Discrete – Time Markov Chain and Poisson Processes

Professor. Subhamay Saha
Department of Mathematics,
Indian Institute of Technology, Guwahati

Lecture 19

Lecture: Stationary Distribution III

Hello everyone, welcome to the 19th lecture of the course Discrete-Time Markov Chains and Poisson Processes. So, in the last lecture we saw the statement of this theorem which says that if $\{X_n\}_{n\geq 0}$ is an irreducible Markov Chain then the following statements are equivalent. That means these statements are if one only if. So, what are the statements? The statement one is every state is positive recurrent. Statement 2 is, some state i is positive recurrent and statement 3 is the Markov Chain has a stationary distribution π. Moreover when 3 holds π i equal to $\frac{1}{m_i}$ for all $i \in S$. So, this theorem basically says that if you have an irreducible Markov Chain, then if it has a stationary distribution then it is unique and it is given by $\frac{1}{m_i}$, $\frac{1}{m_i}$ where $m_i = E_i(T_i)$ and also if it has a stationary distribution then all states are positive recurrent. Now there are various implications of this theorem. So, the first is, so if you are given an irreducible Markov Chain and you are told that, all states are positive recurrent, then you know that Markov Chain has a unique stationary distribution and the stationary distribution is given by $\pi_i = \frac{1}{m_i}$ or $\frac{1}{E_i(T_i)}$. That is one way of interpreting this theorem. Now for finite state Markov Chains, you know that there always exist at least one stationary distribution. We have already seen this result. So, if you are given a finite state irreducible Markov chain, now since it is a reducible and finite state. So, you know that there exists at least one stationary distribution. That is because just it is a finite state Markov Chain, now since it is irreducible by this theorem you know that the stationary distribution is unique and also it tells you that all states are positive recurrent and the stationary distribution is given by $\frac{1}{E_i(T_i)}$. So, if you are given a finite state irreducible Markov Chain, then all states are positive recurrent. Remember again, we saw is another result that if you have a finite state Markov Chain, then it has at least one recurrent state. So, if it is irreducible that tells you that all states should be recurrent. If it is a finite state Markov Chain, we saw the result that it should have at least one recurrent state. It is not possible that all states are transient. Now if in addition you are also told that the finite state Markov Chain is irreducible since this transience, recurrence are class properties, so that will tell you that if you have an irreducible finite state Markov Chain then all states are recurrent. But now via this theorem we get something more stronger, we get that if you have an irreducible finite state Markov Chain then all states should be positive recurrent but remember finite state is important because when finite state you know that there exists at least one stationary distribution and then you can use this theorem. So, if you are given an irreducible finite state Markov Chain, then all states are positive recurrent and it has a unique stationary distribution given by $\frac{1}{E_i(T_i)}$. And if it is not finite state, if it is any irreducible Markov Chain then you are the existence of a stationary distribution is not guaranteed but if it exists then it has to be unique and all states has to be positive recurrent. So, these are various implications of this theorem, so it is a very very important theorem. Now we will move on to proof of this theorem. Now so in order to show that all these statements are equivalent what we will show is 1 implies 2, 2 implies 3, 3 implies 1. So, that will complete the cycle and hence all the statements will be equivalent. So, we first start with 1 implies 2 but that is obvious. Because 1 says every state i is positive recurrent and 2 says some state i is positive recurrence. So, 2 is a much weaker statement. So, if every state is positive recurrent then obviously some state is positive recurrent. So, 1 implies 2 is obvious because 2 is a weaker statement as compared to 1. So, 1 implies 2 is obvious.

Now 2 implies 3, so what we start with? We start with that some state i is positive recurrent, now if i is positive recurrent then it is definitely recurrent. So, positive recurrent means it is recurrent plus, so it is have something more that $E_i(T_i)$ is also finite. So, positive recurrence if a state has to be positive recurrent it has to be first recurrent. So, also you are given that the chain is irreducible so that tells you that the chain is recurrent or in other words every state in this Markov Chain is recurrent. Now you we have seen this theorem in the beginning of this start of this module on stationary distribution that if you have an irreducible recurrent Markov Chain this γ^i which was defined in terms of the expected number of visits is an invariant measure. So, we saw this as a theorem. But now what is γ^i, now recall just the definition of γ^i so what was γ^i? So, what was γ^i_j, it was sum over k is running from 0 to $T_i - 1$, $E_i(\delta_j(X_k))$. So, this is basically the expected number of visits to state j between until the first passage time to state i. Now if I sum it over all j, now if I sum this overall j then in between the chain must be in some state. So, it is what so basically what I am doing is, I am doing it sum over all j. When I am doing it sum overall j and then if I do k running from 0 to $T_i - 1$ this thing E_i then now see, in between so it is when you are doing γ^i_j you are looking at the number of visits to state j between 0 to $T_i - 1$ but now when you sum overall j in the state space that means this sum should be what, because see this δ, each X_k will be equal to some j in the state space. So, when you sum over all j, this sum will be basically equal to, so like this $\delta_j(X_k)$ will be equal to 1 for some j.And you are summing over all j, so finally this sum will be nothing but just $E_i(T_i)$ which is just equal to m_i. So, if you sum γ^i_j over all j that is equal to m_i. That is just from the definition and since you know that i is positive recurrent you know that $E_i(T_i)$ is finite, so m_i is finite. Now, so again remember what the difference between invariant measure and invariant distribution? So, invariant measure so the difference is, in case of an invariant distribution, the sum of the entries should be equal to 1. So, now for an invariant measure if I know that the sum of the entries is finite, then if I just divide by that sum if I divide each entry by the sum now that will become a distribution because now the sum will become 1 because what I am doing? I am taking an invariant distribution for which I know that the sum of the entries is finite. Now if I divide by the sum that means what is basically called normalization. So, if I divide by the sum, now if I look at this new invariant measure, it is actually an invariant distribution because I have divided each entry by the sum. So, if I look at this new π_j which is γ^i_j over m_i where m_i is the sum, then this gives an invariant distribution. So, starting with a positive recurrent state, we have shown the existence of an invariant distribution which we get in this way so since it is invariant sorry, since this is irreducible and recurrent, we know that this gamma and i is positive recurrent, so we know that this γ^i is an invariant measure now since i is further not just recurrent. So, if it is just recurrent and irreducible then we know that γ^i is a stationary measure, but we know something more here that i is not just recurrent but positive recurrent. So, we now know that the sum of the entries of this invariant measure is finite, so we divide by that sum and make it an invariant distribution and thus we have shown the existence of an invariant distribution.

So, we have shown that 2 implies 3. Now we will show 3 implies 1. If the Markov Chain has a stationary distribution π, then every state is positive recurrent. So, we are now trying to show 3 implies 1. So, consider any state k we need to show that k is positive recurrent. Now we are assuming that the Markov Chain has a stationary distribution π, so since this is a stationary distribution the sum of the entries is equal to 1. Now if the sum of the entries is equal to 1, there exists an i such that $\pi_i > 0$. Because if the sum has to be 1 everything cannot be 0. So, π_i is strictly greater than for there exist an i such that $\pi_i > 0$. Now also recall, although we do not need it here but recall that we have shown that if it is an irreducible Markov Chain then if one state is and if you have a stationary measure, so that in that case either the stationary measure will have all the entries 0 or all the entries positive. You will see, the similar kind of argument we will use here, since this sum is 1 there should exist at least one i such that $\pi_i > 0$. Now by irreducibility there exists an n such that $p_{ik}^{(n)} > 0$. Remember, so we have started with some k and why this is true? Because this is irreducible, so you can go from i to k in finite number of steps with positive probability which means there exists an n such that $p_{ik}^{(n)} > 0$. But now if I take π_k, π_k again is, you know this property of π_k that π_k sum over $j \in S$, $\pi_j p_{jk}^{(n)}$. So, this is not exactly the property in the definition but what we get that $\pi P^n = \pi$. So, this is also true if π is a stationary measure or so here is a stationary distribution and stationary distribution is obviously a stationary measure so this thing is true but now you know, this obviously, this sum will be strictly greater than 0 why, that is because one of the terms here is $\pi_i P$. So, one of the terms here is $\pi_i p_{ik}^{(n)}$ correct. But now that since $\pi_i > 0$, $p_{ik}^{(n)} > 0$ so we get that this is greater than 0 so if one term is strictly greater than 0 and these are all non-negative things. So, the sum is strictly greater than 0. So, what we got is $\pi_k > 0$. So, this is basically the same argument which we used to show that, if the Markov Chain is irreducible, then if you are given a stationary distribution which is not 0 that means not all entries are 0 then all entries should be positive.

So, if you have a reducible Markov Chain and if you are given a stationary measure either all the entries will be 0 or all the entries will be positive. So, but here since we are looking at a stationary distribution. So, it is not possible that all the entries are 0 because if all the entries are 0 it will not sum up to 1. So, now that tells you that all the entries should be positive. So, this is the same proof which we saw earlier as well. So, what we get is $\pi_k > 0$. Now we define λ_i to be π_i over π_k. Now why am I doing this? Now then λ is, because see, π was a stationary distribution so in particular it is a stationary measure and if you multiply a stationary measure with a constant, it remains a stationary measure. This I have already told you before, so if π is a stationary measure then if you multiply with some c real number, some constant then that is also a stationary measure. Now this λ_i so since we are multiplying with so π_i was a stationary measure it was something more. It was a stationary distribution but never mind so π_i over π_k is again an invariant measure or a stationary measure whatever, one and the same thing but when I am doing this π_i over π_k then $\lambda_k = 1$, because what is λ_k? λ_k will be π_k over π_k which is equal to 1, but now what we have this is an irreducible Markov Chain and λ is an invariant measure or it is a stationary measure with $\lambda_k = 1$. So, now we saw in the previous theorem that in that case λ should be greater than or equal to γ^k, where γ^k is that stationary measure defined in terms of those expected number of visits. So, we saw this in a previous theorem. If it is irreducible and if you have an invariant measure with this property that the k-th component is equal to 1 then $\lambda \geq \gamma^k$ that means $\lambda_i \geq \gamma_i^k$ for all $i \in S$. So, this follows from previous theorem that we saw. So, what we are getting here, so remember we are trying to show that every state is positive recurrent. So, we need to show that, actually once we show that, this so this $\pi_k > 0$ now this $\lambda \geq \gamma^k$.

Now, if I now do this m_k, what is m_k? m_k is γ_i^k. Now this we have already seen, why? So, this is basically the same thing that we saw here that if I sum over all the entries of γ_i^j then it is m_i but in this case we are looking at k, so m_k is equal to this that is just because of the definition of γ_i^k but this γ_i^k but what we saw here that this $\lambda \geq \gamma^k$ and λ is defined in this way so using this what we get that this is less than or equal to this now since each $\gamma_i^k \leq \frac{\pi_i}{\pi_k}$ so this sum should be less than or equal to this sum. But now what is this sum remember π is a stationary distribution, so this is sum over i so this k will actually come out of the sum so this is just basically sum over $\frac{1}{\pi_k} \sum_{i \in S} \pi_i$, but since this is a stationary distribution, this sum is actually equal to 1 so you get that this is equal to $\frac{1}{\pi_k} < \infty$. Why? Because you have already shown that this $\pi_k > 0$. So, m_k remember what was m_k? m_k by definition is $E_k(T_k)$. So, in order to show that k is positive recurrent we need to show that $E_k(T_k)$ is finite and that is precisely what we have shown that is because this in the previous slide we already showed that π_k is strictly greater than 0. So, m_k which is by definition the sum of these $\sum_{i\in S}\gamma_i^k$ but now this $\gamma_i^k \leq \lambda_i$ which we defined in the previous slide. So, this sum is less than or equal to this sum but this sum is nothing but $\frac{1}{\pi_k}$ using the fact that this π_i is a stationary distribution so its sum is equal to 1, so we get that and since $\pi_k > 0$ we get that $m_k < \infty$.

So, thus k is positive recurrent. So, that is what we wanted to show that, if the Markov Chain has a stationary distribution then every state is positive recurrence, so k was a generic state we started with any state k and we showed that k is positive recurrent. Now remember, so that completes the equivalence of these statements but we still need to show this uniqueness of stationary distribution but that is very simple. Why? Now we have already shown that every state is positive recurrent, so in particular that means every state is recurrent. So, now again go back to that same theorem from where we got this so it was if it was just irreducible, we got this greater than or equal to but in the state if you recall the statement of the theorem it said that if actually it is not just irreducible but also recurrent then this stationary measure whose kth component is equal to 1 is actually equal toγ^k. So, now we know that this chain is reducible as well as recurrent so this λ is not greater than or equal to but actually equal to γ^k. Hence now this thing where so here what we got inequality here will get an equality and finally we get that m_k is equal to $\frac{1}{\pi_k}$ which is precisely what we wanted to show. So, that completes the proof of this theorem. So, never mind the proof but the theorem is very very important because it gives us so many implications. But before going to the example, just let me repeat the situation again. So, if you have an irreducible Markov Chain, then if it is finite state then it has a unique stationary distribution. All states are positive recurrent and the stationary distribution π_i is given by $\frac{1}{E_i(T_i)}$. But if it is not finite state, if it is an it is reducible but an infinite state Markov Chain then existence of stationary distribution is not guaranteed. But if the stationary distribution exist then it is unique and all states are positive recurrent. So, for a finite state irreducible Markov Chain you already know that there exists a unique stationary distribution but if it is an infinite state Markov Chain, then one way of showing, so there are these two ways so if you can actually show that there exists a stationary distribution then you know that okay all states are positive recurrent. But if you also can show that all states are positive recurrent then also you can claim that it has a unique stationary distribution which is given by $\frac{1}{E_i(T_i)}$. So, the situation is much simpler in for a finite state Markov Chain, so if it is irreducible then there exists a unique stationary distribution but if it is an infinite state Markov Chain there may or may not exist a stationary distribution but if it exists it is unique and all states are positive recurrent. Similarly, if you can show that for an irreducible infinite state Markov Chain all states are positive recurrent then also you can just claim that the Markov Chain has a unique stationary distribution. So, for an infinite state Markov Chain, if you have to just claim the existence of a stationary distribution then either you need to show by like straight from the definition of invariant distribution or stationary distribution that you can find a solution to those set of equations or if you can show that every state is positive recurrent. Again, you do not need to show every state with positive recurrent because it is irreducible if you can show one state is positive recurrent then you know that every state is positive recurrent also. Again, that follows from the previous theorem. So, you know if you are given an irreducible infinite state Markov Chain and if you can show that at least one state is positive recurrent, then all states are positive recurrent and it has a unique stationary distribution given by $\frac{1}{E_i(T_i)}$. Now, till now for we have not given any example of an infinite state irreducible mark of chain where a stationary distribution exists. Remember the example we have given are all like where stationary distribution does not exist. So, we have given examples where stationary distribution does not exist. Now we are going to see an example where of an irreducible infinite state Markov Chain where stationary distribution do exist.

Now let us look at that example, so what is that so again this is the Markov Chain with state space $S = \{0, 1, 2, \ldots\}$ and transition probabilities. So, you can go from 0 to 1 with probability 1 and then for all $i \geq 1$ you go from i to $i+1$ with probability $\frac{1}{2}$ or you just return to 0. So, how is the chain? So, the states are 0, 1, 2, so 0, 1, 2, 3, 4, so from 0 to 1 you go with probability 1. Now from 1 you either go to 2 with probability $\frac{1}{2}$ or you go to 0 with probability $\frac{1}{2}$. Similarly, from 2 you either go to 3 with probability $\frac{1}{2}$ or you actually return to 0 with probability $\frac{1}{2}$. So, that is how the Markov Chain is, so it is an infinite state Markov Chain and again it is easy to see that it is irreducible you can go from any state to any state. So, you go from 0 to 1 with probability 1 and for any $i \geq 1$ with probability $\frac{1}{2}$ either you move to the right so you go from i to i plus 1 or with probability $\frac{1}{2}$ you return back to 0. That is the transition mechanism of the Markov Chain. Now, again as I said, it is easy to see if you just look at the transition diagram. It is easy to see that this is an irreducible Markov Chain but it is an infinite state Markov Chain. Now let π be a stationary distribution. Then first of all, now again remember, so I said like how do you write the equations for stationary distribution. So, in the expression for π_i the sum, so what is basically π_j? π_j is sum over i, $\pi_i p_{ij}$. So, the sum is over all those states from where you can go to j because if you cannot go from i to j then this π_j will be 0 and hence it will not contribute to the sum. So, the sum is over all those states from where you can go to j. Now if I look at π_0, you can go to 0 from any state, not from any state from 1, 2 because from each state you go to 0 with probability half. So, the equation for π_0 is this. Now, if you look at π_1 you can go to π_1 only from 0 and that is with probability 1. So, you go from 0 to 1 with probability 1 so for π_1 the equation looks like this. Now π_2, from π_2 you can go from 1 only, you go from 1 to 2 with probability half. So, $\pi_2 = \frac{\pi_1}{2}$, there is no other state from where you can go to 2. So, $\pi_2 = \frac{\pi_1}{2}$ but $\pi_1 = \pi_0$ so this becomes $\pi_2 = \frac{\pi_0}{2}$. Similarly, π_3 you can only go from π in state 3 you can only go from state 2, so $\pi_3 = \frac{\pi_2}{2}$ but $\pi_2 = \frac{\pi_0}{2}$ so you get $\pi_3 = \frac{\pi_0}{2^2}$. So, in this way if you proceed you will actually get that $\pi_n = \frac{\pi_0}{2^{n-1}}$, for $n \geq 1$. Now, since it is a stationary distribution, we need that this sum should be equal to 1. So, now if I sum it over so π_0 now π_1 is π 0, π_2 is $\frac{\pi_0}{2}$, π_3 is $\frac{\pi_0}{2^2}$ now this should be 1 but now I take this first π_0 out. Now if I look at this part of the sum then this is nothing but a geometric series which will add up to 2 so you get $\pi_0(1+2) = 1$ or which tells you that $\pi_0 = \frac{1}{3}$. Now once you have this you know, you now get the state again since this is an irreducible Markov Chain, then obviously it has a unique stationary distribution and so the unique stationary distribution is now given by this because now you know what π_0 is and remaining you know in terms of π_0, π_1 is π_0, π_2 is $\frac{\pi_0}{2}$, π_3 is $\frac{\pi_0}{2^2}$, similarly $\pi_n = \frac{\pi_0}{2^{n-1}}$, so you get this as the unique stationary distribution. Now, if we use that theorem we also can tell that each state of this Markov Chain is positive recurrent. And also you can say for example, if I ask you what is E_i or say $E_0(T_0)$ so starting from 0 what is the expected time to return to 0 then that should be equal to 1, that should be equal to 3. Why? Because π_0 is equal to 1 over, sorry, let me write this little clearly. So, $\pi_0 = \frac{1}{E_0(T_0)}$ and that $\pi_0 = \frac{1}{3}$, that tells you $E_0(T_0) = 3$. Similarly, you can find for other i,$E_i(T_i)$. So, this stationary distribution also give you those information that say if you consider the set 0, what is the expected time to return to 0 it is equal to 3. So, you get all those information from the stationary distribution using the theorem which we proved today, in today's lecture. So, this is an example of an irreducible infinite state Markov Chain which has an unique stationary distribution and we have explicitly calculated what that unique stationary distribution is. So, till now we have seen now all kinds of examples, we have seen example of a finite state Markov Chain which has a unique stationary distribution. We have seen example of a finite state Markov Chain which has infinitely many stationary distributions in that case it was not an irreducible Markov Chain, you can go back to that example and check because remember the matrix there was, it was the identity matrix

$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$. So, basically it had two states and both states are absorbing states so it had two classes. So, it is not irreducible. And then we saw examples of infinite irreducible infinite state Markov Chains which does not have a stationary distribution and then finally we saw here an example of an irreducible infinite state Markov Chain which has a unique stationary distribution. So, if it is irreducible for finite it has a unique stationary distribution. If it is infinite state either it will have no stationary distribution like we saw for say simple symmetric random walk or it will have a unique stationary distribution like we saw for this particular example. So, we will stop here today. Thank you all.