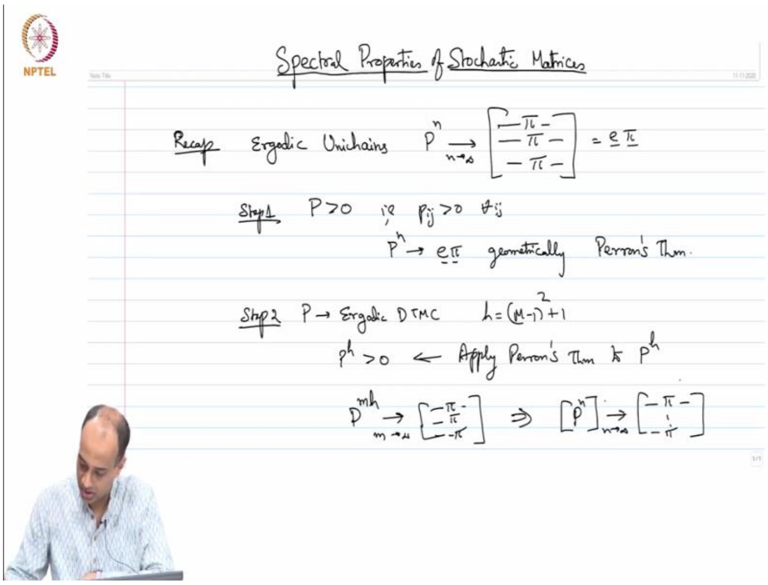
Stochastic Modeling and the Theory of Queues
Prof. Krishna Jagannathan
Department of Electrical Engineering
Indian Institute of Technology – Madras

Lecture –51
Spectral Properties of Stochastic Matrices - Part 1

(Refer Slide Time: 00:14)



Welcome back. Till the last module we were looking at the long term behaviour of finite state DTMCs. To give you a recap the basic result we proved was that for Ergodic Unichains the matrix P power n converges to a matrix of pi, pi, pi matrix where all the rows are pi as n tends to infinity. So, proving these roughly 3 main steps. The step 1 was considering P greater than 0 ie all P ij were greater than 0 for all i, j.

In this case we looked at the difference between the maximum element in any column of P power n and the minimum element and we said that the difference between the maximum and the minimum element goes down geometrically fast as n goes to infinity. So, in this case we could prove P to the n goes e pi so this is just e pi geometrically fast and this result is sometimes known as Perron's theorem.

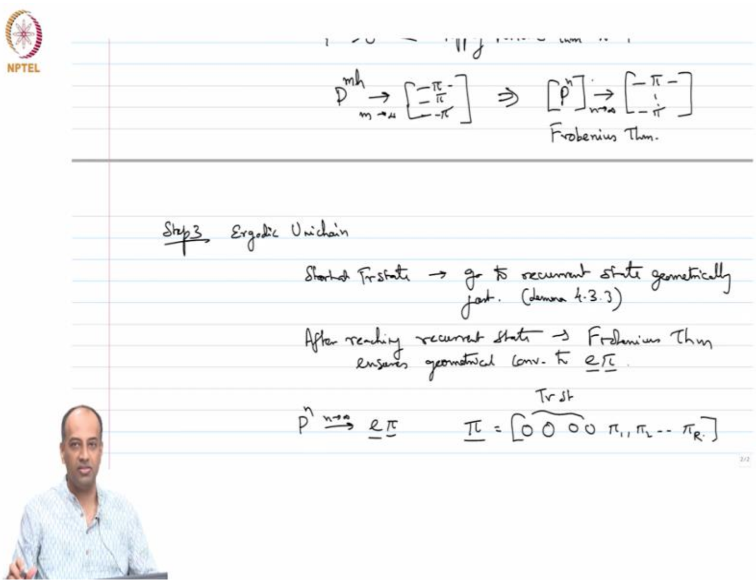
I did not mention this name last time. This result is known as Perron's theorem. Then for step 2 we considered the matrix P corresponds to ergodic DTMC. What that means is that the entire Markov chain consists of a single aperiodic recurrent class. In this case we took h = M – 1 square + 1. M is a number of states and we said that matrix P power h is strictly positive.

This can be proven then you apply Perron's theorem to P power h which is a matrix with all strictly positive entries.

Then we can show that P power mh goes to the e pi pi, pi, pi as n tends to infinity geometrically fast in m. Then we said that the maximum entry in each column and minimum entry in each column they are both monotonic in n the maximum is decreasing and the minimum is increasing. Therefore, this result implied that for all n P power n has to converge by monotonicity.

And the limit has to be the same as the limit along the subsequence mh as n tends to infinity oh sorry this was m tends to infinity and this result is known as Frobenius theorem.

(Refer Slide Time: 03:58)



This is known as Frobenius theorem. So, Frobenius theorem is a generalization of Perron's theorem, but it is also a consequence of Perron's theorem in the sense that Perron's theorem is used improving Frobenius theorem which is more general than Perron's theorem. Then finally step 3 we took an ergodic unichain. So, finally we want to prove that for ergodic unchain P to the n converges to e pi.

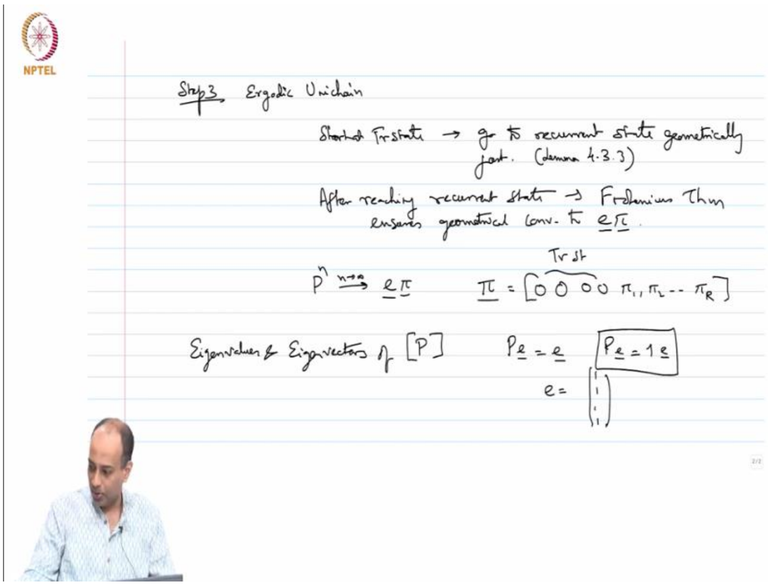
And ergodic unichain remember is a finite state DTMC in which there is one recurrent class which is aperiodic plus possibly some transient states. In this setting what happens is that if you start a recurrent class you are going to stay in the recurrent class and step 2 Frobenius theorem would apply. If you start at a transient state we manage to prove that we will eventually get to a recurrent state geometrically fast.

So, we said that the probability of starting at a transient state and remaining it one of the transient states after n steps goes down geometrically fast. So, you go from a transient state to a recurrent state geometrically fast and once you get to a recurrent state again you converge geometrically fast to pi according to Frobenius theorem in the recurrent classes aperiodic. So, here also we could prove that; so here we again had two steps.

So if you said start at transient state so we go to recurrent state geometrically fast the first proven lemma and then once you get to a recurrent state of course Frobenius theorem will ensure that geometrically fast convergence to e pi. So, in this case it is a two step thing. So, again P to the n converges as n tends to infinity to e pi where pi will consist of a bunch of 0s for all the transient states.

So, these are transient states and then so if there are recurrent states we will go to pi 1 to pi R which is the Markov. So, pi 1 through pi R is the row vector a stationary distribution vector for just the recurrent class and the transients states are 0 at a stationary probability. So, we have seen these 3 steps. Now this is just a recap.

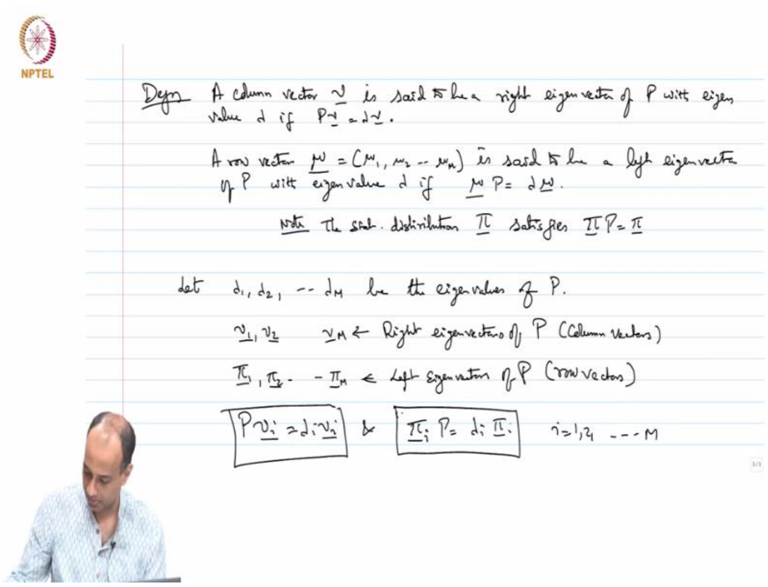
(Refer Slide Time: 07:36)



In this lecture, we will look at a closer look at I mean we will take a closer understanding we will get a closer understanding of how this convergence happens by looking at the Eigenvalues and Eigenvectors of stochastic matrices. A stochastic matrix is simply the transition matrix of a finite state DTMC. For stochastic matrix all the rows sum to 1 because given the bottom line being that if you are in any state the probability of going to if you will sum over P ij over all the j starting at any i it has to be equal to 1 because you have to go to some other state with probability 1 and therefore sum of all rows is equal to 1 which we can write as P e equals e.

Or rather just to be more clear so this can be written as P e equals 1 e where e is the column vectors of all one's. So, this already looks like a right eigenvalue equation. So, this looks like A x is equal to lambda x where A is just P here and e is the Eigen vector and the Eigen value is 1. So, this is the study that we are going to undertake now in little more detail. So, first I want to define right eigenvectors and left eigenvectors.

(Refer Slide Time: 09:28)



Definition a column vector nu is said to be a right eigenvectors of P with eigenvalue lambda if P nu is equal to lambda nu. So, P is a square matrix M / M matrix nu is a M long column vector and P nu equals lambda nu you say that nu is a right eigenvectors with eigenvalue lambda likewise a row vector now let us call it pi. Here I do not particularly mean that this pi is the stationary distribution maybe I should call it pi is a specific it is a specific left eigenvectors so I do not want to call it pi.

So, a row vector let us call it let us say mu so this is like mu 1, mu 2, mu M is said to be left eigenvectors of P with eigenvalue lambda if mu P equals lambda mu. So, P is the square matrix ambient square matrix mu is the row and mu P equals lambda mu then you say that mu is a left eigenvectors with Eigen value lambda. Now, note that the stationary distribution; note the stationary distribution pi satisfies as you know it satisfies pi P equals pi.

So, this is a left eigenvalue equation as you can see with lambda equal to 1 which is why I did not call the generic row the left Eigen vector as pi. So, pi is the specific row eigenvectors which it is a left eigenvectors which satisfies pi P equals pi namely with eigenvalue equal to 1. So, generally this is a M / M matrix let lambda 1 so it will have M eigenvalue be the eigenvalue of P.

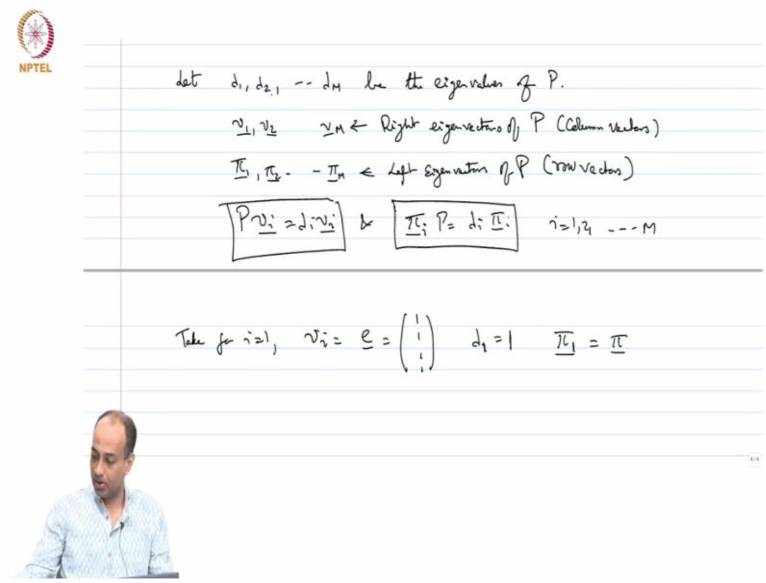
Please note that the eigenvalues of P may do not depend on whether you are looking at the right eigenvectors or left eigenvectors because the right Eigen vectors and left eigenvectors are both solved using the equation determinant. So, the right Eigen values you will solve it using determinant P – lambda i = 0. and the left Eigen vectors you will solve it using determinant P transpose – lambda i = 0.

But you know the equations those two questions are one on the same because determinant P – lambda i and determinant P transpose – lambda i are one on the same. So, when you are talking about eigenvalues of P we do not have to worry about whether it is a left eigenvalue or a right eigenvalue you will have a same set of eigenvalues. Of course, the Eigen vectors will be different.

So, what we will do is we will call nu 1, nu 2, nu M to be the right eigenvectors of P corresponding to lambda 1 through lambda M these are column vectors and we will call pi 1, pi 2, pi M to be the corresponding left eigenvectors of P these are row vectors. So, each eigenvalue lambda i is associated with a right eigenvector nu i which is a column vector and the left eigenvector pi i which is a row vector.

To be more explicit we are saying that P nu i = lambda i nu i and for the left eigenvectors we have pi i P = lambda i pi i this is for i equals 1, 2 dot, dot, dot, M. This is the right eigenvalue equation and this is the left eigenvalue equation for each i equal 1 to M.

(Refer Slide Time: 15:42)



And also we know that for i = 1 we will take nu i = vector e which is just vector of all 1’s and lambda 1 = 1 and pi 1 is the stationary distribution pi. So, that we already have P e equals e and pi P equals pi. So, we are looking at i equal to 1 is a special index which I am without loss of generality taking to be lambda 1 equal to 1 sorry right eigenvector is the vector of all 1s and the left eigenvector is stationary distribution pi.

Now, I am going to these eigenvalues lambda 1 through lambda M these guys need not be distinct. They could be repeated. However, the best intuition comes by looking at the case where all these lambda 1 through lambda M are distinct. So, that is where you can get the clearest intuitive idea of what is really happening to these eigenvalues and eigenvectors and in turn you can clearly understand how this P to the n converges to pi.