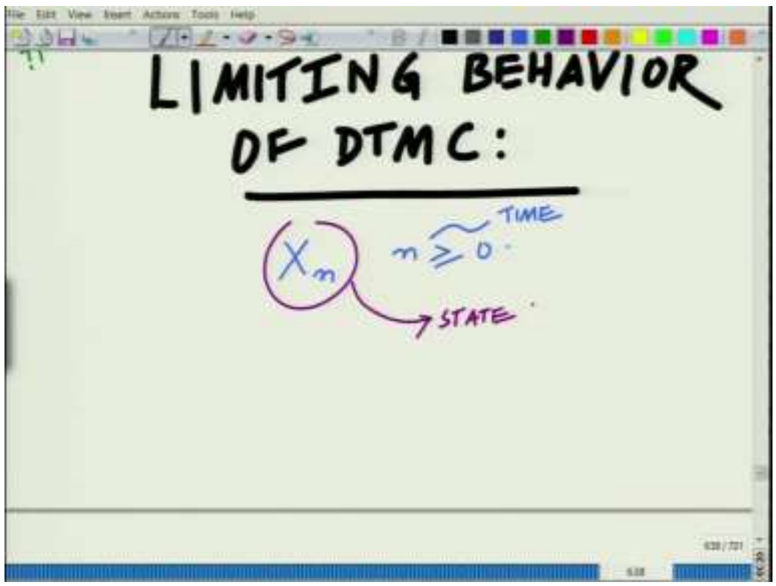
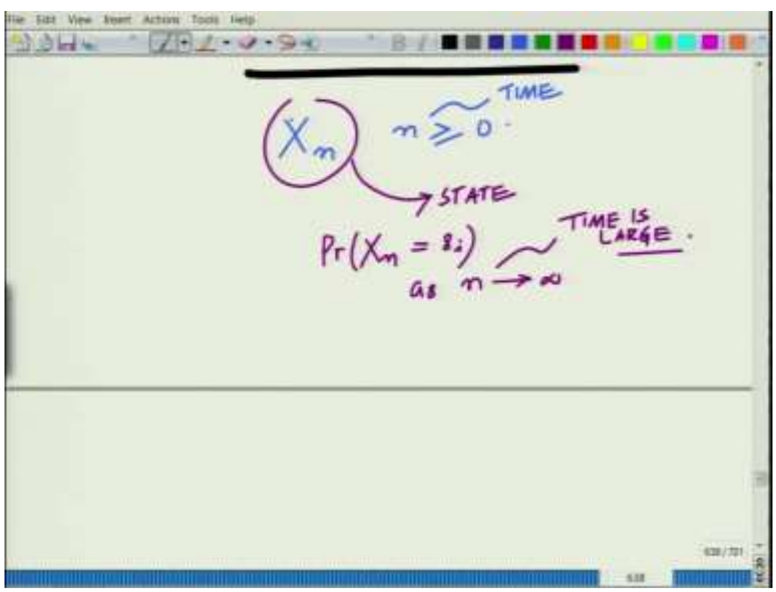
Applied Linear Algebra for Signal Processing, Data Analytics and Machine Learning
Professor Aditya K. Jagannatham
Department of Electrical Engineering
Indian Institute of Kanpur
Lecture 60
Limiting Behavior of Discrete Time Markov Chains

Hello, welcome to another module in this Massive Online Open Course. So, let us continue our discussion on this DTMC Discrete Time Markov Chains. And let us look at, the limiting behavior of a DTMC discrete time markov chains.

(Refer Slide Time: 00:28)

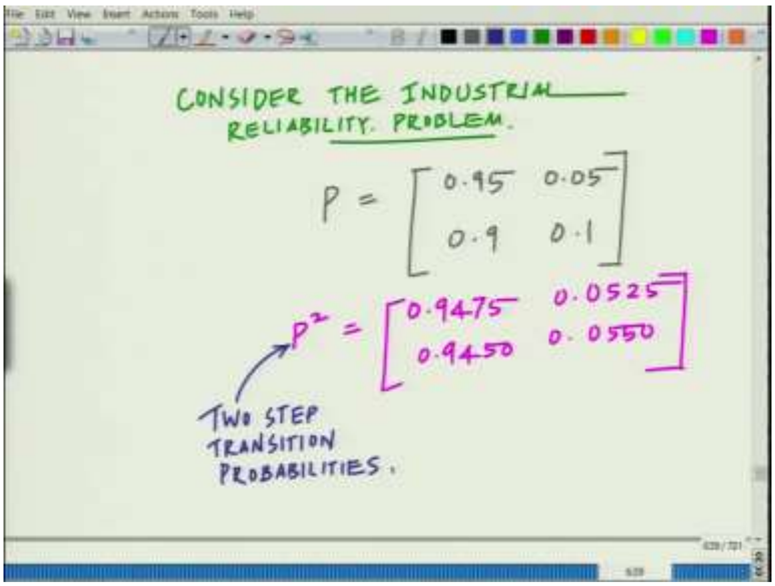




So, we want to look at the limiting behavior of, limiting behavior of DTMC we will start with our. So, what we will do let us look at, X n which is the state of the DTMC for n greater than equal to 0 remember this is our time and this is our, this is basically our X n is our state.

Now, we would like to ask the question what is the probability X n equal to S i as n becomes very large as time that is time is extremely large that is after a long period of time if, you observe this probability is the probability that the state is S i can you comment about that? Can you comment about that? Can you comment about what is the probability that the state will be in, the state will be S i after a long duration of time n.

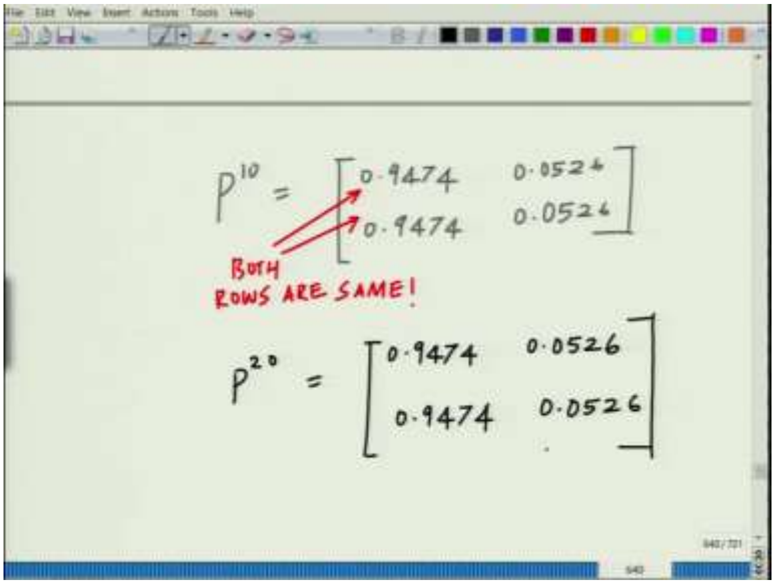
(Refer Slide Time: 02:22)

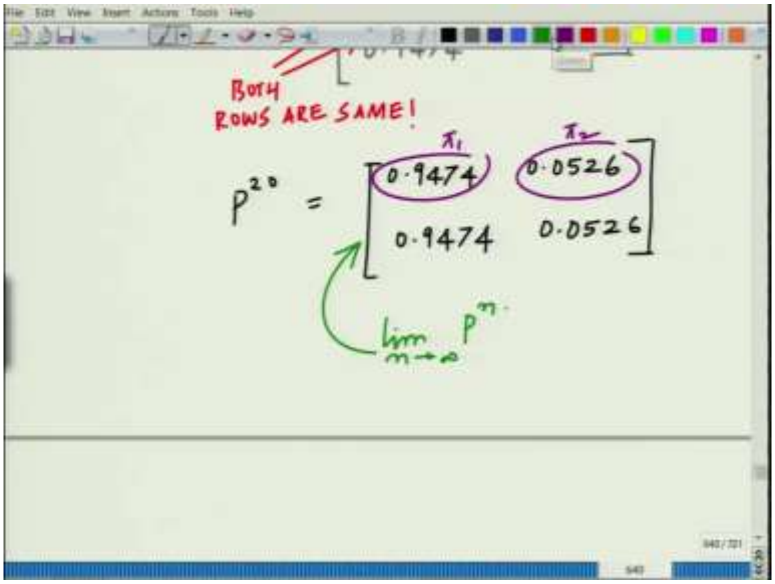


Now let us use this, let us under use an example to understand this better. So, for instance consider our industrial reliability problem where, you have the machine which has two states operational and faulty. Consider now, what we have here is if, you look at the one step or the transition probability matrix this is basically you are .95, .05, .9, 0.1

And now, if you look at P square this is going to be .9475, .9450, .0525, .0550. And we know, that P square gives the two step transition probabilities that is probability X n plus 2 equal to S j given X n, given X n equal to S i that is the element P i j. So, these are the two step transition probabilities. So, I hope all of you also appreciate that. So, these are the, these are the two step transition probabilities.

(Refer Slide Time: 4:43)



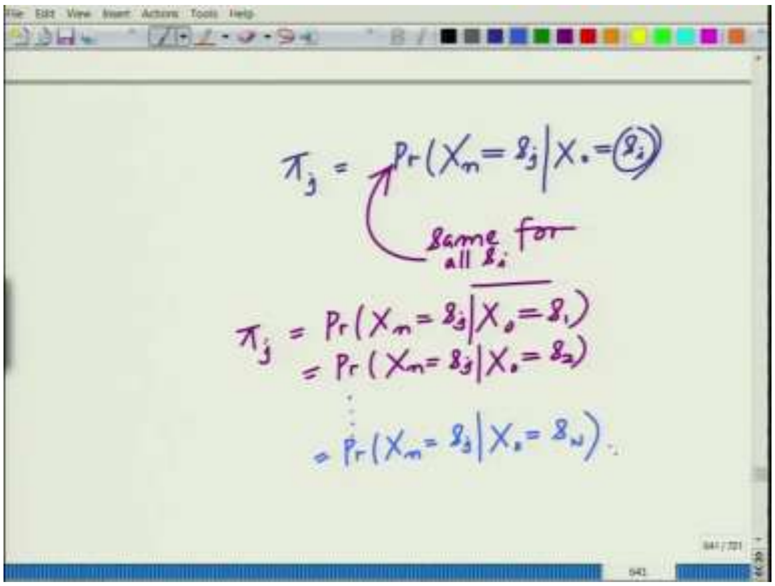


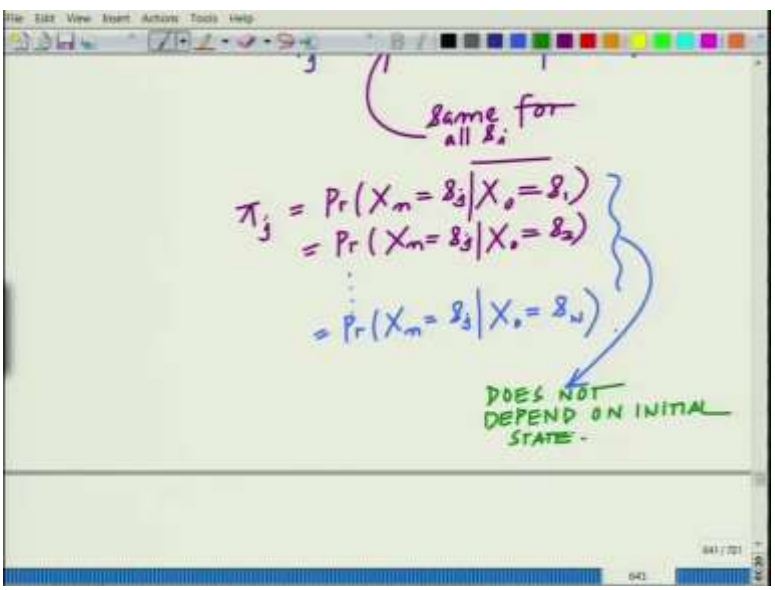
Now let us look at, P to the power 10, this gives the 10 step transition probabilities and you will observe something interesting. So, you will observe that this is .9474, .0526, .9474 and this is also point what you starting to observe is that both the rows, both rows are, both rows are the same that is if, you look at the transition probabilities you will observe something interesting that is in this matrix the transits 10 step transition probability matrix P raise to the power to 10 both the rows are same.

And further interestingly now if, you look at P 20 what you will observe is that, this is again essentially it is the same, the same matrix. of P 20 approximately equal to what you can see is it does not change significantly there might be some minor numerical variations. But, you can see by and large they are the same. And therefore, you can say this is basically the limit n tending to infinity P raise to the power of n that is the n step transition probability.

And if, you call this as the vector pi bar that is your call this as pi 1 and you call this as pi1 what you can see is essentially this is your pi 1, this is your pi 2.

(Refer Slide Time: 07:06)

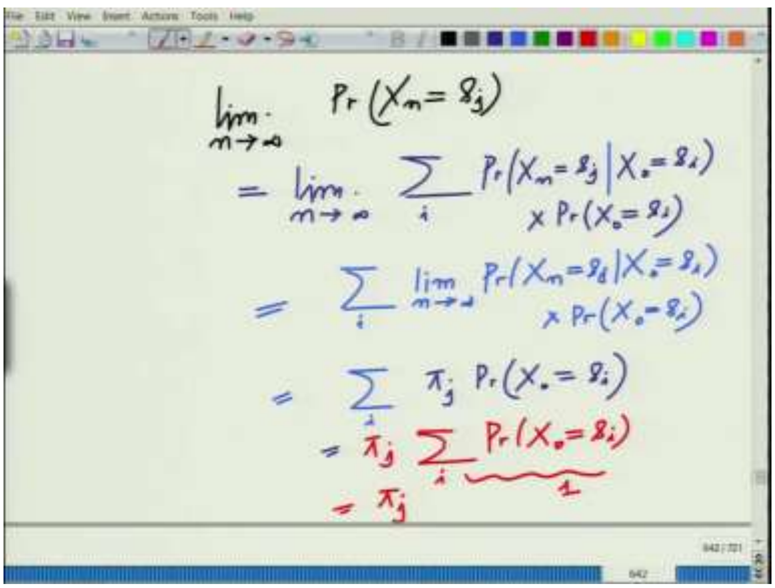




And, what you can see is this pi 1 or this pi j equals probability X n or you can even say this probability X n equals S j given X 0 equals S i that is the starting state can be, the starting state is S i. And you can see this is the same all the rows are same implies this is the same for all S i. So, this probability is same for all S i that is irrespective of the starting state and this is what you would expect because, at time n equal to infinity it does not really matter which state you have started from the probability that is in state S j is same irrespective of the starting state S i and this is what we are denoting by pi j.

So, this is what we are denoting by pi j. So, pi j equals probability X n equals S j given X 0 equals S 1, this is equal to probability X n equal to S j given X 0 equals S 2, so on, this is the probability X n equals S j given X 0 equal to S n. So, essentially this does not depend on, does not depend on the initial state, does not depend on the initial state, which is a very interesting property.

(Refer Slide Time: 9:20)



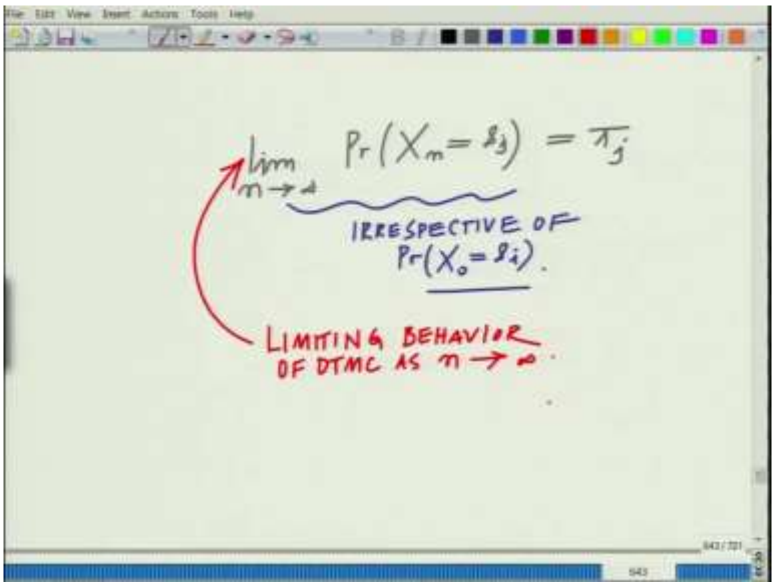
Further now, if you look at the limit n tending to infinity. Now if, you ask the question what is the limit n tending to infinity probability X n equal to S j that is at n equal to infinity what is the probability that this discrete time markov chain will be in state S j.

There is a very simple way to evaluate that is probability limit n tending to infinity summation over all i probability X n equal to S j given X naught equal to S i times probability X naught the initial probability x naught equal to S i. Now, we have seen this quantity is essentially pi j for all j this quantity does not depend on i.

So, this becomes and now if, you take the limit n tending to infinity. Now first let us, take the limit n tending to infinity inside. So, this becomes summation over i limit n tending to infinity probability X n equal to S j given X naught equal to S i times probability X naught equal to S i this becomes our I probably X n equal to S j given X naught equal to S i this becomes this is your pi j remember this is essentially your pi j times probability X naught equal to S i.

And now, you take this pi n outside this becomes summation over i probability X naught equal to S i. Now, this sum of all probabilities is equal to 1. So, this becomes equal to your pi j this probability x naught equal to S i for different i is nothing but, the probability density function or the probability mass function of the initial distribution of the state probabilities. And, what this shows is limit n tending to infinity.

(Refer Slide Time: 11:51)

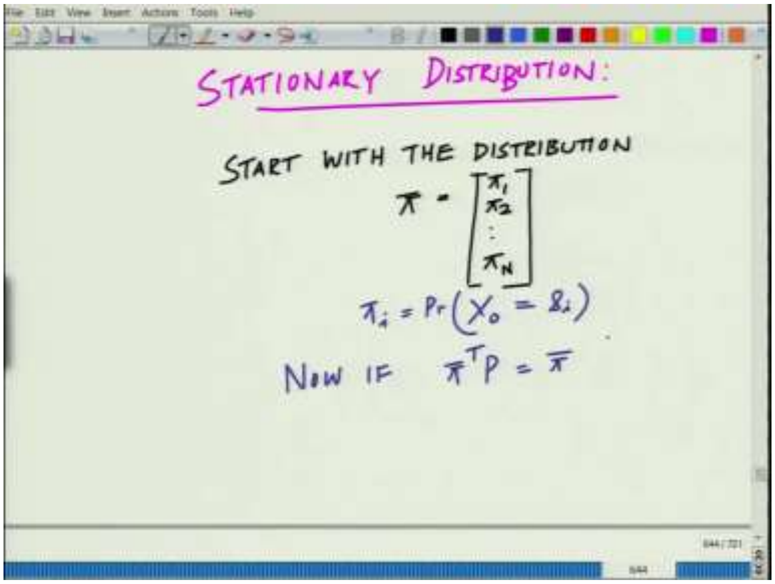


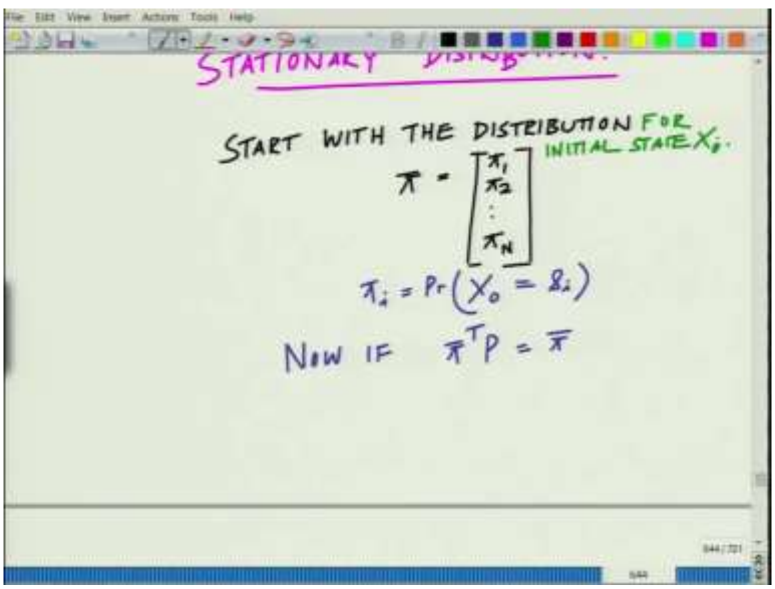
And, this shows a very interesting result that is after a very large time if, you ask the question what is probability X n equal to S j this is equal to pi j and therefore, this is known as the stationary and this is basically therefore this is essentially what this is? So, the probability at n equal to.

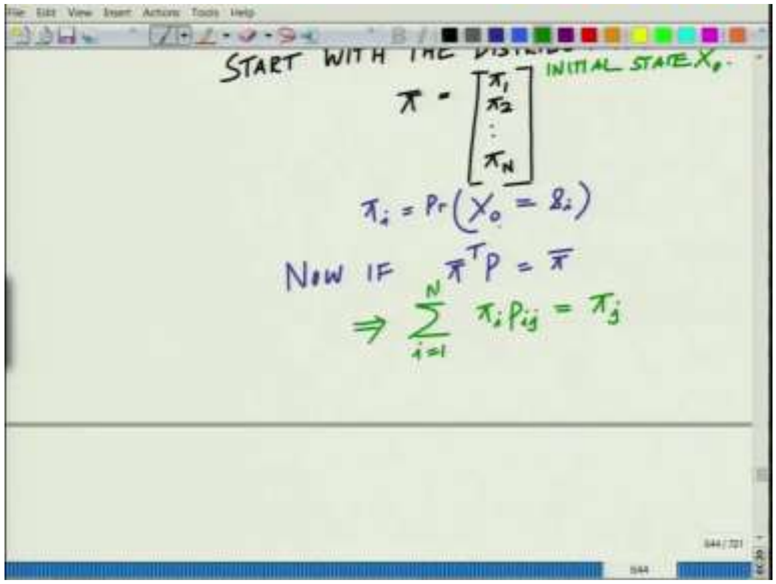
So, the probability at after long duration that is and this is irrespective of the and the interesting thing is this is irrespective of probability X naught that is irrespective of the initial probability distribution or mass function probability that X naught is in state S i after a very large duration probability X n that is limited ending to n tending to infinity the probability that the discrete time markov chain X n discrete time markov chain can be found in state S j equals pi j.

Where, remember pi j is basically the jth element of any row of the n step transition probability matrix as the limit n tends to infinity. So, that is the interesting. So, this is essentially what we call as the limiting behavior, this is the limiting behavior of DTMC as n tends to infinity.

(Refer Slide Time: 13:47)





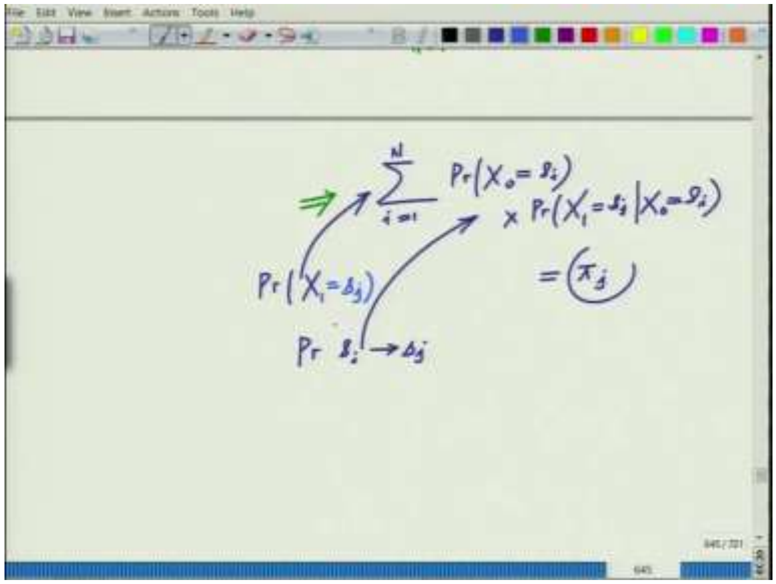


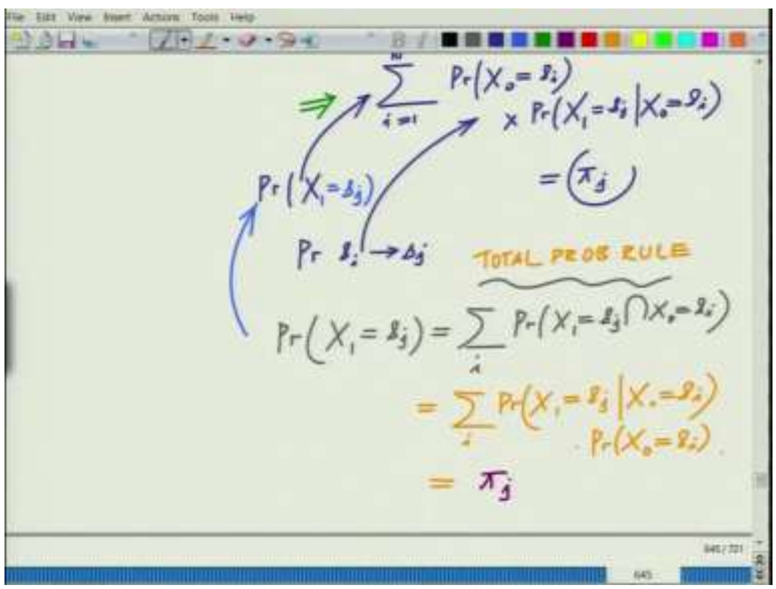
Now, we have the notion of what is known as a stationary distribution for any DTMC we also have what is known as the stationary distribution which is again a very similar concept we have this notion of a stationary, we have this notion of a stationary distribution which is essentially that the what is the stationary distribution?

Now if, you look at this let us start in the distribution. So, let us start in the distribution pi bar equal to pi 1, pi 2 up to pi n that is these are where pi i equals probability that X n equal to S the probability that X naught to S i this is the probability that X naught equal to S i then what happens is. Now, if pi bar Transpose times P equal to pi bar.

Now let us, consider a probability, probability distribution, probability distribution this is the any probability initial, probability mass probability distribution for the initial state. So, the start with the distribution. So, distribution for the initial state. Now if, pi bar Transpose P equals to pi bar this implies that if, you look at any i j th element this implies pi i this implies that pi i or this implies that pi j or this implies that, pi i times pi j equals pi j, pi i times P i j equals pi j where you have i equal to 1 to n.

(Refer Slide Time: 16:51)





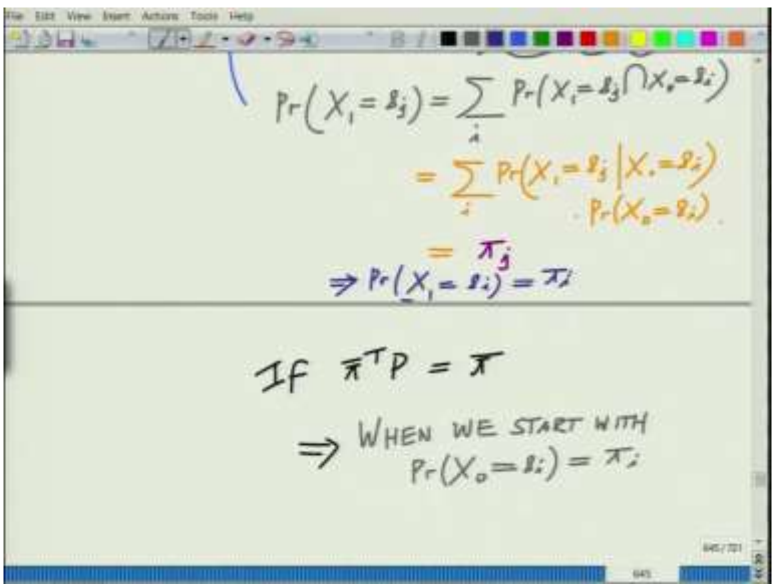
This implies that if you remember summation i equal to 1 to n this implies that probability of the probability X naught equal to S i times probability the transition probability X 1 equal to S j, X 1 equal to S j given X naught equal to S i equals pi j. But, this is nothing but initial probability of x naught equal to S j times the transition probability that it from X naught it transitions to X 1 equal to S j summed over all possible S i. So, this is the probability of transition probability.

So, this is probability of S i to S j sum over and this weighted and summed over the probabilities of all S i and this equal to this equal to P j and this is nothing but, essentially if you look at this quantity this sum that is this summation over the initial state S i multiplied by X 1 equals to S j given x naught equal to S i this is nothing but, the probability that x this is from the total probability rule there is a probability that S i or X 1 equal to S j.

Because, the probability X 1 equal to S j you can write it as follows if, you look at probability X 1 equal to S j probability X 1 equal to S j is the summation you can write it as the summation probability X 1 equal to S j intersection X naught equal to S i over all possible states or all possible i this basically follows from the total probability rule.

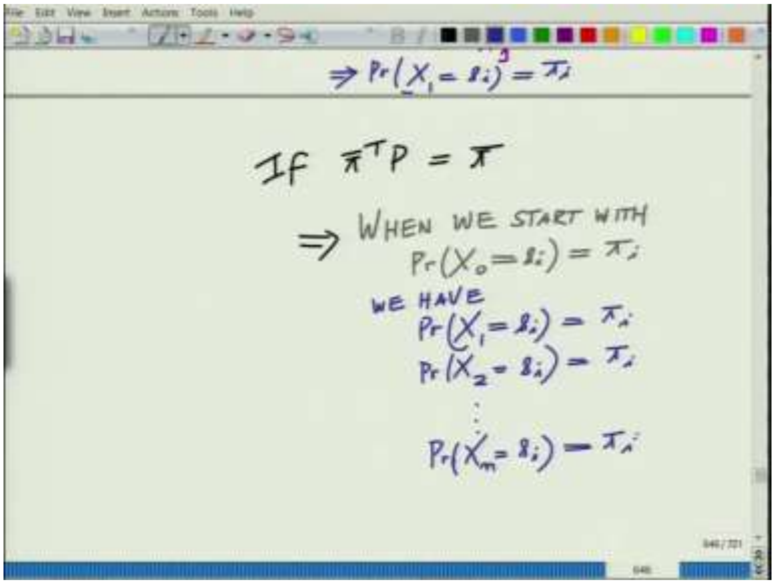
Which is basically equal to nothing but summation over i probability X 1 equal to S j given X naught equal to S i times probability X naught equal to S i. And, that is essentially what we have here and that we can see is equal to. Now, from this property of this vector that is now, equal to pi j.

(Refer Slide Time: 19:44)



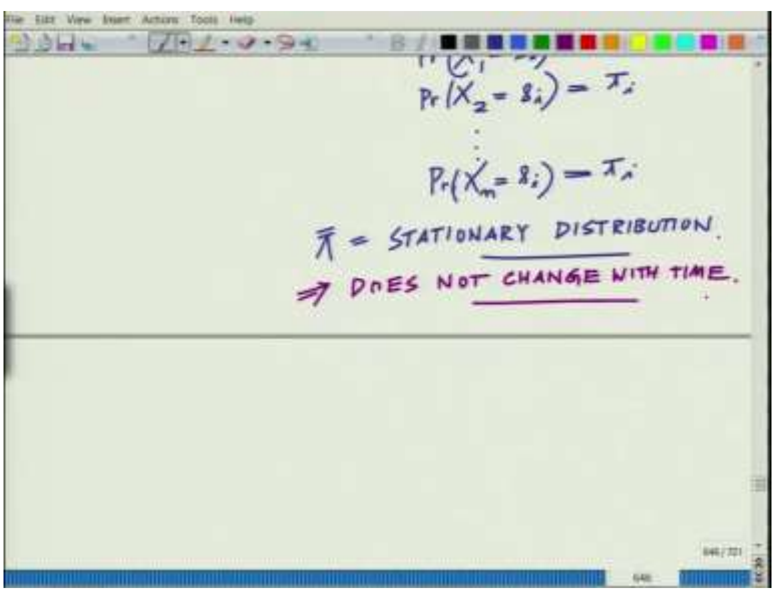
Therefore, what it means is if pi bar Transpose P if you find a vector pi bar Transpose P equal to pi bar this implies that if, we start with or when we start with probability X naught equal to S i equal to pi i we will end up with if you look at this probability X 1 equal to S j equal to pi j this implies that probability X 1 equals to S i equal to pi i.

(Refer Slide Time: 20:48)



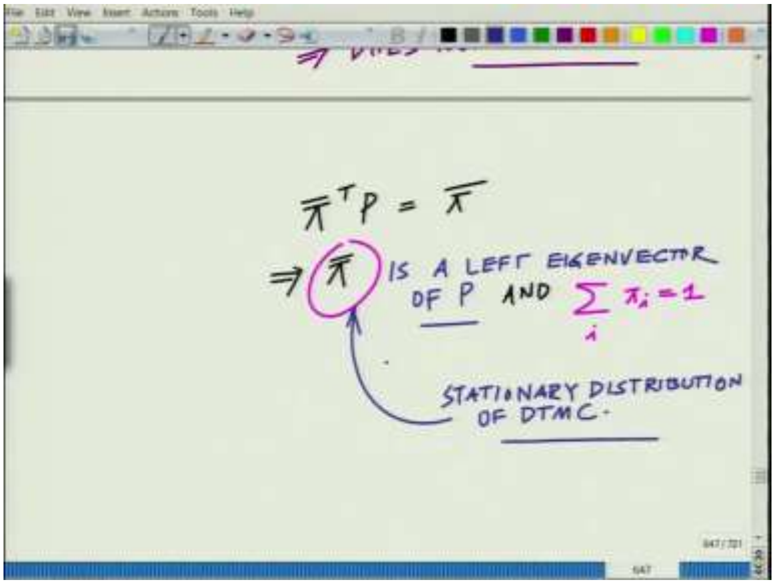
So, this implies that when we have probability X 1 equal to S i equal to pi i. Similarly, probability X 1, X 2 equal to S i will also be equal to pi i and so on and so forth probability any X n equal to S i equal to pi i and therefore, what this means is this is basically this is stationary the probability that X n equal to S i does not depend on n. So, this distribution is a stationary distribution. So, if you start with the distribution pi bar that distribution remains for the states at the next possible time instant at the time instant after that and so on.

(Refer Slide Time: 21:53)



So, this is distribution this pi bar is a stationary distribution, this pi bar is a stationary distribution that is does not change with time, this does not change with time.

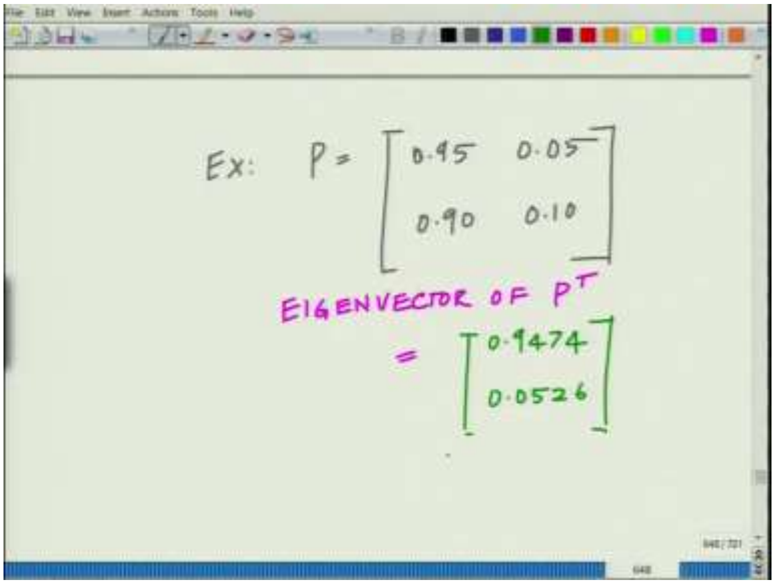
(Refer Slide Time: 22:36)

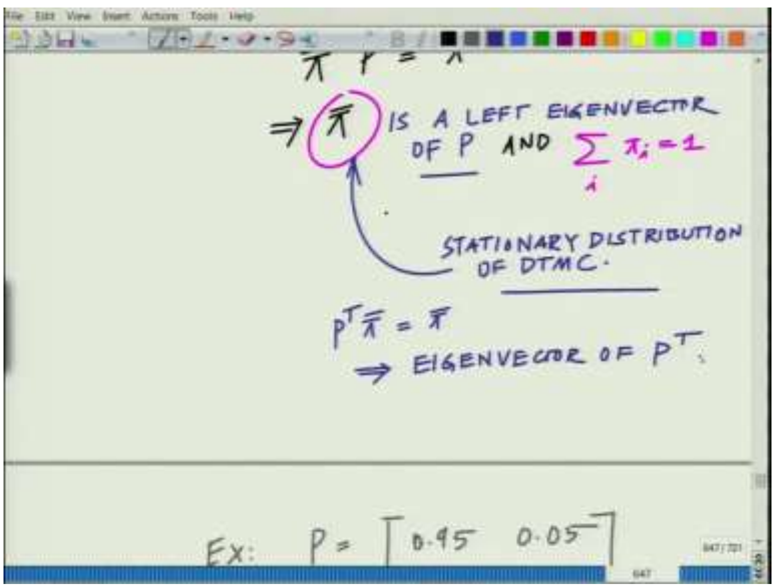


And when does that happen this happens with remember pi bar Transpose P equal to pi bar this implies pi bar is a left Eigen vector pi bar is a left Eigen vector of P, pi bar is a left Eigen vector of P and remember the Eigen vectors they can differ by scaling now to make it a probability mass function we have to ensure that it sums to identity.

And, summation over i pi i equal to 1 that is what we do is find the left Eigen vector of P such that summation of the individual elements is equal to 1 that gives you the stationary distribution of the DTMC. So, this pi bar is a stationary distribution this then becomes a this then becomes the stationary distribution of the DTMC.

(Refer Slide Time: 24:15)

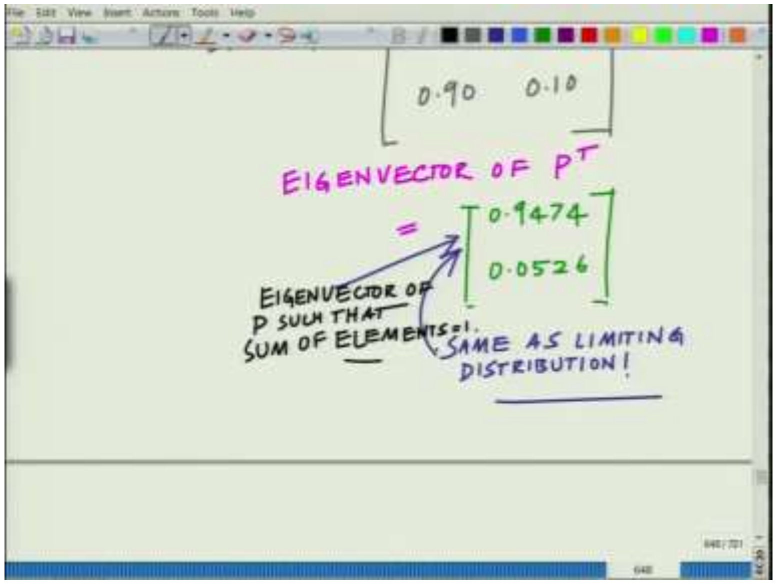




Example if, you take an example if, you take a look at again we go to our industrial reliability example which is our P equal to .95, .05, .90, .10 you can say. So, you can either say this Eigen left Eigen vector of P or if you take the Transpose you can also write it as P Transpose pi bar equal to pi bar implies the right Eigen vector or we simply call this as Eigen vector of P Transpose.

So, now take Eigen vector P Transpose and find what we ask the question what is the Eigen vector this is equal to you can see once again not very surprisingly this is basically .9474, .0526 which is nothing but the first row of P as n tends to infinity same as the limiting distribution by.

(Refer Slide Time: 25:38)



And, this is by and large true this is same as the limiting distribution that is this is the Eigen vector such that Eigen vector of P such that remember Eigen vector of P such that sum of elements is equal to 1 this is an important property to keep in mind Eigen vector of P such that remember you cannot take any Eigen vector.

Because, it is a probability distribution function probability mass function. So, sum of elements must be equal to 1 otherwise, it is not a valid probability mass function. So, Eigen vector of P such that sum of elements equal to 1. So, this is the same as the limiting distribution and that is essentially what you find.

So, this is a very interesting aspect of the discrete time markov chains what you find as P tends to n that is for a very large duration as P tends to n that as n tends to infinity P n that is the n step probability transition probability matrix approaches a matrix which is structure that all the rows of this are identical the each row of that then gives the limiting distribution of the DTMC that is the probability that after large time instant this the state x after large duration n the state this X n, the DTMC the state at time instant n is basically S j that is basically pi j.

And, what we have seen is that this stationary, this limiting distribution also is corresponds to what is known as the stationary distribution that is the DTMC starts in this initial distribution of states, initial probability distribution of the states that remains unchanged for any n and this stationary distribution can be found as the left Eigen vector of the transition probability matrix P.

So, this is a very-very interesting property of the DTMC and its relation to linear algebra and you can see as you can see linear algebra plays a very important role with matrices, products of matrices and so on. And, the Eigenvectors of it plays a very important role in the analysis of DTMC. So, let us stop here and let us continue in the subsequent modules. Thank you very much.