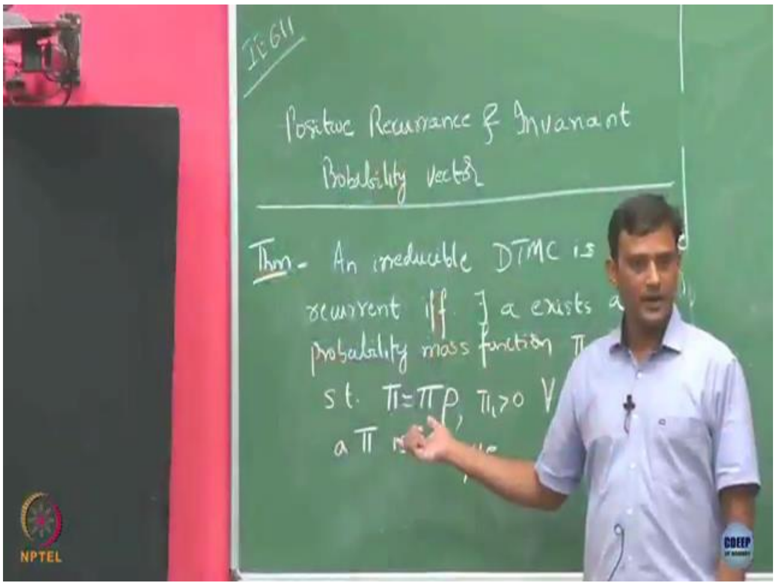
Introduction to Stochastic Processes
Professor. Manjesh Hanawal
Industrial Engineering and Operations Research
Indian Institute of Technology Bombay
Lecture 38
Positive Recurrence and The Invariant Probability Vector

So, next we are going to focus on some property called invariant probability vector associated with a Markov chain and this invariant probability vector has some nice properties, when my Markov chain happens to be positivity recurrent. So, suppose let us say I have my Markov chain that based on my communicating class relationship, splits into different classes and let us say one of the communicating class turns out to be closed.

Now, can I, after I do this, can I ignore all the other possible, other possible classes and then just focus on this class? And I assume that my Markov chain is strictly to this states in this class, I can do this, because I am always circulating within this class and I am not going to visit any other states.

So, I could ignore all other states and I can just think of my Markov chain is just to this state and when I do that, I can now think of my Markov chain restricted to that communicating class I can just think, that my on that class my Markov chain is irreducible. Because it is only on this states which are already closed and communicating, close communicating class.

(Refer Slide Time: 02:01)



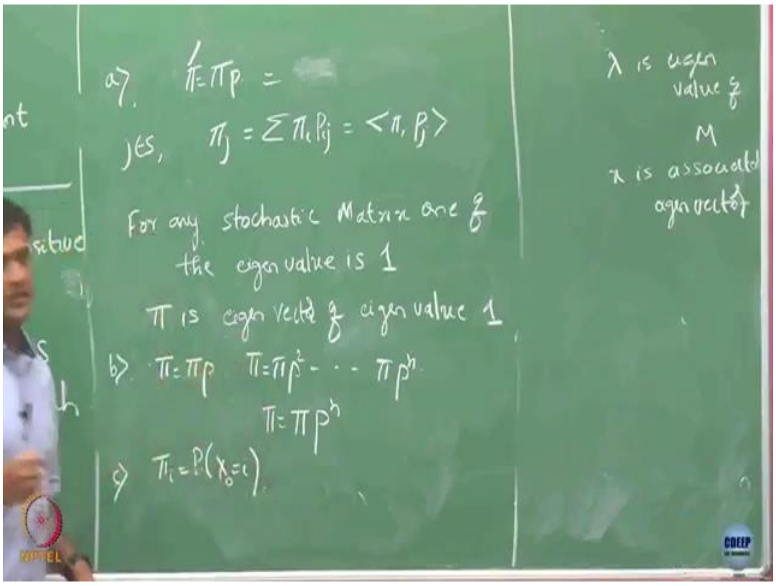
So, that is why I am now henceforth going to assume irreducible Markov chains. So, that means basically I am saying that, if you have Markov chain has many multiple classes and if it has any of the many of these classes happen to be close communicating class, then I could just focus or attention on that class and on that class my Markov chain, I can just think it as an irreducible Markov chain.

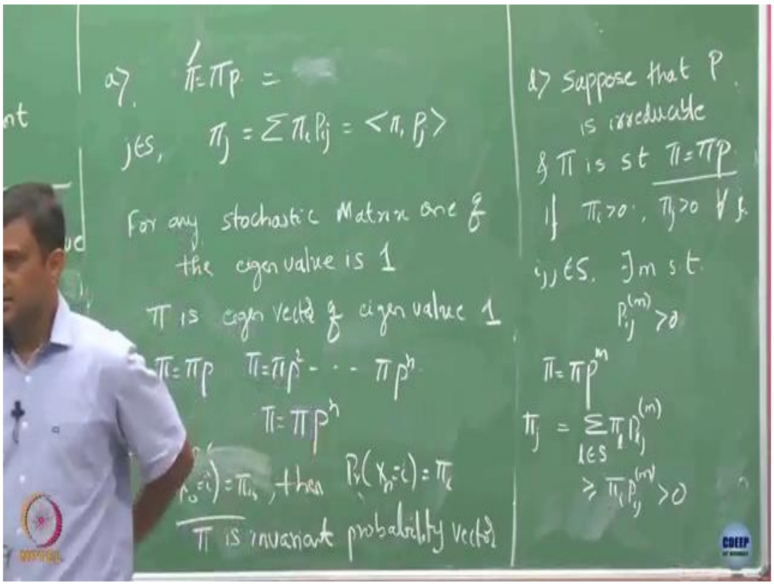
So, this is main theorem, so, whenever my Markov chain and that Markov chain, so that class if it happens to be positive recurrent, a closed communicating class, I am not making any assumption of finiteness it could be arbitrary. So, if it is finite, I know it is already positive recurrent, if it is not, I do not know it could be anything else. But suppose, let us say it is positive recurrent. I know if it is finite, it is already, going to be positive recurrent, even if it is not, let us assume it is a positive recurrent.

So, in that case, we have this nice property. So, if let us say I have a Markov chain that is irreducible and that is positive current if and only if, this is both necessary and sufficient condition, if there exists a probability mass function Pi on my state space such that Pi is a solution of this relation Pi equal to Pi P and all these Pis are positive and it says that further this such a Pi is going to be unique.

So, it is clear, so if my Markov if my Markov chain irreducible Markov chain happens to be positive recurrent it has this nice property that there exists this probability mass function which satisfies this solution. So, now it is if you see that this Pi has lot of nice interpretation as we go along, first let us look into some of its properties.

(Refer Slide Time: 06:17)





So, what is p here? It is the transition probability matrix of my DTMC. So, this expression basically Pi equals to Pi P, What is this? I return in a compact form, but what is this?

Student: (())(06:38)

Professor: So, if this is actually, so, this is going to be what? Many simultaneous equations are there at in this, it is a basically a collection of simultaneous equations. So, if you are going to look at state j. So, what is Pi j is?

Student: (())(07:04)

Pi i or S, Pi is a probability mass function on S. What does that mean?

Student: (())(07:25)

Professor: Can it be? So, Pi is a basically probability mass function. So, it has to be vector, it is going to give probabilities to each of my element in my state space S. So, the way we are going to see is this is a column vector,

Student: Row vector

Professor: Row vector and this is another row vector that is multiplied in a matrix and now, if I want to find the jth element in that row vector, what I have to do? I have to take the jth column in my matrix p and the dot product of that 1 with my Pi. So, this is basically saying, so this is nothing but Pi and my Pj. So, suppose if I interpret my Pj as the jth column of my transition probability matrix, then my Pi j is nothing but this quantity and this is for all j.

So, how many simultaneous equations I have here? So, this is going to be the cardinality of s. And so, if I am going to treat this Pi as, let us say variables and want to solve this. So, how many, so how many simultaneous solutions I have and in how many variables. So, this j is coming from s. So, this, so there are actually cardinality of S such equations and there are also, that many variables. So, we have as many equations as a number of variables here.

Now, this P here we know that is a transition probability matrix, this transition probability matrix we know this is a Stochastic Matrix. All the nice property of all these Stochastic Matrix is they have an eigenvalue, whose value is what? So, they will have many eigenvalues, but one of the eigenvalues will have a value equals to 1.

So, this is for any Stochastic Matrix, you know what is eigenvalue? So, you can check this this is just a property, which I am stating here. So, if I have Stochastic Matrix, that is, if it is rows add up to 1, all the rows and all the elements in that matrix are positive, so you can verify that it will help one of the eigenvalue size 1. So, if 1 of 1 is the eigenvalue, what can we say about this relation in for Pi? So, if lambda is an eigenvalue of, let us say lambda is eigenvalue of M and let us say x is associated eigenvector. So, how can I write?

Student: (())(11:37)

Professor: So, m x equal to lambda x, using that relation, can you what you can say about Pi equals to Pi P, eigenvector what?

Student: (())(11:49)

Professor: Of what?

Student: (())(11:52)

Professor: So, Pi is eigenvector of eigenvalue 1, we can say that. So, what we trying to seek is? Eigenvector associated with my transition probability matrix whose eigenvalue is 1 and further we want this Pis to be positive because that is the requirement. Now we know that Pi equals to Pi P, I can repeat this iterations and write, Pi equals to P Pi square and like that P equals to Pi to the power n. So, basically saying that Pi equals to Pi is basically saying that this relation should also, be true for any n greater than or equals to 1.

Now, why is the name invariant probability that we are calling it. Suppose you take some Pi which is which is the value of your initial probabilities of your Markov chain. I said that

Markov chain is going to be characterized by its initial distributions and the transitional probability matrix. So, you said, so we know that suppose you take this Pi equals to Pi P solution, whatever that Pi is, and set the initial distribution of your Markov chain to be that Pi value.

Then it is so happens that I am going to say this such, then it so happens that the probability that in the nth round, your Markov chain taking value i also happens to be the same probability. So, if you start your Markov chain to have initial distribution that corresponds to the solution of this relation Pi equals to Pi P then it, so happens that in every round probability that your Markov chain is going to particular state is going to remain the same distribution. So, that means if you start your Markov chain with this initial distribution Pi in every round, the same distribution continues to hold, we are not saying that this is probability.

So, you start your Markov chain initially. So, if I start my Markov chain, let us say I am going to start it in state i with probability Pi i. Now, let it run, let us say let it run for 10 rounds. Now the probability that we are going to see it again in state i is going to be the same Pi i and now we are going to run it 100 rounds and then ask what is the probability that it will be in state i that is still going to remain the same Pi i. If this Pi is are selected such that they are the solution of this Pi.

So, that is why this Pi is called invariant probability distribution. So, this is Pi is called. So, check this, this is easy to verify, if you are going to start from this relation just check that if I want to do it in the next round, whether this relation holds and then try to see that this holds for any possible m using this relation, Pi equals to Pi n, here whatever you have written.

Next property suppose that this Pi is irreducible and by Pi is such that Pi equals to Pi P. If I have a solution Pi which is coming out of the relation Pi equals to Pi p, then it so happens that if in the solution P i is greater than 0, if one solution is positive it must be the case that for all i and for all j.

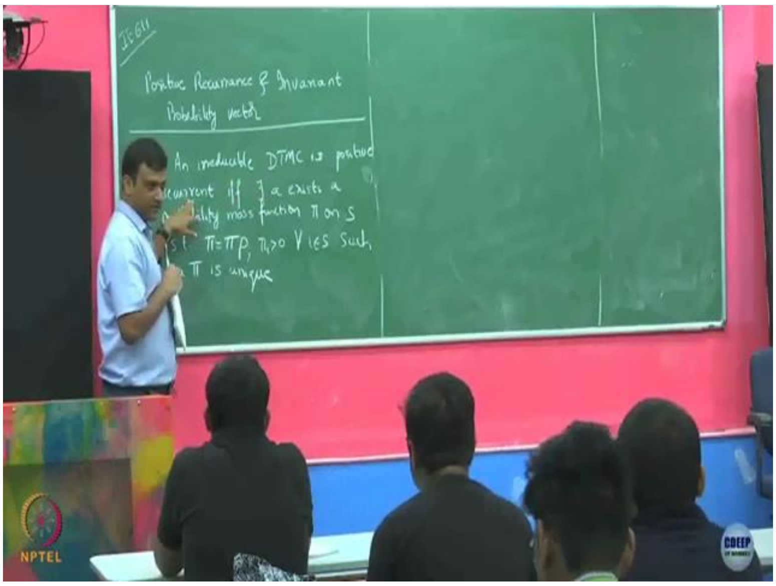
Say previously, if I am going to set this Pi to be all 0 vector, then the relation holds. But let us say I have one solution, where one of a component is not zero, one of the component is not 0 or positive, it must be the case that the solution is such that every component there is going to be non-zero. Why is that? So, let us take a state i and j belong to S and now I am looking into a individual class, I know that there exists some m such that Pij of m is positive, because I know i and j communicate, they belong to a same class.

And now if this relation is true, then this relation is also true for the same m. I have just recursively use for Pi and then I can write this relation now P i now let us say some j is equal to this is going to be L S and this is like from L, I am going to state j, so, this is L to j in m number of steps. So, this is our all possible states. This is my definition of this, this is the meaning of this relation here.

Now, in this summation, just focus on the one term here, so this term I know that this is going to be greater than or equals to Pi i and Pij m. I am only looking at the term where L takes the value j. I know that I initial my assumption is this Pi i is positive and now I already shown that because if j another state there must be some probability like this Pij of m is going to be greater strictly greater than 0. So, now I have two terms were both of them are strictly positive. So, then it must be the case that this guy is also going to be positive.

So, any j any Pi j should also I have be taking positive value. So, one element in this vector Pi is non-zero or strictly positive that means all the elements in this vector i should be positive. So, these are some of the properties which I think we should be comfortable in using about this P. Now let us see why this holds, so what I will do is, this is both this proof in works in both sufficient and necessary conditions. So, we will only show the necessary part.

(Refer Slide Time: 21:52)

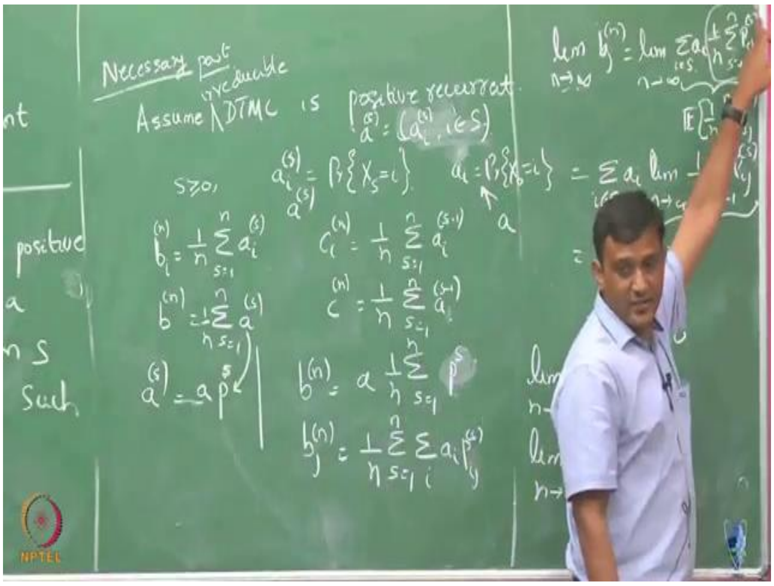


So, what we will show is so, it says what is that said? If and only if there exist a probability mass function on this such that this. So, if there exist some Pi which holds this relation then we are going to say that it is going to be?

Student: (())(22:08)

Student: Irreducible but positive recurrent the DTMC. But suppose we will start with the case that, we will start with the case where we start assuming that it is a positive recurrent DTMC then we will try to show that there exist a Pi which satisfies this relation and further that go5ing to be unique and also all the elements in that are going to be non-negative.

(Refer Slide Time: 22:41)



So, let us, we will use this notation. So, for any S greater than or equal to 0, we will use this notation that, a i of S, is basically probability that my Markov chain takes value i in in S step. So, this one, I am simply what is this? This is the probability that my Markov chain takes state S, in step in round S, that I am going to denote it as a i S.

And now what we will be looking at is these two quantities. So, what is b i n looking at? b i n is looking at the average of this probability, this probabilities at for the first n steps. So, I am just looking at what is the probability of, that I am going to take state i in the first step, what is the probability that I am going to take state i in the second step. What the probability that I take state i in the nth step, and I take add all of them and then take the average, this is the same quantity, but I am looking at the delayed versions, one step before instead of S step I am going to start at s minus 1.

So, here then S is equals to 1, this is a i to the power 0. No, which is my, whatever my initial distribution of my Markov chain. When a i of 0 here is x, this is probably that X 0 equals to i, that is my initial distribution. So, this guy c i n includes my initial probabilities, but this b i n does not include, because it is starting from 1 to n and actually, I am going to denote my, instead of that, so I am going to simply denote probability of x naught equals to i as simply a i. Instead of writing it as a i to the power 0, I will just write it as a i that is my initial distribution.

So, now can I write my. So, now these are all for a particular state, now I want to write it more generally instead of i, I am going to now write it simply b i of n, 1 by n, S equals to 1 and n, and then simply write a of s. So, now these are like vectors whose ith component is this. So, when I looking at the ith component of this b n, I will look only ith components of this a s and get this.

So, similarly this is also, going to be c n of n. So, are you comfortable with this notation? Just like instead of writing component wise I am writing them in vector form now. Now, I am going to see that. So, this one also, like instead of this is component wise when I want to write it as a vector, I am going to just write it as a of s.

Student: (())(27:26)

Professor: So, this is now a vector, if you look in the ith component in that, that is going to be a i of s, that is called tell me what is the probability that I take state s in the sth round.. So, just to be clear, what a of s is? Nothing but a of S 1, a to S 1, like whatever like, or like I can just say that this is nothing but this is a i of S where my i belongs to where? Capital S, set of my states.

So, now let us write to see my set of, so this vector, what is this vector is telling me? This letter is basically telling me, what is the probability that in a sth round, I am going to take different, different states and that probabilities are captured in this vector.

So, this one, I can always write it as simply a, this is my, so, now a is a vector for this notation, which is the vector of this probability. So, I can write it as a times p of x, what is p? My transition probability matrix. So, what is a of S?

Student: (())(28:59)

Professor: So a of s, let us focus a particular component in that let us say a i s. So, that is basically telling, what is the probability that I take state S in the Sth round? Now, how can I write this? I can say that, you are going to start with different probabilities and in the Sth round you are going to get to some state? How I am going to get that is by multiplying a with p to the power S. So, what is p to the power S is giving you?

Student: (())(29:32)

Professor: A step starting probability. But now if you are going to start from, what are your initial stage and then multiply it with your resistant probability that should give you what is your states with what probabilities are going to reach different states in the Sth round.

So, just I mean, just to try to follow this notations, then what we will do is. Now we are just going to plug back this relations what we have so far. Now in this case b of n is going to be a times 1 by n. So I have simply plugged in this relation here and I have just pulled out a outside and this is what I have here. So, if you want to write it, so this is the compact notations in terms of the vector.

If you want to look at the particular jth component in this, this is how it is going to look like. So, this is going to look like, so this is a is a vector. So, you can pull out 1 outside and what is this is going to be then look like is summation S equals to 1 to n, then summation i, a i and then P i of j of S. Now, what we will do is? We will now look in the limit of this.

So, I want to now look at limit as n tends to infinity of this quantity bjn, which is nothing but limit as n tends to infinity just slightly reorganize this. What I will do is? This summation a i summation 1 by n I summation Pij of S the, and what is S here? S is going from 1 to n and this i is our state space. The same thing here I have slightly reorganized it by bringing this summation outside and taking 1 by n inside.

Now, I want to interchange this limits, can I? So see that this a i is what? These are probabilities, a i is your probability that in the you start in state i. So, in that sense, if I look into this, this is nothing but some expectation of this 1 by n summation s equals to 1 to n of i j, what is i change this i to capital I here, because now this is the random variable and that is going to be taking the probabilities as per this distribution a i, because now this is the.

What is this? What is this a i is? Probability on the state i and for that i this is the value I have. So, I can now think of this as an expectation term here. But now the question is, now if I want to change this summation and limit it is same as asking the question, can I change this limit and this expectation here? Now, now, let us come back to our things we have studied, if I want to interchange this limit and expectation can I do here? Or if at all I can do I am I can do it here. So, is that, any of the three theorems we studied to interchange limit and expectation, applies here.

So, now what? So, this is my distribution and this is like the value taken by my random variable, you have to map before you apply those results, you have to see, what is the random variable here? What is the distribution here? So, here you can say that with probability a i this is the value taken, so that is why this is the expected quantity.

Now, so the first thing we studied in our results where we wanted to interchange expectation limitation was, first one was what? Bounded.

Student: (())(35:14)

Professor: Bounded convergence theorem, is bounded convergence theorem applicable here? So, what is the values of the random variable taken? The value the values of the random variables are this, 1 by n summation Pi j of S and that is changing with different values of i, is this value bounded?

So, P S j here are probability terms, this summation here cannot be more than n, but we are already do it good by n, so this is going to be less than 1, so this is going to be a bounded. So, all my random variable the values here are already bounded, with by 1 with probability 1. So, I could use my bounded convergence theorem and interchange this part.

Now, if I know that my state is recurrent, in this case I know more, I know this is I have started with assuming my state is positive recurrent, what I know about this quantity? This might just my state is positive recurrent.

Student: (())(36:48)

Professor: Greater than?

Student: 0

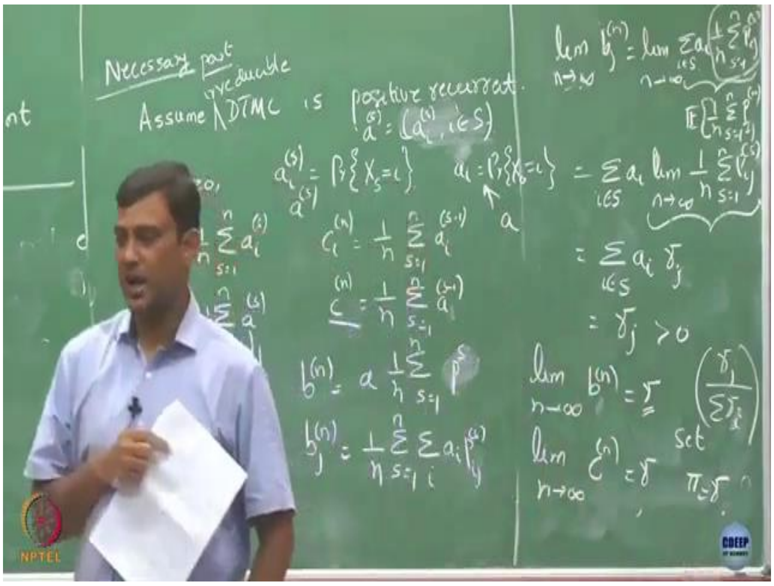
Professor: 0 and we denote it to be, let us say that some this is some quantity gamma i. So, it is greater than 0 that means it is some value, let us say that is gamma j, which is strictly positive. So, now, this gamma j, so this is gamma j here, and this is independent of all the index of the summation. So, I can pull out this gamma j outside and then what is summation of a i is going to be? What is the summation of a i is going to be? 1, because this a i is other probabilities.

So, this is simply going to be gamma j in this case, which I know is strictly positive by my theorem. So, what I basically did is? This I did for a particular j, now, I can do it for all possible j states and write limit as n tends to infinity. So, b j of n equals to gamma. So, now notice that this b n is a vector now, for the jth component I have done it, now I am looking at all the components put together in a vector and that limit, I am going to call it as gamma, where the jth component is this gamma j.

Now what we dealt with this b ns here, what we can do is, whatever the way we did for this b ns, the c ns almost are the same except for one delayed version. You could repeat this argument and also show that actually, limit as n goes to infinity, the c n of n is also gamma. So, do verify this, like you have to just again go back and write, a s is equals to this format if you just plug in here, we are going to get a into here.

So, the whole analysis will remain the same except for the fact that this P i j s will be replaced by what? S minus 1. So, because this you will see let you will also end up with the same limit in this case because we are only looking in the asymptotic region here. Now, how does this help? So, what is the relation between now c n and b n, I know in the limit, they go to the same value. What is the relation between b n and c n?

(Refer Slide Time: 40:07)



Is that true that b n and they are just one step delayed? b n is one step delayed version of here. So, I can I write b n to b, c of n and P, this is one step delayed version of this. So, I take that and add the next step multiply that is means I am going to the next time by multiplying this transition probability matrix, so that is what I am going to get b n as. Now, what about this gamma you got? This is a vector? Let us take that as my Pi. All I want to show is there exist a Pi such that Pi equals to Pi P. Let us take this as my Pi, set Pi equals to gamma.

So, now in this relation this holds for any n, let n goes to infinity on both sides. What is this? This is gamma and this is gamma P, or I have said this gamma to be Pi and basically I have showing you that there exist and now, we have also guaranteed that this gamma is such that all the components in this gamma are positive. So, what we have is? I have a solution Pi which is, so have shown that if I start assuming that my DTMC is positive recurrent, there exist a solution such that Pi equals to by Pi P and where all the components of this Pi are themselves positive.

Then the next question is this Pi a probability vector? That is what we said? When I stated the theorem, we said that there exist a probability function on S such that Pi equals to Pi P. Now, it is just a probability vector. Why?

Student: (())(42:35)

Professor: So, they will add up to 1, how you know gamma is are less than 1? Because gamma is this ratio, this limit. We know that each of the terms here is going to be less than 1, this is because this summation is or end terms is going to be at most n, you are dividing it by n. So, this every term means for each 1 is going to be less than 1.

So, in that sense, we can have that gamma is are less than 1, that is fine. But, why it is a probability vector? So, these are the issues now, how to ensure that this Pi is, now how to argue that this Pi is indeed a probability vector, then how do you argue that this is indeed unique? When I make it a probability vector that this is going to be unique, how can I do that? So, is that fine?

Student: (())(43:33)

Professor: So, if I take this gamma j to the limit of this, this is going to be unique, in that sense is this gamma j is going to be unique, that is fine. Does that prove uniqueness? No? Yes? In a way yes, because I know that limits are always going to be unique, for the, if I have a limit, the limit if I have a sequence its limit is going to be unique. What here basically I had a sequence here, which is basically deterministic sequence here, which is defined in terms of your P i js. So, this gamma j is our limit, so in that way, uniqueness is coming for granted for us.

Now, the question is why is this a probability vector? So, for that what we need to argue is that instead of this gamma is the way we have, we can argue that we can take the normalized versions. What I mean? You can take this gamma j, whatever vector we have, add them and then divide each of these gamma j by that quantity by that summation. So, in that way it is already probability vector.

So, instead of this gamma j if I am going to look at gamma j by summation gamma j, and now look at this vector. So, gamma i let us call this and if I look at this vector, this is going to be probability vector. So, that would be that, but the question is than I need this summation to be finite, if the summation happens to be unbounded, then this division does not make sense. So, how to ensure that this the values of gamma is here I have if I add them, they would not be blowing up, they will be still less than 1 less than some quantity.

Let us see if we can quickly argue that. So, what is b i n? So, let me just quickly write the steps. I know that Pi i is going to be greater than 0, Pi i are said to be basically gamma i and I also know that this summation i equals to 1, let us take kK and then if I look at b i of K, this is going to be 1 and then we have to show that, summation i equals to 1, K Pi is less than or equals to i and this is true for all K and now if you let K goes to infinity it will be the also case that this guy, i equals to 1 to K, this is going to be less than or equals to 1.

So, I am going to leave this for yourself to be verify. So, what we want to finally argue is that this summation of this Pi is or summation of the gamma is that we are going to deal with is going to be going to be less or equals to 1 in fact, it is not going to blow up. So, because of that what we can do is in this case, as I said, we can just normalize this gamma is by this summation of the gamma is and then it would going to be still be a solution to these equations.

So, in this case, you just gamma is. So, you just divide both sides gamma i and summation of gamma i, I had just divided both sides by this. So, this is going to be a new vector for me and it is still going to be solution to this equation. So, because of this, we had a Pi which is derived from this gamma which is going to satisfy this equation Pi equals to Pi P and we also, said that that Pi is going to be probability function here and also from this argument it follows that that is going to be unique here. So, please check this yourself that I can add the p is which happens to be strictly less than or equal to 1 that is why I could normalize. So, let us stop here I will.