NPTEL

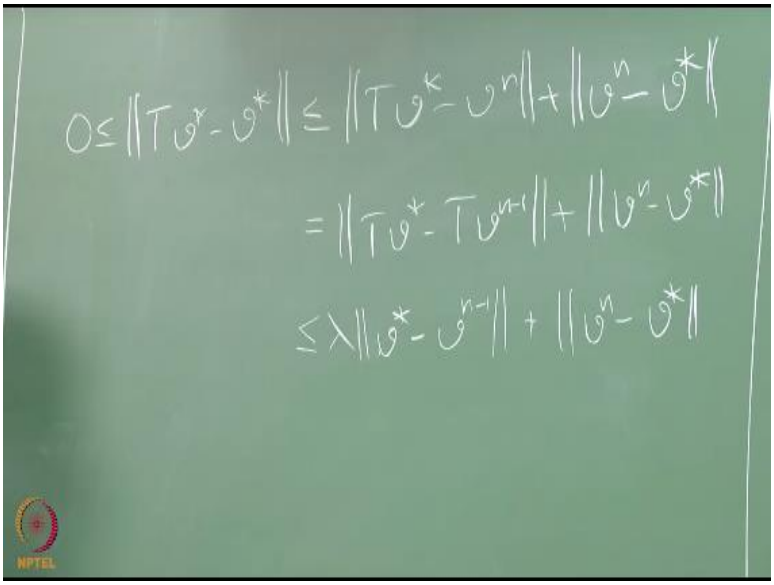
NPTEL ONLINE COURSE

REINFORCEMENT LEARNING

Convergence Proof

Prof. Balaraman Ravindran
Department of Computer Science and Engineering
Indian Institute of Technology Madras.

(Refer Slide Time: 00:18)

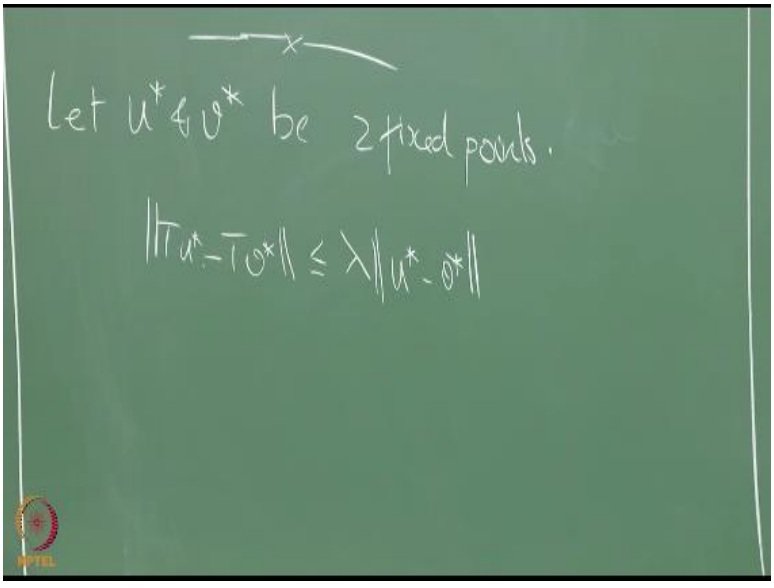


Right so this has to be true if we star is truly a fixed point right something they star is truly a fixed point then this will be 0 so there should be something greater than zero right so because of the fact of this being a norm there is no case of less than zero so this is finest all of this is the thing so now what we will do is we will use our friend the triangle inequality again like I say it is very simple proof right all we are doing is just using the triangle inequality again and again I souse a triangle in you call it again so what we get so what do we know we know that V^n is Cauchy so we know that V^n is Cauchy.

So this sequence is going to this thing right V* - V^n - 1 is going to go to zero I mean that is my definition of Cauchy sequence V* - V^n - 1 will go to 0 likewise V^n – V* will also go to zero by virtue of the fact that V^n is Cauchy that is a definition we had earlier it so that is going to converge right so it will go to 0 no V* is a convergent point right so we know that V* because of the being Barack you know that v* exists in u right so that is the thing we had right limit as n tends to infinity V^n will go to V* that is what we mean by saying it is convergent weight limit as n tends to infinity V^n is going to V*.

So limit as n tends to infinity V* – V^n will go to 0 limit as n tends to infinity that is the crucial point here right it would not be doing this forever and if the difference is going to be something small when he goes far away that means is not convergent that is the assumption we have made see what I am trying to prove is that what I am trying to prove is that.

(Refer Slide Time: 03:39)



TV* equal to V* is a fixed point that V* is a fixed point of T is what I am trying to prove right the fact that V^n is Cauchy right when she tells me that if I applied T or infinitely often enough I will anyway converge to V* when we reach V*. What I am trying to show is once is v* you do not go anywhere you just stop there under T that make sense so the sequence I am generating V^n

VN +1 VN + 2 and so on so forth I am showing that it is Cauchy right so that means if I applied t infinitely often enough I will eventually reach infinitely often infinite should be enough so if I apply the infinitely often I will reach v*.

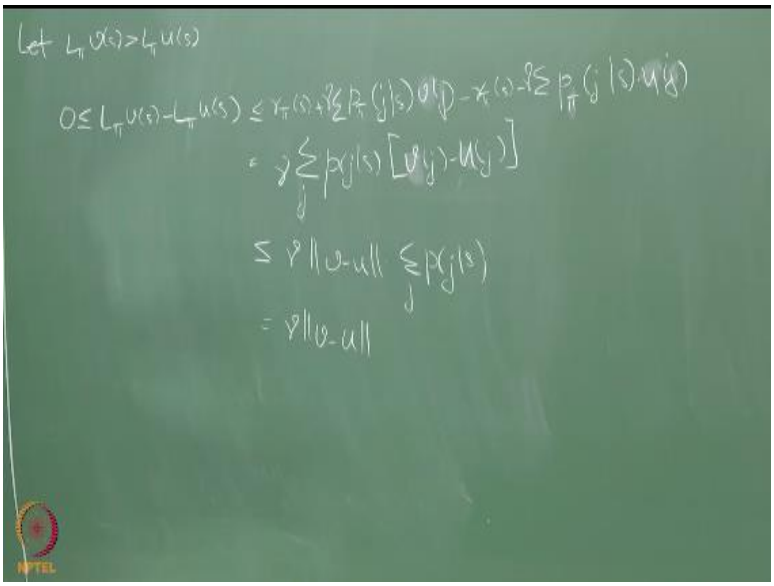
So that process is that what we are trying to say is okay v* is actually a fixed point once you reach me star you are going nowhere take a little bit n tends to infinity for we want one what would you have gotten like in n = 1 and then m = 1 tending to infinity take a little bit yeah what would have got it just becomes zero okay and then you show that it is the successive differences become 0 that is the $V^n + 1$ is can be written as $TV^n - V^n$ will be less than equal to 0 yeah.

So essentially you show that asset tends to infinity it will become 0 you become it will become 0 but you have to show that we V* reaches V* to stop that that is all we are showing here when it reaches V* it would not go anywhere beyond V* that is it right so we are just showing that it will stop at V* that is basically what we are trying to show here and since this is trivial in fact we can actually show it that way also as limit n tends to infinity okay now the uniqueness from the contraction property.

So U* and in V* or two different fixed 0.40 from the contraction property we have TU* - TV* should be less than or equal to λ times u* - V* norm of U* - v* but since they are six points the right hand left hand side is also U* - V* so I'm asking Q for U* - V* should be less than or equal to λ times U* - V* for λ non 0 λ that is satisfied only for only if V* = U* e start to stop well sorry if λ is not 1 then it satisfied for yeah so even if λ is zero it has to be that U* = V* okay.

This does not matter anyway just using the contraction property just this is sufficient to show that U* has to be equal to V* rights as long as λ is not one and anyway we have excluded λ1 here right so as long as λ is not equal to 1 this implies that.

(Refer Slide Time: 08:30)



Great so what we have done so far is shown the at the benefits point theorem is correct that we need to show that L π is a contraction so if you so L π is a contraction you get everything else for free right so we need to show α a is a contraction so let us do that and I will stop after that ah I do not people stop me what okay when I do B - U V- okay so far it just algebra okay so now what I am going to say is well I do not care what j I am considering the notation so of U and V are vectors but US and Vs are scalars but all the contraction mapping results are four vector norms it so mean all the contraction mapping is so far vector norms here.

So do element wise do they correspond to that I mean the maximum of that particular vector has to be lesser than the maximum of another vector but each element wise they need not be lesser element wise they need not be lesser no that is I not needed yes only the max noun ways yeah the max of this should be less than that expression is again for a scalar expression again for a scalar expression whatever I am writing so far is a scalar I will come to the vector form of this case so far I am writing scalar now so the tricky thing you know I am going to say so I am summing over all j right so I am going to look at this expression.

So what I am going to serve is this expression right will always be bounded by B – U like the max norm alright so this is a pretty loose bond right so I am taking each element here I am looking at this that each element of this will be bounded at most by the maximum right so it cannot be any greater than the max norm difference therefore I will say that this expression is lesser than or equal to γ times the max norm into that right so that anyway goes to zero so this is essentially equal to okay.

So at every point this will satisfy that this is one way so go the other way me let into the other way easily enough right you will end up with the same thing norm of V – U the same of Norm of U- V so we will end up with the same expression right so essentially I can say that so I am assuming that you guys will finish the other way around okay.

So what is otherwise is when L π V(S) is less than L π V (S) okay so greater than less than so I am just assuming you guys will do that right putting these things together gives us what right so what can infer from this point place it is being drawn closer right by γ so the max difference should also have come down point where is being drawn closer even if the max difference is at different points right it should have shrunk all right so essentially that so this essentially means that implies that L π is a contraction.

So next thing to do is to start the same thing for the optimality equation right so we will go here in the next class and then I have to also talk about so assuming that at every point at every point the functions are being drawn closer so I am saying that see L π VS- L π us is closer right then the max difference that you had in V –U right and this holds for all state S right and that the otherwise the right-hand side is infinity norm left-hand side is just absolute value so we also have to show algorithms for solving this right and so we will do all of that in the next class possibly it might spill over into the subsequent class as well.

IIT Madras Production

Funded by

Department of Higher Education

Ministry of Human Resource Development

Government of India

www.nptel.ac.in