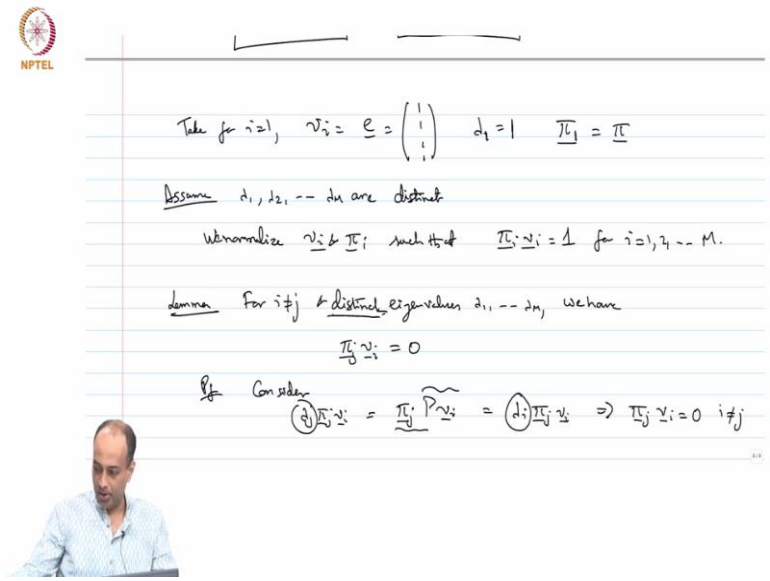
**Stochastic Modeling and the Theory of Queues**
**Prof. Krishna Jagannathan**
**Department of Electrical Engineering**
**Indian Institute of Technology – Madras**

**Lecture –52**
**Spectral Properties of Stochastic Matrices - Part 2**

**(Refer Slide Time: 00:14)**



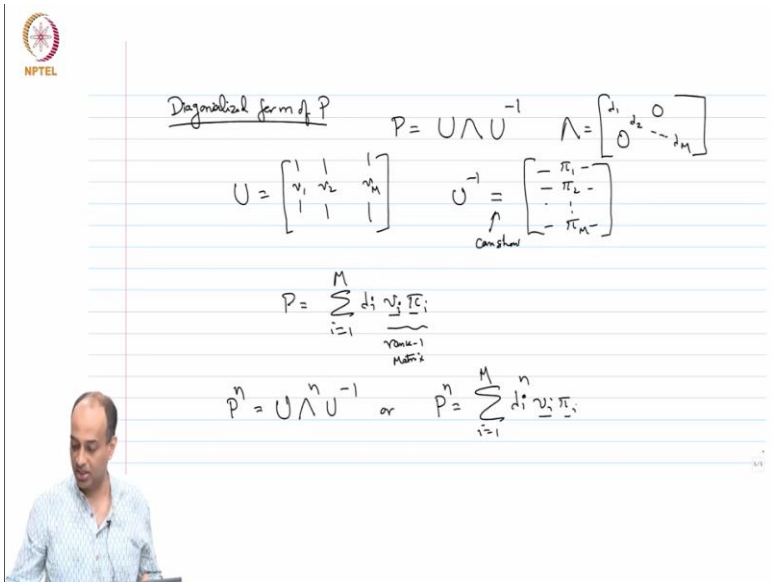

So, we will assume that lambda 1, lambda 2 dot, dot, dot lambda M are distinct. This is a special case not always the case that this is true, but this is the most intuitive case, but also we normalize nu i and pi i such that pi i nu i = 1 for all i. These eigenvectors can be scaled such that pi i, nu i = 1 for all i. For the special case i = 1 which is already true. This is just saying that for i = 1 this is pi e = 1 which means that the sum of all the stationary probability is equal to 1 which we already knows correct, but we are going to do this for all the eigenvectors.

Now, here is a interesting lemma say for i not equal to j and distinct eigenvalues lambda 1 to lambda M we have pi j nu i = 0. So, if you take the dot product between these left eigenvectors and right eigenvectors \for unequal indices you get 0. The way you prove this is you consider pi j P nu i. So, one way to write this is to realize that P nu i is just lambda i nu i. So, I can write this as lambda i pi j nu i.

And the other way to write this is to take the left eigenvalue equation I will write these as lambda j pi j nu i and these lambda j and lambda i are not equal. You are assuming distinct

eigenvalues this is the special case of distinct eigenvalues. This implies that this implies what you want implies pi j nu i = 0 for i not equal to j.

**(Refer Slide Time: 03:27)**




So, this will be useful in writing out the diagonalized form of the matrix P. Since the matrix P is assumed to have distinct eigenvalues you can always write a diagonal representation of P or diagonal decomposition. Namely P can be written as U lambda U inverse where lambda so I should tell you what these matrices U and lambda are. This lambda is simply the diagonal matrix of all the eigenvalues.

U is the matrix of all the right eigenvectors made into a matrix. You can always show that this sort of a Diagonalization can be done as long as the lambdas are distinct. Now, what is U inverse? So, what you can show using the above lemma is that U inverse (()) (05:12) equation will simply turn out to be the matrix in which all the left eigenvectors are arranged in their rows.

This follows from the fact that so you can easily verify U inverse what I claim to be U inverse is in fact the inverse of U by simply noting that the pi i lambda i = 1 for all i and sorry I think I made a mistake I beg your pardon these are eigenvectors so I should be writing nu I made a mistake. So, you know, that pi i nu i = 1 and pi i nu j is equal to 0. From that you can prove that the structure of U inverse is like this, this is using the lemma above.

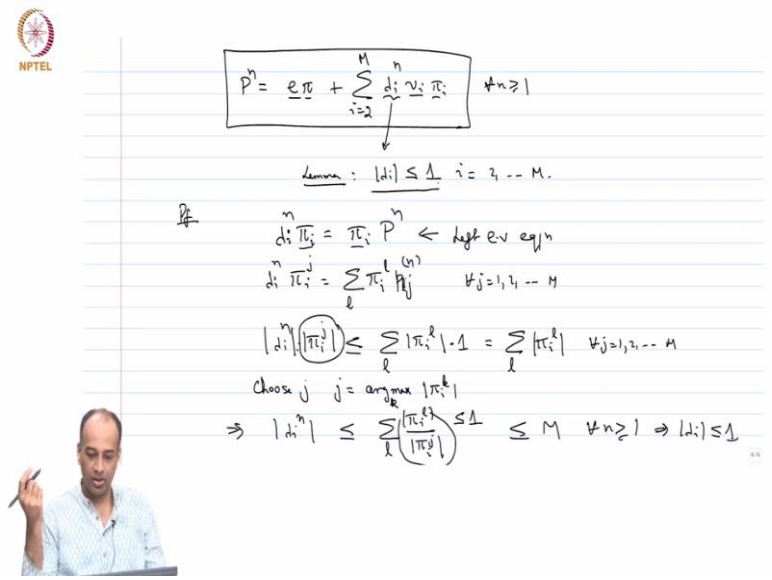
So, this is great. So, this kind of a diagonalized representation is possible. If you write out this matrix P another way of writing P = U lambda U inverse is simply to write sum i equals 1

to M lambda i nu i pi i. So, this nu i pi i is rank 1 matrix. This is nothing, but the spectral decomposition of the matrix P or you can just look it as a rewriting of the equation P = U lambda U inverse.

Next, how about P power n? Can you write a similar spectral decomposition of P power n. Interestingly the matrix P power n has eigenvalues lambda i power M where lambda i are the eigenvalues of P, but the eigenvectors will be the same. You can show that P power n nu i will be lambda i power n nu i and similarly for the left Eigen vectors. So, all in all you can write the matrix P power n will have a diagonal representation U lambda power n U inverse.

So, only the diagonal matrix get raise to the nth power this is basic linear algebra or you can write it as P power n = sum over i equals 1 to M lambda i nu i pi i. This is the spectral decomposition of the matrix P power n, but we know that lambda 1 = 1 we have taken lambda 1 = 1 and nu 1 = e and pi 1 is equal to the stationary distribution pi.

**(Refer Slide Time: 08:46)**




So, I can just write that out I can just pull that as a special term so I will have e pi + sum over i = 2 to M lambda i power n nu i pi i. So, this is an important equation this is nothing, but the spectral decomposition of P power n. So, P power n so this is true for all n greater than or equal to 1. So, for all n greater than or equal to 1 we have an explicit expression for P power n which is the matrix of all pi, pi, pi plus a bunch of terms.

Now the question is what happens to these terms i = 1 to M lambda i to the n nu i pi i. So, if this sum from i = 2 to M goes to 0 then we can easily conclude that P to the n converges to e

pi. Now what happens is the following see these eigenvalues lambda i which are the eigenvalues other than the one eigenvalue these can be shown lemma that to be having magnitude less than or equal to 1.

Actually lambda 1 = 1 and the others have magnitude less than or equal to 1. The way to proof this is the following. Remember that the lemma only says that the absolute value less than or equal to 1 for i = 1 to M. I can just write i = 2 to M because I know that lambda 1 = 1 so maybe I should just write i = 2 to M. Now, if all these lambda i from 2 to M have a magnitude strictly less than 1 then I will have convergence of P to the n to e pi.

We know that in the case of an ergodic Markov chain these convergence happens we already proved it in Perron's theorem. So, for an ergodic Markov chain we can conclude that the absolute value of lambda i is actually less than 1. In general, if you do not have ergodicity the eigenvalues are less than or equal to 1. The way you prove this lemma is as follows. You consider the left eigenvalue equation lambda i power n pi i = pi i P power n.

You consider the jth component of this. This is just the left eigenvalue equation for pi i. So, if you consider the jth component of this you will get lambda i power n pi i jth component = sum over l pi l P lj n. So, pi i l P lj n. So, the way you prove this lemma is to write out the left eigenvalue equation lambda i power n pi i = pi i P to the n. This is just the left eigenvalue equation for P power n.

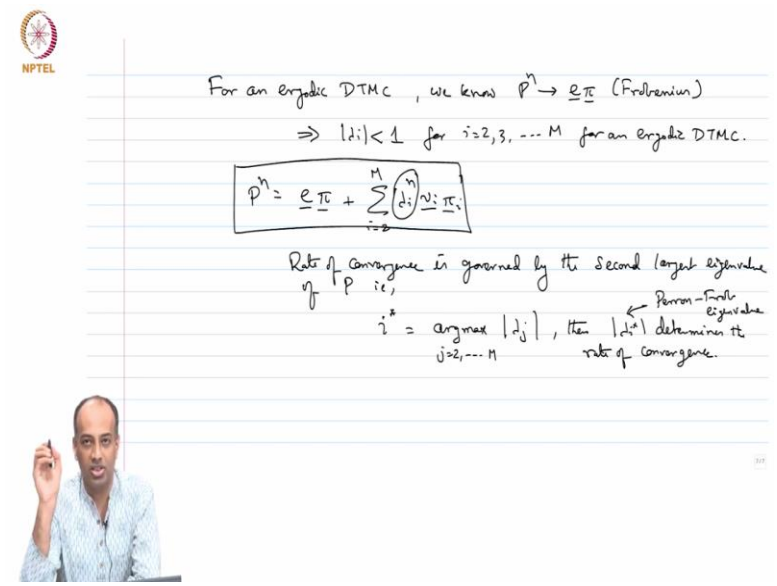
Now, you look at the jth component of this equation you get lambda i to the n pi i j is the jth component of the row vector pi i = sum over l pi i l P lj n. This is true for all j equals 1 to dot, dot, dot M. From this we can get lambda i to the n is equal to so maybe I should put absolute value I can write is less than or equal to absolute value of lambda i to the n pi i j is less than or equal to sum over l absolute value of pi i l and the absolute value of P lj n is less than or equal to 1 so I can upper bounded by 1.

So, this just turns out to be sum over l absolute value pi i l. This is true for all j equals 1, 2, dot, dot, M. Now you choose j this is true for all j the above equation so choose j such that j = Arg max absolute pi i k over all these so to take the j for which that guy is this absolute value is the biggest then what happens is that you will get absolute value of lambda i to the n is less than equal to sum over l absolute value of pi i l over absolute value of pi i j.

And you have chosen j such that these ratios are all less than or equal to 1 because you have chosen the j for which absolute value of pi i j is the largest. So, each of these term (()) (15:46) is less than or equal to 1 which means that this sum is less than or equal to M and this is true for all n greater than or equal to 1 and the only way that absolute value of lambda i to the n is less than or equal to M for all n is if absolute value of lambda i is less than or equal to 1 because if absolute value of lambda i were bigger than 1 then sum nth power will be bigger than M and that is not possible.

So, we have shown that all the eigenvalues are less than or equal to 1. The first eigenvalue is of course equal to 1 and then the subsequent eigenvalues are all less than or equal to 1.

**(Refer Slide Time: 16:30)**




Now we know that for an ergodic Markov chain we know that P power n converges to e pi as n tends to infinity. How do we know this? This we know this from Frobenius theorem which means that if you look at this spectral decomposition which means that the sum from i equal to 2 to M must go to 0 and that means that all these lambda i for 2 to M must be have absolute value strictly less than 1.

So, this implies absolute value of lambda i is strictly less than 1 for i = 2, 3 dot, dot, dot M for an ergodic DTMC. So, you have this kind of a convergence so P to the n I am just rewriting this equation P to the n is e pi plus sum over i = 2 to M lambda i to the n nu i pi i and all these eigenvalues have absolute values strictly less than 1 and therefore all these terms go down exponentially fast geometrically fast.

And therefore we have a very clear understanding now that P to the n converges to e pi geometrically fast and we even have the rate of convergence. The rate of convergence for an ergodic Markov chain is governed by the second largest eigenvaluee of P ie you look at the i star for which has the largest magnitude of these lambda j where j is from 2 to M. Remember 1 is equal to 1 you forget that.

And you look at all the other eigenvalues and take the eigenvalue with the largest magnitude. Overall, this is the eigenvalue with the second largest magnitude. The largest eigenvalue being 1 and then there is a whole bunch of eigenvalue so absolute value is strictly less than 1 for a ergodic Markov chain and you look at that index for which the absolute value of the eigenvalue value is the largest.

Then absolute value of lambda i star will determine the rate of convergence and this eigenvalue which has the second largest magnitude they are sometimes known as the Perron–Frobenius eigenvalue. So, what is the Perron–Frobenius eigenvalue of P? There is a eigenvalue which is equal to 1 you forget that and look for all other eigenvalues other M – 1 eigenvalues you take the 1 with the largest absolute magnitude.

Of course, the lambda i star itself may be complex, but you are looking at the eigenvalue with the second largest absolute value and that is known as the Perron–Frobenius eigenvalue and this Perron–Frobenius eigenvalue determines how fast the matrix P power n converges to e pi. If the Perron–Frobenius eigenvalue has a magnitude very close to 1 than the convergence is very slow.

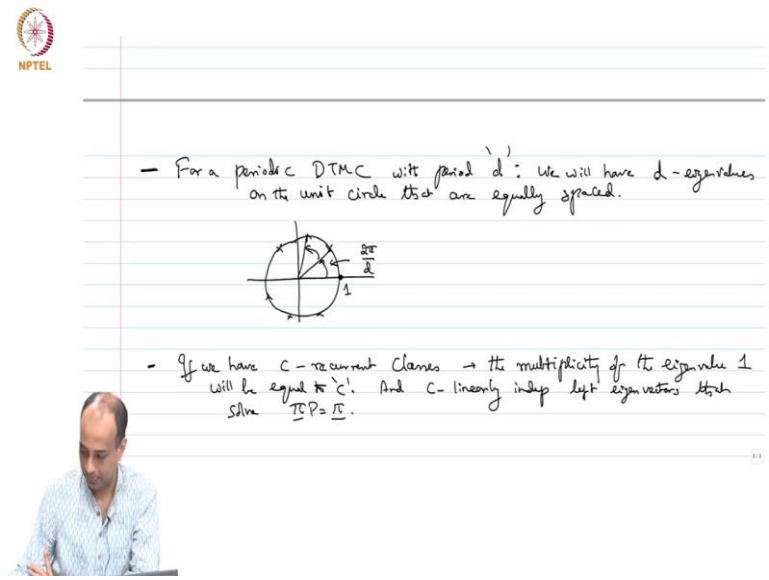
And conversely if the absolute lambda i star this Perron–Frobenius eigenvalue if it is fairly small in magnitude then the convergence to steady state pi is very fast. So, this is a good exercise, it basically just uses linear algebra in the sense of this spectral decompositions and we get a very fine grained understanding of how this P power n behaves. For ergodic Markov chains we have a very precise characterization of the rate of convergence.

We know that is geometric already from Frobenius theorem, but the exponent in the geometry convergence is exactly characterized to be the second smallest largest eigenvalue in magnitude. If the Markov chain is not ergodic there could be other eigenvalues lambda i with

absolute value = 1. In that case P power n will not converge to e pi. So, I will briefly discuss what happens in those cases?

So, in the ergodic case in the aperiodic case you have convergence to e pi which is geometrically fast. In the periodic case what happens is you will have other eigenvalues whose magnitude is equal to 1 you will have eigenvalues on the unit circle.

**(Refer Slide Time: 22:36)**




So, these I will just state without proof for periodic DTMC. Let us say I have one recurrent class and the recurrent class is periodic. In this case we will have d eigenvalues on the unit circle in the complex plane that are equally spaced. This I am not going to prove what I am saying is that if you have this let us say this is the unit circle on the complex plane. I know that there is an eigenvalue at 1 always for a stochastic matrix.

You will have other eigenvalues which are equispaced on the unit circle. You have d such eigenvalues and the spacing between all these guys will be equal and this angle will be 2 pi over d. So, we will have d equally spaced equally spaced eigenvalues on the unit circle for a period d Markov chain. So, you will not have convergence if this expression out here this spectral expansion here is still valid except that this lambda i power n some of these lambda i have magnitude equal to 1 they will lie on the unit circle.

So, if the nth power does not converge to 0 that is what happens in a periodic Markov chain. The way you prove this is using the fact that in a period d Markov chain you can partition the states into a d separate subsets where all the transitions go from S 0 to S 1, S 1 to S 2 and so

on till S d – 1 to S 0 we saw this result already and using that result we can argue that the matrix P will have to have a certain rotating structure among these blocks going from S 0 to S 1 and so on.

And then we can prove that these eigenvalues lie on this unit circle equally spaced. This is there is a guided exercise is Gallagher book for this, but I am not going to prove it here. The other case I want to point out is that if you have c recurrent classes what will happen is that we will have the multiplicity of the eigenvalue 1 will be equal to c. Remember that in the previous discussion I have only covered the case when all the lambda i are distinct.

That is not the most general case if these lambda i were not distinct then you will have a bit of an issue in the sense that eigenvector lambda i so the eigenvector nu i and the left eigenvalue pi i may not span the space span R M. In that case this sort of a diagonal elevation U lambda U inverse may not work. So, if you have repeated eigenvalues you may not be able to appeal to the mechanism that I have talked about using diagonalization.

You have to go into a formalism called Jordan form in linear algebra which is in my opinion not very directly intuitive for the study of Markov chains. I am just saying that you can do it using Jordan's form. I will not get into that in detail in the specific case when you have c recurrent classes then the multiplicity of the Eigen value 1 will be c. So, there will be c repeated Eigen values at lambda equal to 1.

And you will have c linearly independent left eigenvalues or let us say eigenvectors that is all pi P equals pi. So, there will not be a unique pi like this. There will be c linearly independent solutions to pi P equals pi and of course all the linear combinations of these linearly independent solutions will also be solutions. So, if there are c recurrent classes we will have c repeated eigenvalues at lambda equal to 1.

And correspondingly we will have c linearly independent solutions to pi P equals P. So, these two things I am not going to prove so this is where I will stop.