I have some i 1, where up to i n minus 1 and i that belongs to S and I will take another j blocks to S and I want to compute this. That Y 0 equals to i 0. I want to compute this. So, let us see what is this value. So I have written Y n as Y n minus 1 plus X n, but I could then this is also say through that Y n equals to Y n plus X n plus 1. Now let us try to understand.

So now let us say this is just a relation. So now let us try to compute this. What this value is going to be? Suppose I know that if Y n is going to be some value i, what are the possible values of Y n plus 1? Either it is going to remain at i or it can take the value i plus 1. Suppose now, let us say if this guy j equals to i, what is the value of this probability? 1 minus P. Because that means tail must have happened because of that. That X n plus 1 is 0.

So that must be 1 minus P. And if this j equals to i plus 1, it should be P. And for any other value of j which does not belong to i, i Plus 1, what could be the value? Any other value just, so I have been told Y n equals to i. Let us take j is equals to anything other than these 2 values. Just let us take j to be i plus 2 can take, can j can Y n plus1 be i plus 2? No, right. Can it be i minus 1? No, and can it be any value other than i and i plus 1.

So then this is going to be 0. So, this value to compute this what I needed to know, what is the current value of Y n and nothing else I needed to know, So because of that, can I write, so these guys are immaterial. To compute this I only need to know what happened at, I only need to know what is the value at Y n. So then is this a discrete time Markov chain, Y n? That is right because it exactly satisfied by condition.

Now, what it basically showed is, in this Y n plus 1 was basically Y n minus 1 given Y n. So like say Y n plus 1 earlier I had said that this X 1 plus 1 was independent of all this, but you can take one at a time in this. This Y n plus 1 was independent of Y n minus 1. As long as you know Y n nothing behind it are useful to me. Prior that, I know what is the value of Y n.

So we are saying that Y n plus 1 and Y n minus 1 independent, but this independence is conditional, conditioned what? Conditioned that I know what is the value of Y n. But if you will just try to, but this does not imply that they are independent unconditionally. So if you just ask Y n plus 1 are independent of Y n minus 1, is this true? Without conditioning on Y n? No. Because if I know Y n minus 1 takes value i, what could be the possible value of S of Y n plus 1.

Suppose let us say, this value, this realization took value i. What are the values of Y n plus 1? It can remain at i or jump by one value or jump by two values at most. So the in that way, if I know this, I already know about this, so they are no more independent. But if I know what is Y n then they are independent of each other.

So as of now we said that Markov, the way we defined Markov property, it kinds of implied that Markov property is just talking about one step memory, that the next step is conditionally independent of everything past given the present one, what about further steps in the future? So let us say if I know something today, I know tomorrow is going to be independent of my past given my Markov property.

But does this also imply that given today, my day after tomorrow is also going to be independent of past?

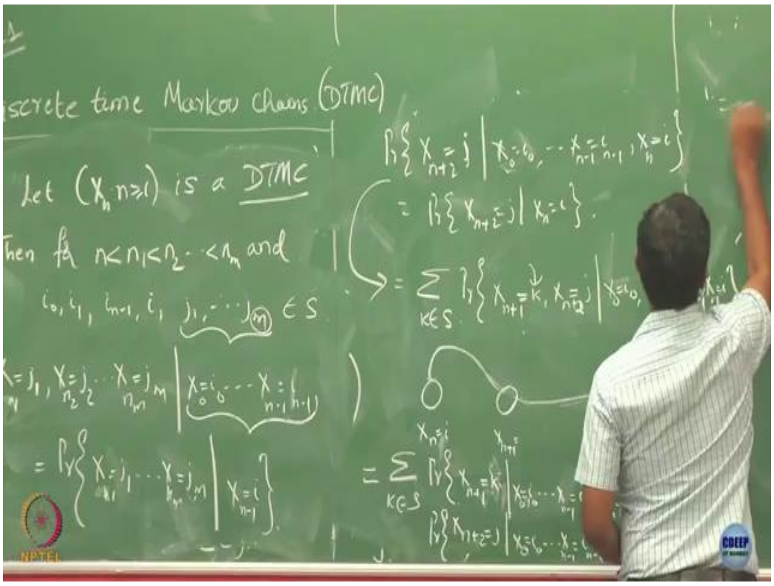
Student: Yeah.

Professor: Is that true? I mean, do you feel that by Markov property that should imply? If tomorrow is going to be independent of my past given today is day after tomorrow is also going to be independent of my today given my past?

Student: Yes.

Professor: Yes. Let us formally verify that. And that is true. Like all my future steps, however further they are, they are going to be independent of my past given my current state. So we are going to state this formally.

(Refer Slide Time: 27:09)



So let us first try to digest these notations, because we are going to use this notation again and again now in the study of Markov chains. So let us take n, when I write n, it will be usually varying my current index, my present and when I write this n1, n2 which are indexes in the future. So suppose n equals to my current day, let us say it is the 10th day. n 1 could be 12th day and n 2 could be 15th day, like that. They are the indexes in the future.

And then I would write this i 0, i 1, i n minus 1 till i to denote these are the states taken in the past, that is on X 0 took the value state i 0, X 1 took state i 1, X n and minus 1 took state n minus 1. Sorry i n minus 1. And my current day nth, I took the state i and then these are my futures. In my future date, in my next n 1th, n 1 index random variable, I am going to take state j 1 and all the way.

Let me call, so call this as n m and in that n m th date, I am going to take state j m. So this is the notation. So at time 0, I took state I naught at time n minus 1 I took state i n minus 1 and the current time I took state i and in the future this n 1 which is larger than this n, I am going into state j 1 and after that I am going into state j 2 in n 2, which is larger than n 1 like that all the way up to n m.

So what I am looking is from my current time, I am looking at m steps in the future. I mean these steps are not immediate m step. They could be at some point in the future that could be n 1, n 2, n 12 like. If I am in day 10, it could be 12th, 15th, 16th like that. If I am looking at 3 points in the future. Now it is all it is saying is, this past again does not affect this future probabilities given my current state.

So far my DTMC by the way I define it just said that it did not worry about all this. It only worried about n which was n plus 1. Now if I am saying, okay, this n 1 could be larger than n, sorry larger than n plus 1 and also if you look at all the things that we are going to look into the future, that is going to be independent of this. So if your stochastic process is DTMC, this is also guaranteed. Let us see why this is the case.

So I am doing it going to do it for, but I said I am already in todays.

Student: I do not know what it means.

Professor: That is what I am saying. What I am saying? Today, I know and I am asking about tomorrow. If you want to assume that, today also, I do not know then go back to previous day. Let me know about previous day and ask for today and the same definition applies, and whatever you want also like find. If you do not know today, where you know about yesterday, you can go and do the same thing here.

Student: If you just remove X n equal to i then this…..

Professor: If you remove this part, no, it does not hold. This is true if you know what is today, what is happened, then we are saying future is independent of the past. Now, so suppose we are saying this is not there. Then it should be n minus 1. This will be my latest information. This will hold. We will find the…..

Student: X n also we are going to find the….

Professor: No, this is like whatever like he said instead of i, I only know till the previous day let us say, but I have to assume that, that state is i here. In that case this is about future. This is still going to be independent of my previous days. All the 1 from n minus 2, n minus 3, till 0. So whichever, so basically what it is, you have some information till some point and after that you want to ask about future. Now whatever information you have, to know about the future you can take the latest information in that and throw away all the others. That is what our DTM Markov chain is telling. It need not be like till today and everything. So the generally what we see that whatever information I have only the latest matters affects the future and past I can just throw.

So what I am going to do is I am just going to do it for a special case and that will give you the idea how this should be done for rest of your, for how that can be generalized to this case. So I want to know, do it for this case. X n plus 2 equals to j; given X naught equals to i naught, X n minus 1 equals to i n minus 1 and X n equals to i. So now I am basically asking tomorrow, I am asking for day after tomorrow. Whether probability that my what happens day after tomorrow is conditionally independent of what happened so far given my today.

So how to show this? How to show that this is equals to this? So then if I can show this, you could also extend that and say that okay, even for the future this can be applied. Now how to do this? So let us take this, I am going to write it as, is this correct? So this is like conditional probability. Now what I have done is I have added this X n plus 1 but now I am allowing you to take all possible values. X n plus 1 can take value k and now I am allowing this k to take all possible values and then summing over it.

Then is this is equals to this? So what I am basically doing, let us say I am in this, where this is X n plus X n state and this is X n plus 2. So basically here it was asking okay, X n, now I want to go to X n plus 2 equals to j. This is today and this is day after tomorrow. So before you reach day after tomorrow, you have to go through tomorrow. Let us say this is tomorrow.

And now tomorrow can be any of the possible states in S. So I am allowing it, you go through tomorrow through whatever possible state. I am done if I do this, then so basically I am going from here to here and I am allowing all possible routes. That is different possible values of k. So that is why this probability should be equals to this. So now let us, so after writing this, let us go back to our initial things we studied how to split these probabilities.

I am going to split this property as X n plus 2. May be first this. X n plus 1 equals to k, given X naught, so I am just simplifying this into probability that X n plus 2 equals to j given, X naught, i naught all the way up to X n minus 1 equals to i n minus 1, X n equals to i and X n plus 1 equals to k. Is this correct? The way I have split the probabilities.

So I am just doing probability that a and b is nothing but probability of a into probability of b given a. So I am just taking this as event a and this as event b. Now, so now let us try to apply our Markov property to each one of these terms. Now I am asking probability that I am going to take state k tomorrow given I am in state i today and I am going to take state j day after tomorrow given that tomorrow I am in state k.

(Refer Slide Time: 40:58)

By Markov property this should be into probability that X n plus 2 equals to k given X n plus 1 equals to k. What is j?

Student: X n plus 2.

Professor: Okay. So now what is this telling you? Now you go back to this probability. What he is saying is you from today, you started in state i and you went to state k tomorrow. And then tomorrow you are already in state k and from there you go to state j day after tomorrow. If that is correct like if I sum it over all these possibilities?

So this is basically asking this question. I start from i today, I go to some state k tomorrow and from there further I go to state j day after tomorrow and now again I am looking into all possible routes. So this should be equal to X n plus 1 equals to j given X, sorry 2 equals to j given X n plus 1 equals to i. X n equals to i. So what it is saying is that what happens day after tomorrow only depends on today and I can forget what happens till yesterday. Okay, fine.