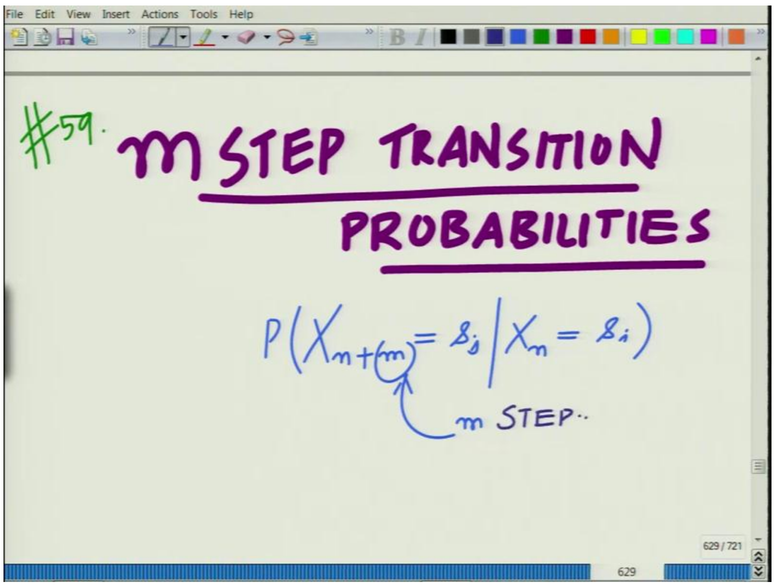
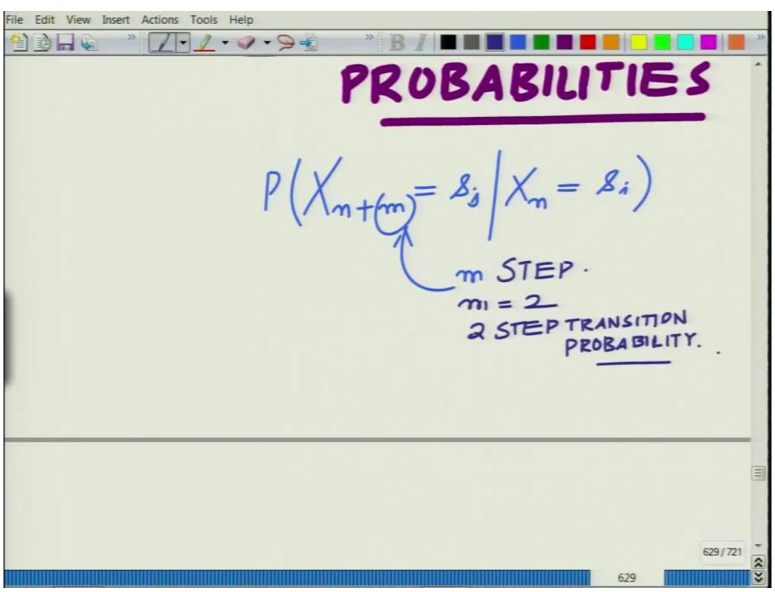
Applied Linear Algebra for Signal Processing, Data Analytics and Machine Learning
Professor Aditya K Jagannatham
Department of Electrical Engineering
Indian Institute of Technology Kanpur
Lecture 59
m-STEP Transition Probabilities for Discrete Time Markov Chains

Hello, welcome to another module in this massive open online course. So, we are looking at discrete time Markov chains and the transition probabilities, how to characterize the discrete time Markov chain and so on. We have seen in particular the one step transition probabilities. Now, let us ask, let us ask the question how to characterize the n step transition probabilities.

(Refer Slide Time: 00:33)

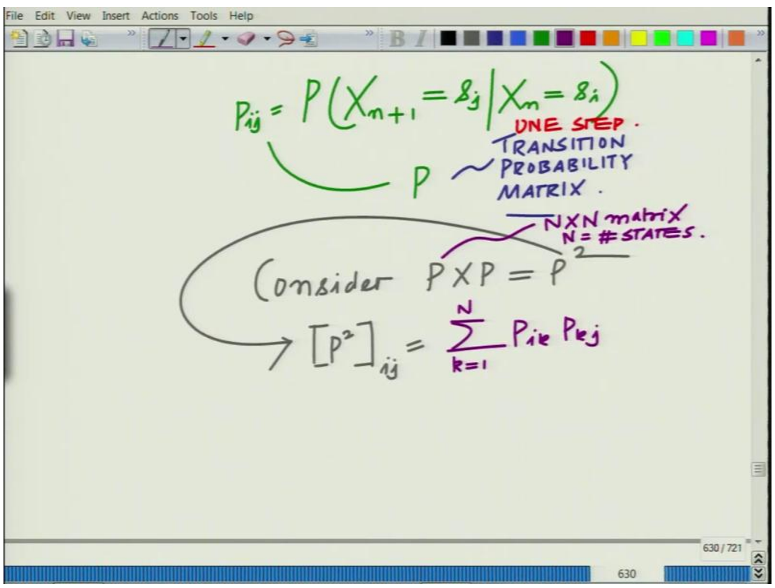


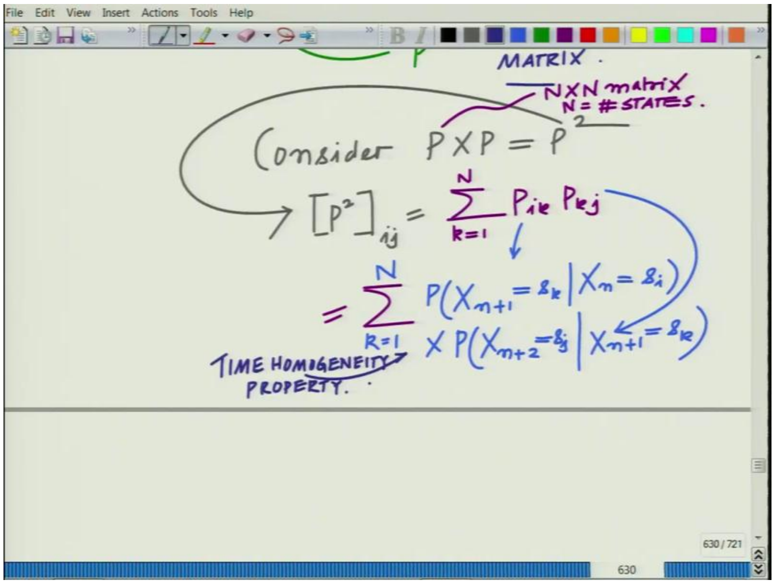


So, we want to look at the n step transition, the n step transition probabilities for this DTMC. That is, we want to ask the question, or let us put it this way, let us say the m step transition probabilities, m step transition probability. So, we want to ask the question, what probability Xn plus m equal to sj given Xn equal to si, that is we are taking m steps. Previously, we were setting one step.

Now we want to look at what is the m step transition probability to begin with, we can set m equal to 2. So, we can ask the question, what is the two step transition, what is this 2 step transition probability? So, remember, we have the one step transition probability matrix, that is we have the matrix probability Xn plus 1 equal to sj given Xn equal to si.

(Refer Slide Time: 02:15)





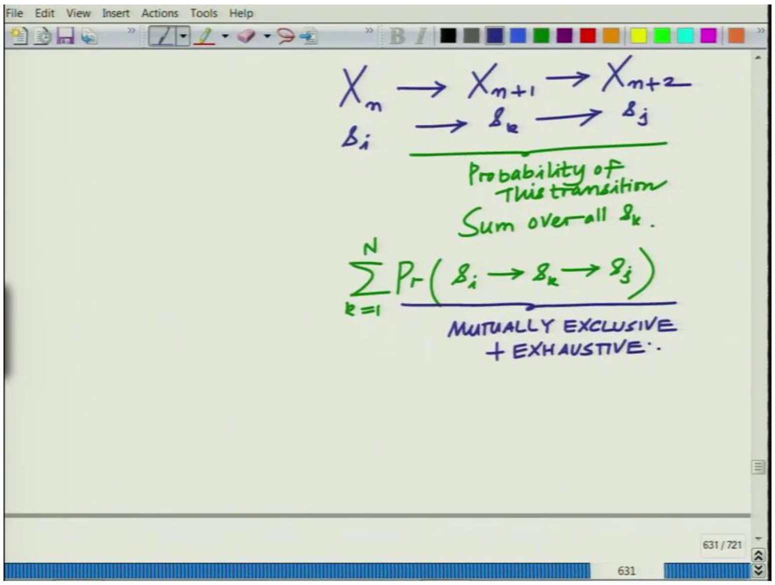
So, we have Pij equals to probability Xn plus 1 equal to sj given Xn equal to si, and we put together in this, in this matrix P, this is essentially our transition probability, the transition probability matrix. And by default this means the one step transition probability matrix that is what we say, we do not mention it explicitly, whenever we say the transition probability matrix, by default it means the one step transition probability matrix.

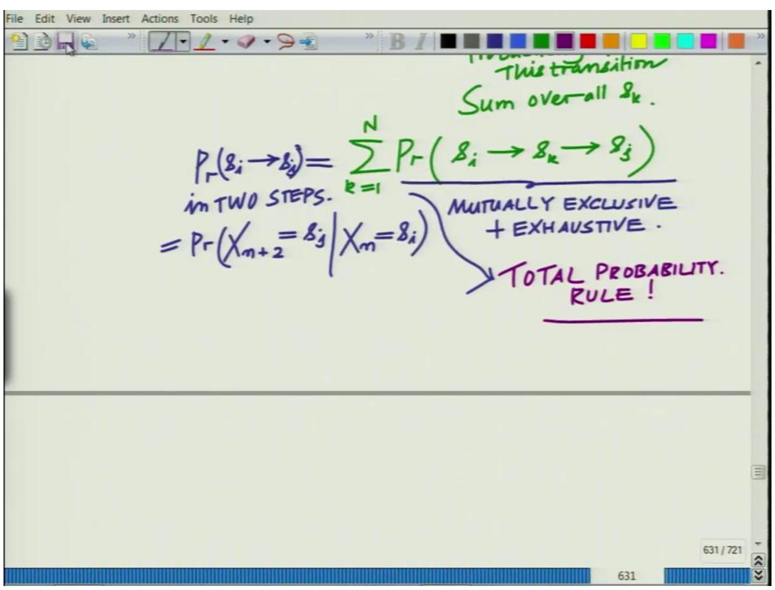
So, I think that is something that you would like to keep in mind, there is not any arbitrary number of steps, but this is the one step transition probability. From which you can derive the n step transition probability matrix as we are going to see. So, consider now, consider P multiplied by P that is equal to P square. Now, if we look at what is the i comma jth element of this matrix, P square and ask that interesting question, what is the i comma jth element of P square, remember this is each P is an n cross n matrix, n equals the number of states.

Now, what is the i jth element of P square? So, this is essentially you sum over k equal to 1 to N Pik times Pkj, Pik times Pkj summation k equal to 1 to n, this is the standard expression for the matrix product. But now, if you look at this, this is interestingly because using the time homogeneity property what happens to this is, this you can write it as the product of these two one step transition probability is P.

You can write it as P Xn plus 1 equals k or sk given Xn equal to si that is this times the probability. Now you can write it as Xn Pkj, you can write it as Xn plus 2 equal to, Xn plus 2 equal to j given Xn plus 1 or Xn plus 2 equal to sj given Xn plus 1 equal to s, Xn plus 1 equal to sk because, remember that we are using here that, the time homogeneity property that is Xn to Xn plus 1, Xn plus 1 to Xn plus 2, these probabilities are the same, the transition probabilities are the same. So, that is essentially the time homogeneity property. So, this follows from the, so what we are using here is the time homogeneity property. We are following the time homogeneity property.

(Refer Slide Time: 06:21)





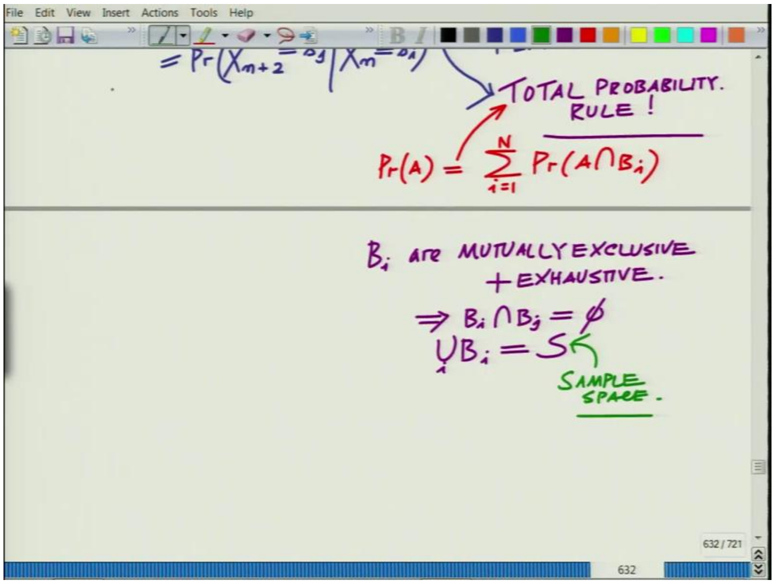
And so, therefore, now, this can be written as and now, you can see this is essentially if you look at this, if you look at this, this describes going from Xn to the state Xn plus 1 to the state Xn plus 2. So, we have Xn in state si Xn plus 2 in state sj and Xn plus 1. So, through sk 1 is going to sj and we are computing this, computing this probability that is, probability of this transition and summing over all, sum over all sk.

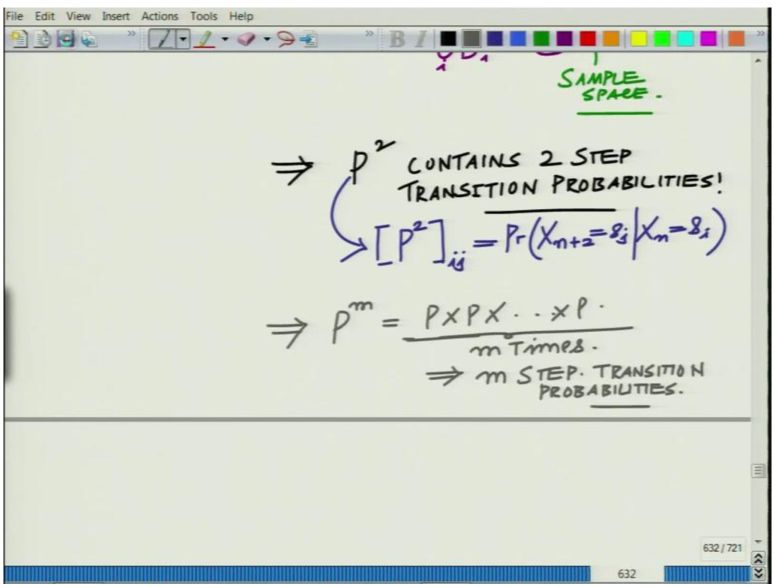
So, essentially what we are doing is, we are doing the probability if you look at it a little philosophically, we are doing probability si to sk to sj that is if you have to go from si to sj in two steps, you have to pass through some intermediate step sk. And now you are taking the sum of these probabilities over all k equal to 1 to n and this is nothing but now, each of these is disjoint. These are disjoint events, so these are mutually exclusive and exhaustive.

So, if you look at in terms of probability theory, these are mutually exclusive plus, mutually exclusive in the sense that if you go from si to sj through s1, and if you go through from si to sj to s2, these two events are mutually exclusive, their intersection is 0, and they are exhaustive. Because you have to go from si to sj through at least one intermediate step in between, so they encompass the complete sample space. So, therefore, one can use the total probability rule to now find the probability of going from si to sj.

This is basically, falls out from the total probability rule. So, using now, the total probability rule, this essentially gives you the probability of going from si to sj in two steps, in two steps. Which we are writing as the probability Xn plus 2 equal to sj given Xn equal to si. So, this essentially follows from the total probability rule. If you are not familiar with this, I urge you once again to revise your, this follows from the total probability rule. So, again, please look at your, refresh your concepts, knowledge of the principles of probability.

(Refer Slide Time: 10:01)





So, we have this total probability rule, which states that essentially if you have, what is the total probability rule? I think it is interesting to just mention it briefly, you can write probability of any event A as the summation of probability A intersection Bi, where i equal to 1 to n, where Bi, the only reason, the only property that the Bi must satisfy is that Bi are mutually exclusive and they are exhaustive.

This implies that Bi intersection Bj equal to phi for any ij and the union Bi over all Bi, this is equal to s, where s is basically the sample space. That is basically if you look at any two events, these two, any two sets that is these two events should be mainly two events, Bi Bj these should be mutually exclusive and exhaustive, exhaustive in the sense that union must span the samples. Because then that is exactly what we have over here, if you look at the transition, two step transition from Xi to Xj, it must go through any intermediate stage or step, any intermediate state sk.

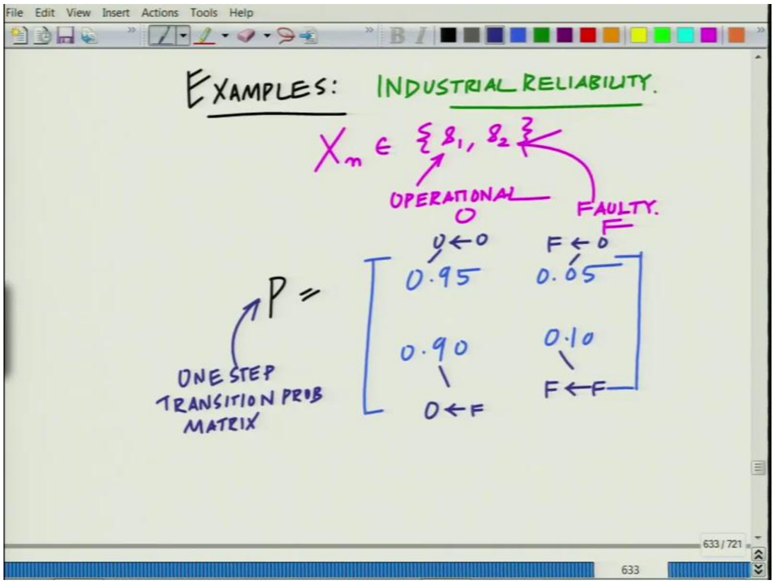
And if you look at all such transitions, these are mutually exclusive, because it, if it is going through any particular state s1 or if it going to any particular state sk it means that it is not going through another particular state sk prime. So, these are mutually exclusive and further exhaustive, because it must go through at least one of these states s1 s2 up to s2 sn in between, in time instant n plus 1.

So, it is very straightforward to reduce this thing and naturally therefore, you can see that the ijth, and the conclusion of this entire, this simple derivation is that the ijth element of P square is the two step transition probability of going from Xn in state si to Xn plus 1 Xn plus 2 it being in state sj, that is the idea. So, so, idea is P square contains the 2 step transition probabilities, a very convenient way 2 step, very convenient way, you simply, simply multiply P with itself and you can do it because P is a square matrix, contains the 2 step transition probability.

And similarly, that is what we mean by that is, if you look at P square of ij this will be equal to the probability as we already said Xn plus 2 equals sj given Xn equal to si. And similarly, extending this, you have Pm that is P times P times P performed m times this gives the m step transition. This gives the m step transition probabilities. So, I think that must be fairly clear from this derivation, in fact something that is worth remembering.

And it is very convenient, once you have that one step transition probability matrix B. Simply take the product m times that is P raise to the power of m and that gives you the m step transition probabilities. Let us look at a simple examples or a couple of simple examples to understand this. Let us go back to our industrial reliability and our stock price examples.

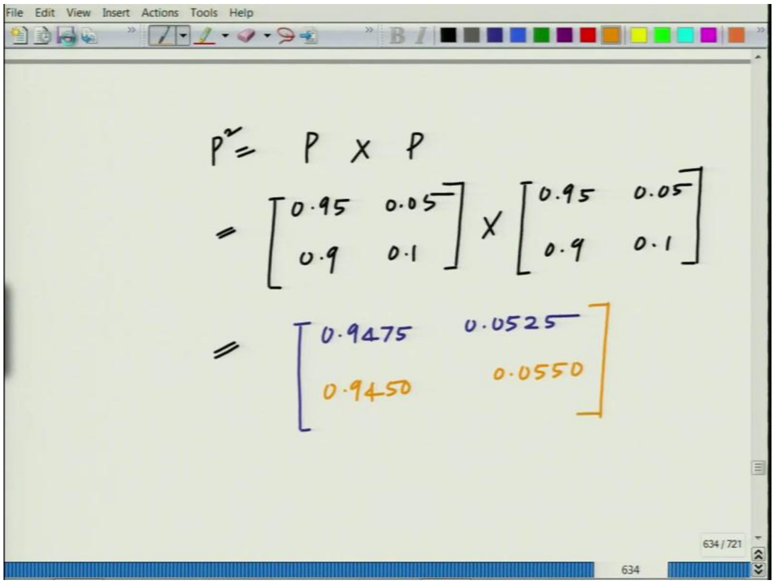
(Refer Slide Time: 14:33)

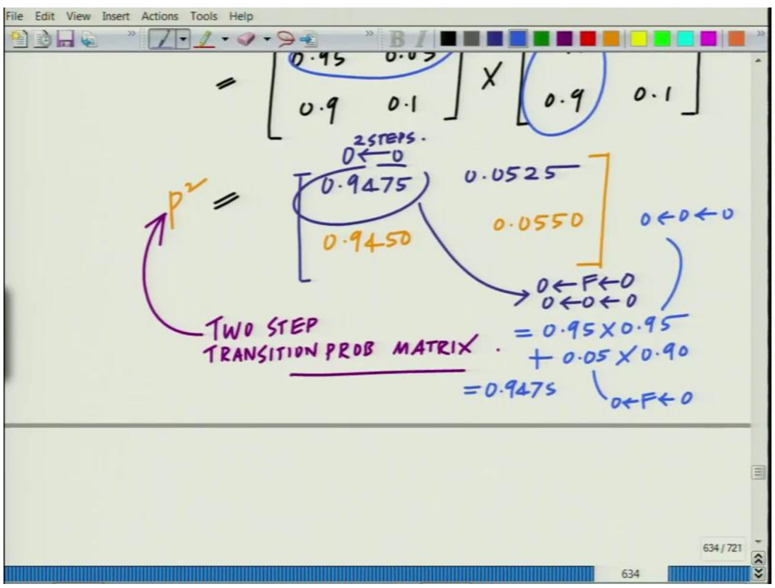


So, let us look at simple examples to understand this. So, the first one is our industrial reliability example. So, we go back to our industrial reliability problem. And we have here Xn belongs to state 1 comma state 2 remember this is your operational, operational or this is the O and state 2, this is your faulty or your F state and the machine can be in one of these two states can be operational or faulty and it can transition from operational to faulty when it breaks down, faulty to operational, when it is repaired or it can continue from operational to operational, faulty to being faulty.

So, now, if you look at the one step and we have looked at this example, in detail the one step transition probability matrix were this, we have seen to be given as in our example 0.95 0.05 0.90 0.10 and this, naturally this is then the operational to operational. This is your operational to faulty, this is your faulty to operational and this is faulty to faulty. So, this is the one step transition probability matrix. This is your one step transition probability matrix.

(Refer Slide Time: 16:52)





And now, if you compute, now if you ask what is the two step transition probability matrix? Now, we have perform P square that is basically P times P which is take this matrix 0.95 0.05 0.9 0.1 and multiply it by itself, 0.05 0.9 0.1 and therefore, this gives you the two step transition probability matrix, that is 0.9475 0.0525 0.9450 0.0550 and this is essentially your, now you can write that this is essentially your two step transition probability matrix, this is your two step, this is essentially a two step transition probability matrix. And so, this is your two step transition probability matrix.

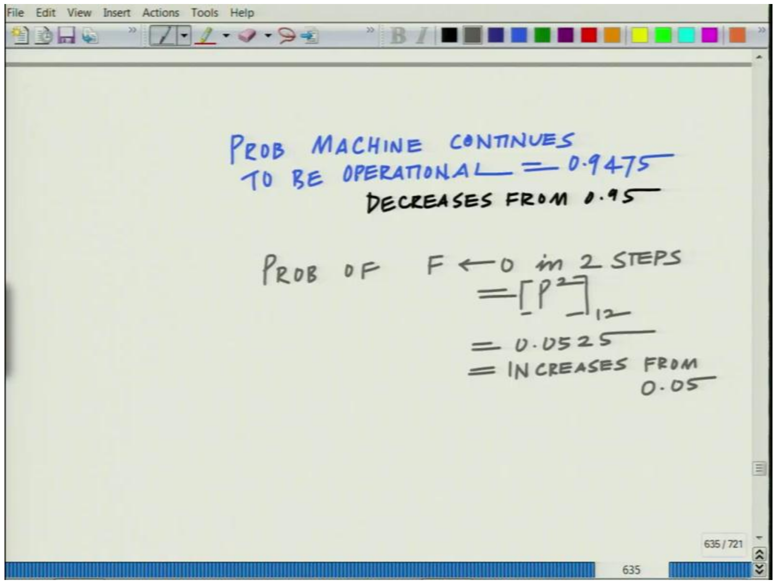
And now, if you look at this, what is this? Let us look at this, this is essentially again it goes from O to O, operational to operational and in two steps and if you see how you get this, you can see to go from operational to operational in two steps either you can go from, you can go two ways you can go to operational to faulty to being operational or you can go from operational to continue to be operational and operational.

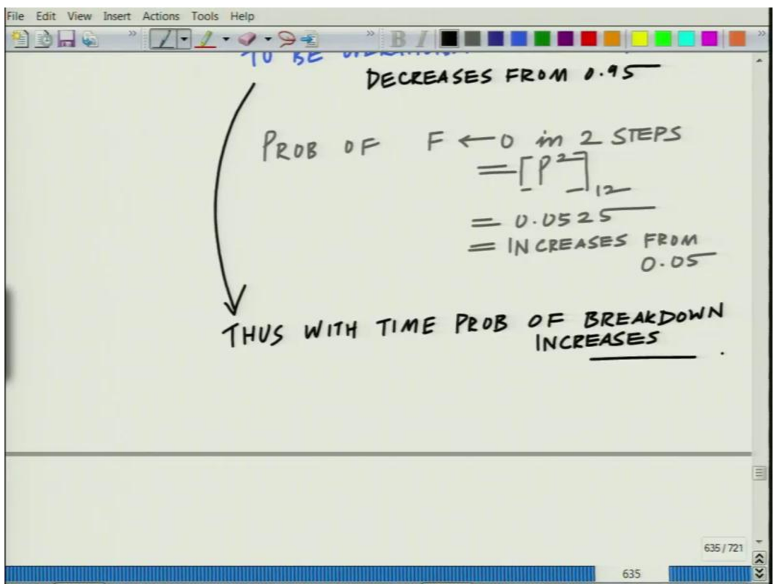
And in fact, if you look at this product that reflects this. So, this is equal to essentially the sum of these two. So, 0.95 into 0.95 which corresponds to operational to operational, that is Xn to Xn plus 1 is operational and Xn plus 1 to Xn plus 2 is also operational. That means, you have the product of the two probabilities, remember these are independent so 0.95 to 0.95 or plus you have, either it is goes from operational to faulty that is your probability 0.05 and subsequently it is repaired.

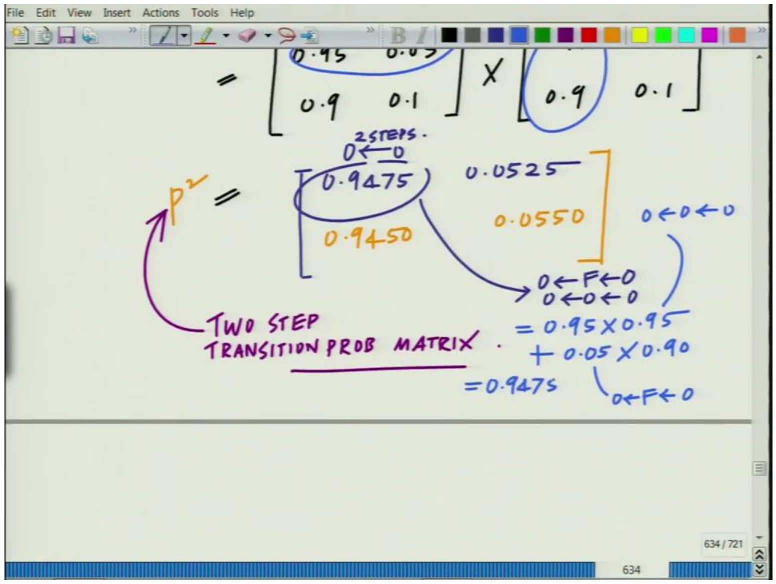
So, it goes from this corresponds to your operational to faulty and subsequently repaired. And that is logical and that you can see is essentially what you get as 0.9475, that is it goes from operational or operational in two steps, remember it does not mean that it continues to be operational, that is the important thing. You can either have been continued I mean, I can either have continued to remain operational or it would have broken down in the first hour and repaired in the second hour.

So, these are the two possibilities and now, we can clearly see the intermediate, the sum over the intermediate steps, there are two intermediate stages. Intermediate stage could have been either s1 that is continuing to be operational or s2 when it is broken down and then repaired. And these are now your two step transition probabilities.

(Refer Slide Time: 20:45)





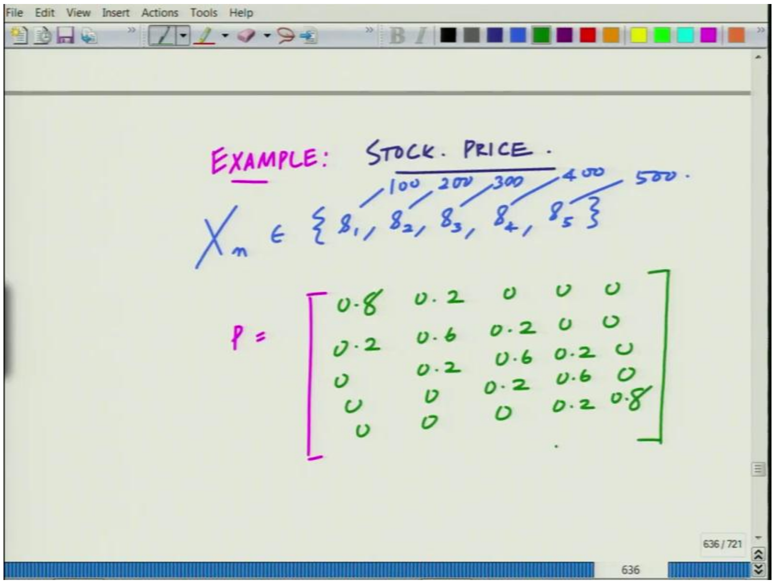


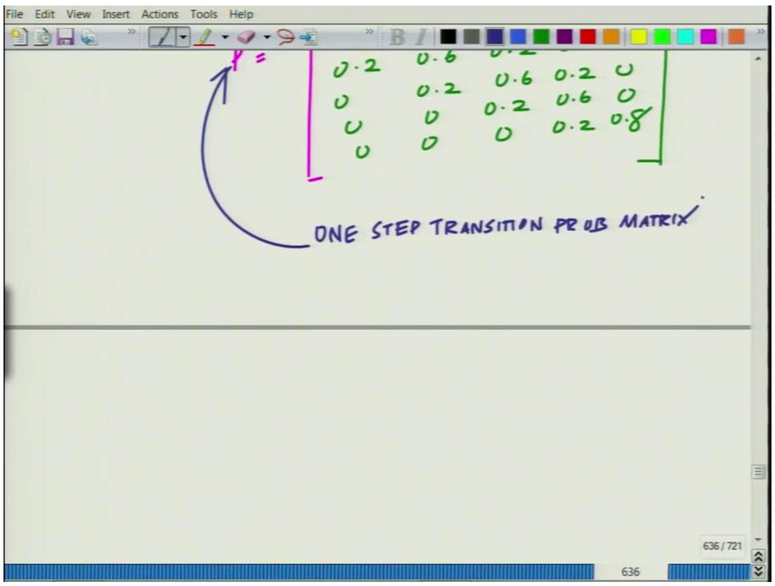
And interestingly you can see that operational operation that it continues to be operational, probability machine continuous to be operational is 0.9475 which decreases from, that is it decreases from 0.95. That is original, remember the one step probability that continues to be operational is 0.95. The probability that it continues to be operational after two hours now, you can see is lower it is 0.9475.

So, essentially it means and look at the probability of breakdown, probability goes from operational to break down in 2 steps, operational to faulty in 2 steps, this is equal to point, you can see this is equal to P, P square of 12, this is essentially your P square of 12. This is 0.0525 which basically increases, increases from the one step value of 0.05. So, originally it is 0.05, the probability it breaks down in two step increases 0.05225. Therefore, now, therefore, as you can see as time increases the probability that it continues to remain operational decreases, probability that it can break down at some point probability goes into breakdown increases.

And that is natural, that is what you expect as time increases as time progresses, probability that breaks down progressively increases. So, the probability that it breaks down and at some point it in fact reaches an equilibrium, that is something that is interesting to hear, that we are, what we are going to see as we go forward. Thus with time, so the summary of this, just of this, thus with time the probability of breakdown increases. So, that is the first example. Let us now look at, go back and look at another example that is our stock price example.

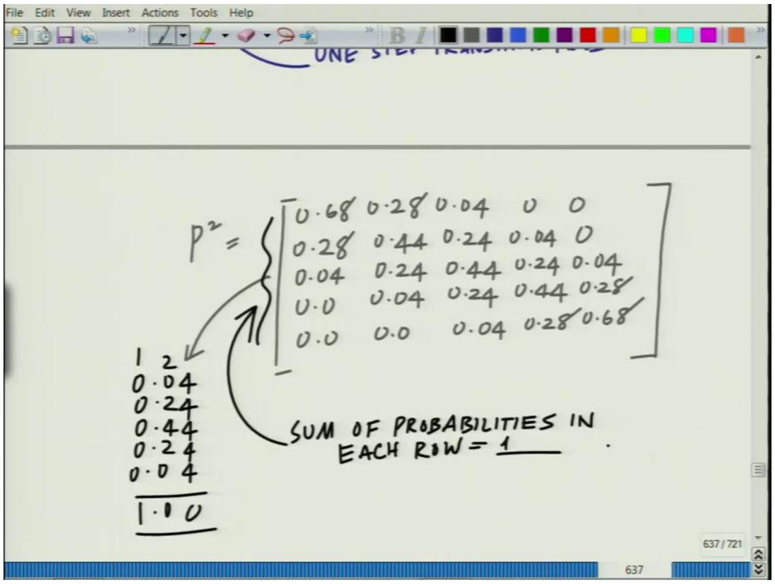
(Refer Slide Time: 24:11)

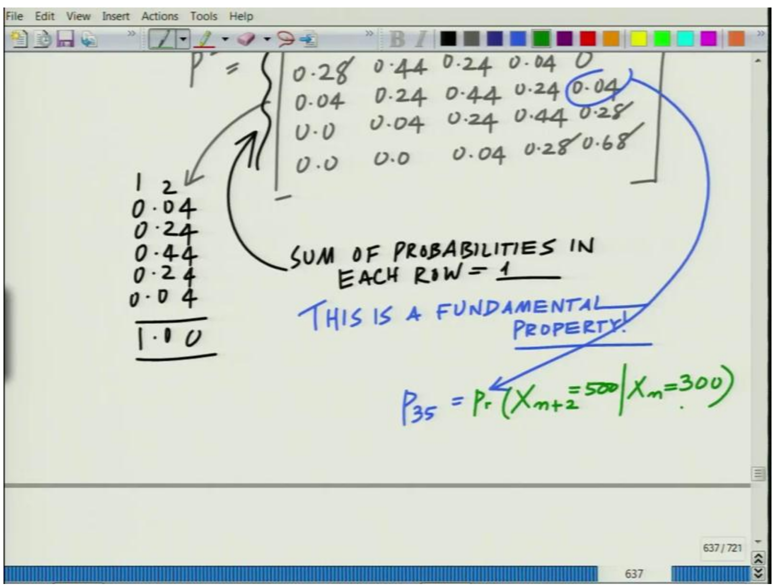


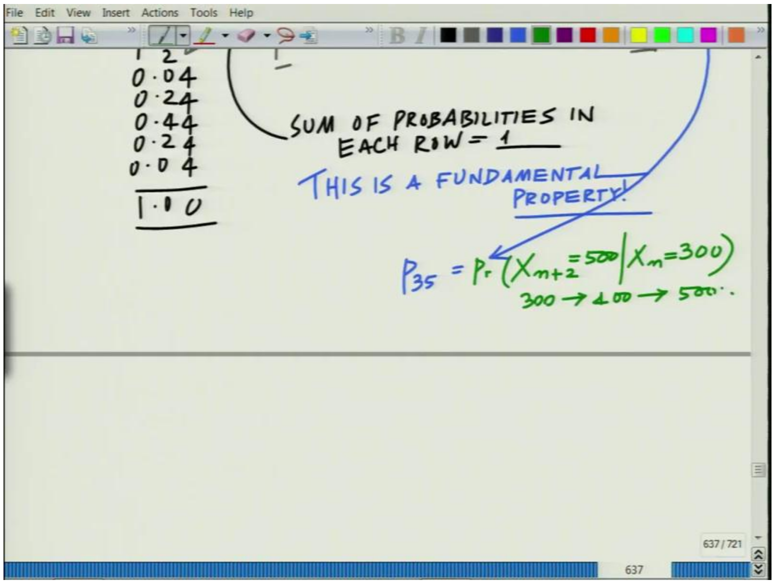


So, let us go back take a look at our another example, that is our stock price. And in this example, remember, the stock price can be in any of five states, this simple example s1, s2, s3, s4, s5 this is 100 200 400 500 and the one step transition probability matrix, remember if you can remember the one step transition probability matrix is essentially, this is your 0.8 0.2 0 0 0 0.2 0.6 0.2 0 0 0 0.2 0.6 0.2 0 0 0 0.2 0.6 0 0 0 0 0.2 and finally, you have the 0.8. So, this is your one step transition probability matrix, this is your one step.

(Refer Slide Time: 26:01)







And the two step transition probability matrix now can be obtained as P square and you can evaluate this thing using a software such as MATLAB, this gives you 0.68 0.28 0.04 0 0. Now, one thing you can still continue to observe, be it either the one step transition probability, two step or m step transition property sum of rows is always equal to 1, that property must always hold.

Because whether it goes to, from ith, state i to sj in one step or it goes from state i to sj in two steps, remember the two step transition probability, three step, for all the transition probability matrices irrespective of the number of steps and the sum of the rows must always be equal to one and that is the fundamental property of the matrix. That follows once again, from the total probability rows, because you are taking the sum of the probabilities over all possible ending states j given in initial state i.

So, that property must always hold. So, I think it is worth something interesting and that is worth remembering, 0.2; 0.28 0.44 0.24 0.04 0 then you have I guess 0.04 0.24 0.44 0.24 0.04 then you will have 0.0 0.04 0.24 0.44 0.28 and then you will have 0.0 0.0 0.04 0.28 0.68. This is the two step transition probability matrix and if you look at this, so for instance, let us take this row and then you have the sum of the probabilities 0.04 0.24 0.44 0.24 0.04 0 that is 2 sum of all the probabilities.

So, you see sum of probabilities in each row equal to, always holds for any transition probability matrix, always holds, this is a fundamental property. This is a fundamental property, which always holds, would serve you well to remember that this is a fundamental property which always holds. And now if you look at for instance, this let us take a simple, let this essentially P 35. That is, we are asking what is the probability Xn plus 2 equals 500, let us state 5 given Xn equal to state 3, that is 300.

The way this can occur is, if it goes from 300 to 400. And from 400 to 500, this is the only possible transition because from 300 you cannot directly go to 500. And no other transition is possible and 300 to 400. Remember, the probability each of these probabilities is 0.2. So, this probability by independence is again 0.2 into 0.2 equals to 0.04. And that is how you get this entry. This is essentially your 0.04.

And similarly, that can explain also all the other entries I mean, you can also finally reduce how you can get essentially from for instance, let us say a, something like, I mean, you can reduce the rest of the probabilities in similar fashion, in this transition probability matrix. So, the summary of this is essentially very simple, that if you have the transition probability matrix, as I have already told you before, and we already seen before that completely characterizes the DTMC discrete time Markov chain.

And now, you also seen an implication of that now, the one step transition probabilities P, 2 step transition probabilities are given by P square, m step transition probabilities are given by P raise to the power of m that is P the matrix P multiplied by itself m times and in particular, if you look at any row of this matrix, any Pm, all the rows have to add up to unity that is a fundamental property. So, after this interesting module, let us take a break here and we will continue with other interesting aspects of DTMC analysis, that is discrete time Markov chain analysis in the next module. Thank you very much.