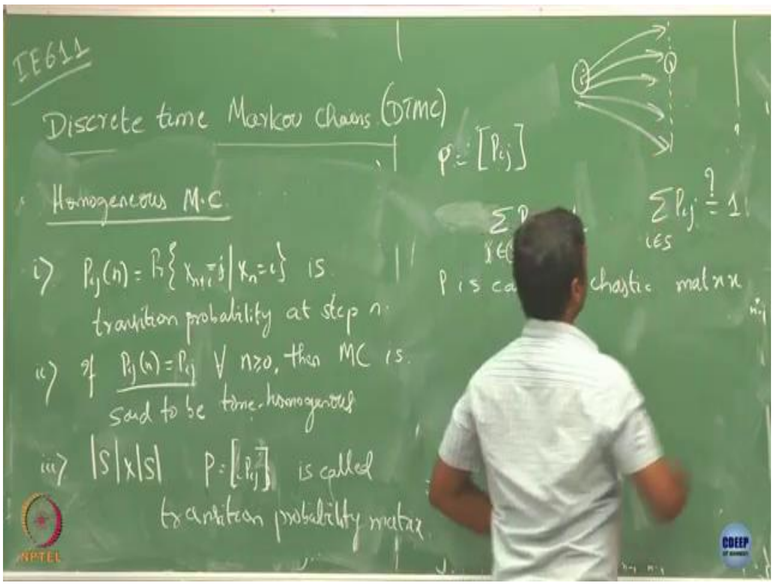
Introduction to Stochastic Processes
Professor Manjesh Hanawal
Industrial Engineering & Operations Research
Indian Institute of Technology, Bombay
Lecture 30 - Transition Probability Matrix

Now we are going to further study some properties of or we are going to further restrict this study of Markov chains to some special Markov chains called homogeneous Markov chains. So, let us understand what we mean by that.

(Refer Slide Time: 0:41)



So, before that we will introduce some definitions. We are going to define, so we are going to say suppose you are at in state n, you are in state i and in the next state, in the next step, we went to state j. You are going to denote this as P i j of n that is you are going to another state j in one step at the nth time at the time index, we are going to denote it as P i j and this is called transition probability at step n. And the way we have defined discrete time Markov chains, this need not be same for every n.

So let us say I am going from state i to j from date 10 to date 11 up to 11th day, that probability need not be same, suppose if you happen to be on the same state i in the 20th today and you going to state j again in the 21st day, need not be always same. Okay, but if it so happens that you going from some state to another state in next step happens to be independent of n. Then you are going to call it as homogeneous Markov chain. Yeah?

At step n, step n and time is same for me. Like, time is index, 1, 2, 3, 4 like that, we can call nth step or nth time slot, whichever you may, at instance n is equals to 1.

Student: What is the probability (())(4:03) which is denoted by P X n equals to….so (())(04:10)

Professor: You are just saying your state i in step n, it is just saying in the next step you go to state j, these are going these are talking about one step jumps. But at which step you are going to jump that is going to be denoted by n, suppose you take n equals to 10. So this is saying, on the 10th step you are at state i, what is the probability you go to state j in the 11th step?

If you take n equals to 100, then you are in state i in the 100th step. But the probability you go to state j in the 101th step and if these probabilities are same irrespective of which step you are, then we are going to call it as time homogeneous Markov chains, okay? That is what like this guy is now does not depend on what is n and we just denote it as P i j.

So now if I am going to construct an S cross x matrix P where this is of P i j. So what does this mean? So I am talking about state i going to state j. Now, using these elements, I can construct a big matrix P where it will tell okay, from state i. So if I take the ith row, it will give me from i different, different states that I can go, that is going to give me that row corresponds to jumping from state i to other states.

If I am going to take this matrix then I am going to call it as transition probability matrix. As I already told you what was that? It is the state space. So, if it is a countable when I say cardinality of, what is this going to be?

Student: Infinite.

Professor: Infinite. So, this S could be finite or it could be infinite, we are allowing it to be countable, but it may so happen that the number of states are only finite. In this case, this matrix is going to be finite dimensional matrix. It will have finite number of rows and columns but if my state S is going to be uncountable, it could it is just going to be large matrix where the number of rows and number of columns are going to be uncountable. And now if I take a row of, so all of you are able to imagine how this transition probability matrix means, looks like or what, or how does it behave?

Suppose let us say, I take one row, so this is my matrix, right? Okay, maybe, the better way to write it is like this. This is a matrix containing the elements P i j. So now, let me take i 1 so let me take one row, ith row and add all the elements in that. Why is this going to sum to?

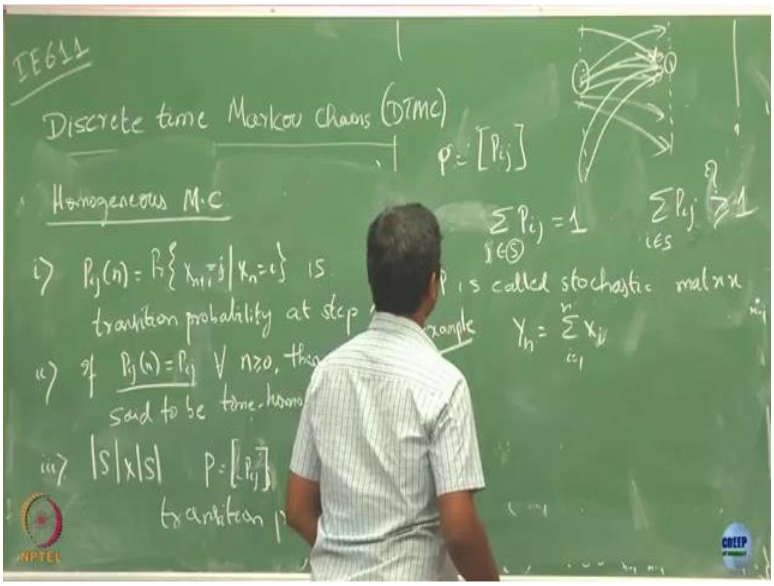
Student: 1.

Professor: 1. And of course, all these elements are going to be non-negative, P i j’s are going to be negative because they are all probability matrix. So such a matrix P is called stochastic matrix. Is that true? So if I take this and now sum over i for a given j, will it sum to 1? Yeah?

This should be j belongs to S, right. So this is I am adding across the columns. No, rows is fixed, ith row is fixed and now adding all the elements in that row, that means I am adding across the column. So what is this? I am asking basically the question i mean state i. Now I am going to state j. And now I consider all possible j’s from a i, I am asking, I am jumping to any other state, what is the probability of that? Once. If you are jumping to any other state, it is going to be. Now here what is this probability saying? What is the probability of jumping to j from different states?

So here final destination is j. And now you are asking what is the probability I reach j from different i's. Is it 1? So what is this? Like say you have state i here and this is j here. And this j is all possible values when I say j is equals to S, right. Now you saying okay, what is the probability that from this, so from this I can I am going to jump to one of this states. That is why this probability 1. Now I am asking the other way question, I want to come to here but from different different values. Now, I am asking okay, what is this probability? So what is the probability that I reached to j from one of the states? But at least is this going to be 1? No. Can this be greater than 1?

(Refer Slide Time: 11:13)



Okay check this, okay definitely this is not 1, can this be greater than 1? It can be 0 also. Sorry, it can be, fine, 0 is fine, 0 is trivially true. Question is, is it can be greater than 1?

Student: One of the state, for example, it is 0.6, so the other state is 0.5. We add this (())(11:39).

Professor: So, let us say from one particular state, it has very high probability of coming to there and it could be greater than half and also it is very likely from another state also it could be coming to that. So these probabilities could add up to 1. So, this if you look at a row, this forms a probability vector.

If you take a one row in your stochastic matrix, that forms a (stoch) that is a probability vector, but if you took at a one particular column, that need not be a probability vector. So because of that it could add up to be greater than 1. Okay, now….

Student: It can be between 0 and 1.

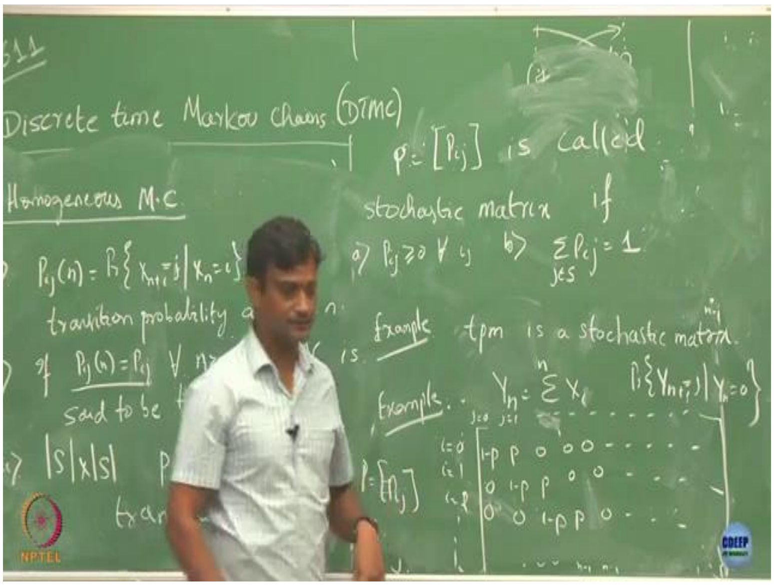
Professor: Yeah, it can be 0 and 1 also. And I am not saying it is all, it is going to be at least 0 and 1 is trivial because each element here is between 0 and 1. The question is, will it add up to greater than 1? That can be possible. Okay, now let us try to understand how this transition probability looks like, for the example we saw earlier. So Yeah. All?

Student: Transition probability (())(13:12).

Professor: Yes. Is it true? All my transition probability matrix are going to be my stochastic matrices, why? No, this is about already like, okay, like a time homogeneous Markov chain. For this, we have this transition probability matrix.

Is this transition probability matrix a stochastic matrix? Yes, right. Because if you take its row, it will add up to one. That is the only condition I need and the other condition that these P i j’s are….

(Refer Slide Time: 13:58)



So, okay, let me formally define, these called, first thing is, and b, summation P i j for j belongs to S is going to be 1. Then transition probability matrix is a stochastic matrix. Yeah, can this be other way around also? Yeah, I mean, if you take a stochastic matrix, that will correspond to translation probability matrix for some Markov chain. Okay, other example, now let us try to understand how this transition probability, yeah?

Student: (())(15:12)?

Professor: It is just like stochastic matrix is a general concept where this has to happen. And now if you have a Markov chain, and it has an associated transition probability….

Student: (())(15:24) need not relate to a Markov Chain?

Professor: No. It was general matrix where the rows add up to one. So basically, what we are trying to show is the property of the transition probability matrix is it is a stochastic matrix. Okay. Let us revisit our example which we did earlier. How does this transition probability matrix look like? So we said that Y n takes value, what? It takes value 0, 1, 2 all the way up to infinity, right? If any sufficiently large.

So now let us try to construct its transition probability matrix. So it is going to be cardinality of S into cardinality of S. So they will going to be uncountable in many rows. But let us try to understand how its one row look like. So let us take the first row which corresponds to i equals to 0. So what does i equals to 0 means? So I am now basically, I want to ask you have observed 0 value and now Y n plus 1.

So this is basically now I am basically trying to say Y i n plus 1 equals to some j given Y n equals to 0. Now, this i is 0 now, now this let us varies this j. If I vary this j, I should be able to get a row for that, right. So, if j equals to 0, what is this value? So, see this is i, this is going to be i 1, i 2 all the way like this and this is like j 0, j 1 and all the way like this.

This is how the stochastic matrix is going to look. So this is because this is a matrix which is P i j for all possible values of i and j. So, i is index in the row and j is index in the column. So, let us say what is this 00 corresponds to, what is this probability?

Student: 1 minus P.

Professor: Why is that?

Student: Because the (())(18:16).

Professor: Right, to remain 0. And what is this?

Student: P.

Professor: And what are the other elements is going to be? They are going to be 00. Now let us come here, if i is 1, can Y n plus 1 be 0?

Student: No.

Student: No.

Professor: So, what is the probability then? And then if i equals to 1, what is the probability that j is 1 again?

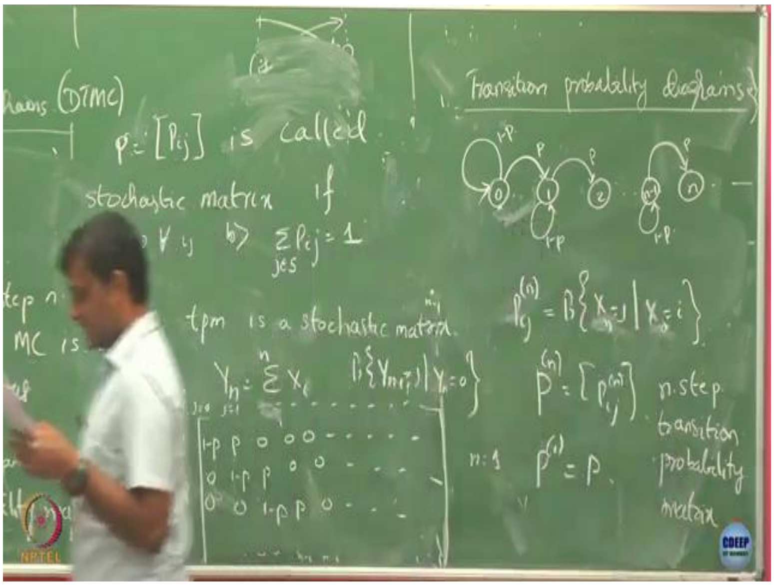
Student: 1 minus P.

Professor: 1 minus P and this is going to be P. And now if I take i equals to 2, what is this value going to be?

Student: 0, 0, 1, 1.

Professor: So, how this matrix look like? Each row will be such that it will be the shifted version of the previous row shifted by one element, shifted towards right. So, this row is shifted here and you feel one like the you keep on shifting and you can fill the entire matrix like this. Okay, now often this kind instead of writing this transition probability matrix like this, one can try to show them pictorially called transition diagrams.

(Refer Slide Time: 19:48)



Transition probability diagrams so, what we will do initially in this is we are going to circle the states like this, let us say this is state n like this. Now, if you are in state 0, I want to see what is the probability that you come back to state 0. So that is 0 going to 0, what is this value?

Student: 1 minus P.

Professor: And now what is the probability that you go from 0 to 1? P and all other things you do not put because they correspond to 0 value. And now, what is the probability that 1 remains 1? What is that it goes to 2? P and what is this probability? 0. So, you can write it say 0 or if it is not if it is 0, you can ignore that link. And suppose if I take this something, what is this n going to n? What is this? So like that, you can show this.

So, you can see that this transition probability diagram has the same information as this transition probability matrix here. Okay, next, now I am going to define one more term here. Let us say I am going to define P i j n. So notice that instead of writing this as a, n as a within this parenthesis, I am writing it as a superscript here. Now, this is a different meaning, what I mean by this is probability that x 1 equals to j given x naught equals i.

So, what is the difference between this notation and this meaning, this notation here? Here you wanted to go from state i to state j in one step in one jump at state n step n, but here you are not jumping from i to j one step, but you want to reach there in n step, starting from origin. So, initially at times 0 you are at state i and now you are asking in the nth step, I am going to reach state j.

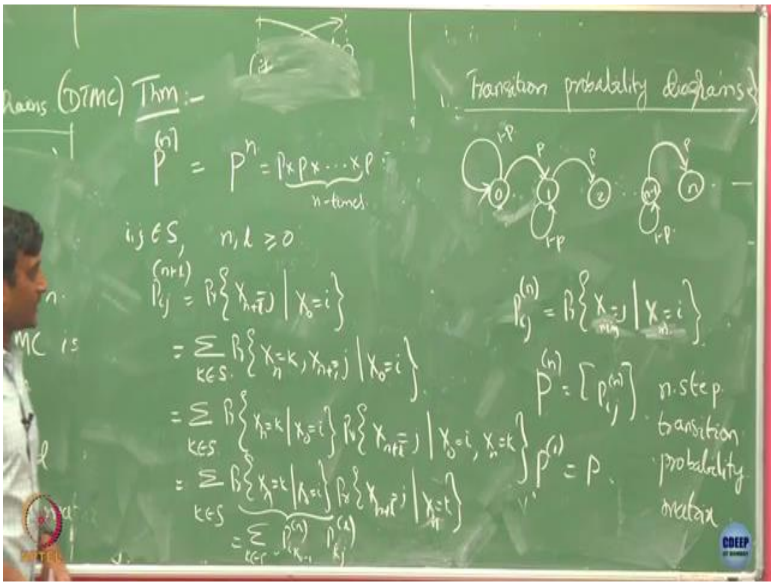
Now, X n equals to this and X. So, you can verify that under homogeneous Markov property, you can indeed write like this for any n. So does not matter from which state you are jumping to which state, what matters is what is that length of steps you are looking at. Yeah.

Student: It becomes stationary.

Professor: We are not talking about that. So, let us not confuse, let us for time being let us take this itself. What I am just saying that you are initially are at state 0, sorry, at time 0, you are at state i and now asking after n steps in the future, you jump to state j. Let us denote this by P i j and then I am going to denote this big matrix B containing all this P i j n.

Now let us try to understand what is the relationship between this P matrix and this P superscript n. So let us take this for n equals to 1, I am saying P equals to 1 is going to be this is P, is this true? For n equals 1, so n equals to 1. So I am just asking for one step jump here and that is going to be. In that case it will fall back to this definition of P. So this, this matrix here is called n step. Now what we are going to show is fine, this is true P of n equals to P.

(Refer Slide Time: 25:16)



Now, we are going to show that P of n, this value is nothing but actually P to the power n, what is the meaning of this? P superscript n in parenthesis is this n step transition probability. P to the power n means, this P multiplied by P times. So, n step transition probability matrix can be expressed in terms of transition probability matrix completely. Okay, let us see why this is true.

So this is we are going to call it as, so let us take some states i j in S and some numbers. Now what I am going to do is, I will be interested in finding x j n plus l. So n plus is some number. So what is the definition of these probabilities? This probability is the probability that i go to X n plus l equals to j starting from X naught equals to i. So this is the definition of P i j n plus l. I am just replacing n by n plus l.

Now I am going to use the usual trick of summing it over all intermediary states X n X n plus l equals to j given x naught equals i. So I did it, I just bought in this extra state here and summed over all possible state it can take. Then X n plus l, so see that I just brought in the state X n and then applied my chain rule of probability here to write in this product form. And then I am going to use my Markov property and the second term here.

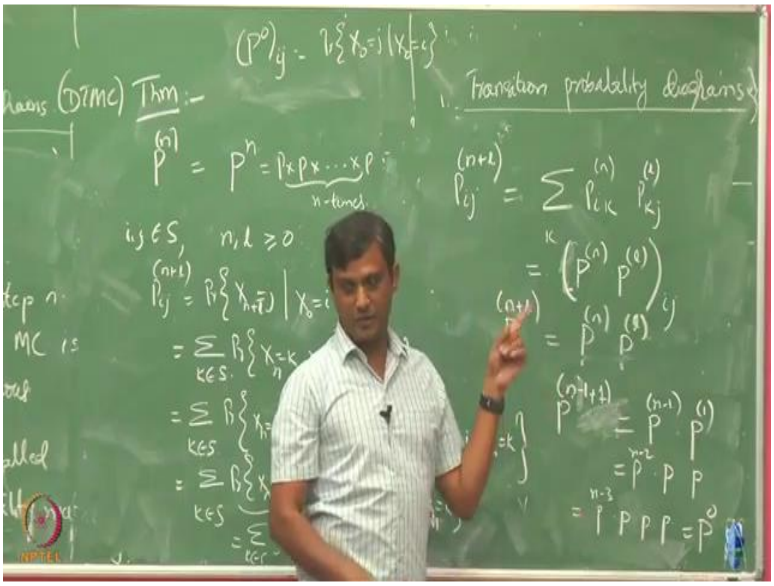
So if I am going to use Markov property, does this probability depends on X naught equals to i? No, right, it only because I am already conditioning X n equals to k, it will be independent of this guy. So it will be, okay. Now focus on this part. So now, this part here, by my definition, this is starting from state i in the 0th round and jumping to state k in the nth round, right?

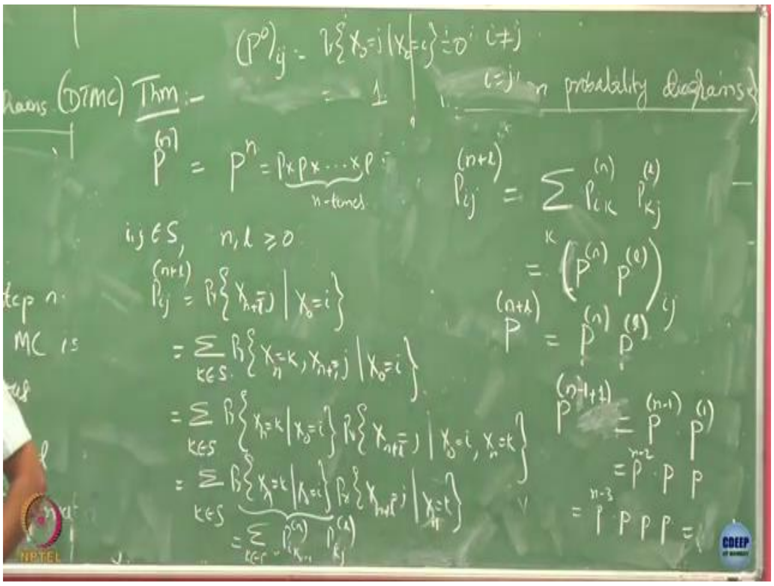
This is by definition P i k in n steps. And then here I am asking, starting from n step. Here we are in state k, I am going to state j in the n plus 1 step that means what how many steps I am jumping here?

Student: l.

Professor: l steps. So you can also argue that this guy is nothing but P I, sorry P k j l so this is so if this n equals to 0, suppose let us fix this n equals to 0. So, this is going from k to j in l steps. But now, we are stating starting from nth round and then jumping to the n plus nth row. So the jump is of l rounds here. So this can be also written as P k j l. So what we are saying is by with this in this definition, if I add like this n plus m and make this sorry, m and n plus m, this can be still with a say, what matters is the number of jumps. So this had to bit argue using our time homogeneity property that this is also true.

(Refer Slide Time: 31:10)





Now what we are doing, we could write P i j of n plus l as summation over k P i k in n steps and then going from k to j in l steps. So, if you are going to think of this as a matrix product, I can write it as P of n and P of l in this of this i j th element in this. Is this true? So, if you take this matrix P n corresponding to the nth transition probability and take this matrix P of l, which is lth transition probability, if you are going to take the product, and then look at the i j th element, this is exactly going to be this.

So, i is that, can you imagine? So, when I am looking at i j th term, what I am basically doing is I am multiplying the ith row with the jth column that is exactly this, right? I am just summing it all possible states that is why this value. So then if every element i j th element of this n plus l can be written like this, then it must be what we are basically saying n plus l is nothing but P of n and P of l.

So, what this relation basically said is, if you take the i j th element of this, it is going to be the i j th element of this product. So, that is why these two matrices are same. So good, what we did is we are able to split this transition probability matrix in this fashion. Now, that will lead us to this result. How? By doing the iterations.

So, I am going to now say this l was here arbitrary, right, I could choose any l I like. So, what I will do is, I am going to take it as I am going to take this n minus 1 and 1 and then I am going to write it as I am going to treat this n as n minus 1 and l as 1 in this case. In that case, I can treat it as P of n minus 1 and P of 1. I am just whatever the property we got here, I am applying that property to this.

So just by say, taking l equals to 1 and n equals to treating n as n minus 1. Now, by our definition, what was P 1? It was P, can I do the same business that I did here on P n minus 1, and write it as like this and I can keep on doing this. Eventually, what I am going to get? I am going to get P raised to the power n actually P 0 and P raised to the power n. So what is P 0? You start from a state and you remain in the same round you remain in that state.

So let me define what is P 0. So what is P 0, P 0, if you take going to take the i j th element, this is going to be probability that x 0 is going to be j given x 0 equals to i, that is the meaning of P 0.

Student: (())(35:55)

Professor: I mean by this kind of iteration.

Student: (())(36:07).

Professor: Yeah, you do that iteration again.

Student: So P 1 is (())(36:12)

Professor: Finally, yeah, so in that case when you end up with P1, you do not want to make it 1 plus 0, right? So in that case, yes, you do not need to do this. You can just end up here, but I still want to define what is this P 0. What is P 0?

Student: Identity matrix.

Professor: You are going to define it identity matrix, why? So when i equals 0 here, this again going to some other state j, is it possible? No, right, this is going to be and if so, this is going to be 0 if i not equals to 0, but if i equals to j, this probability is going to be 1.

So because of this if you are going to now take this P 0, this is simply going to be identity matrix. So both ways are correct, we can just stop at P of 1, or even if you do this, our to make the things consistent our P 0 is going to be defined like this. Okay, so let us stop here. So in the next class, we will look into the other properties of Markov chain called a strong Markov property.