

# NPTEL

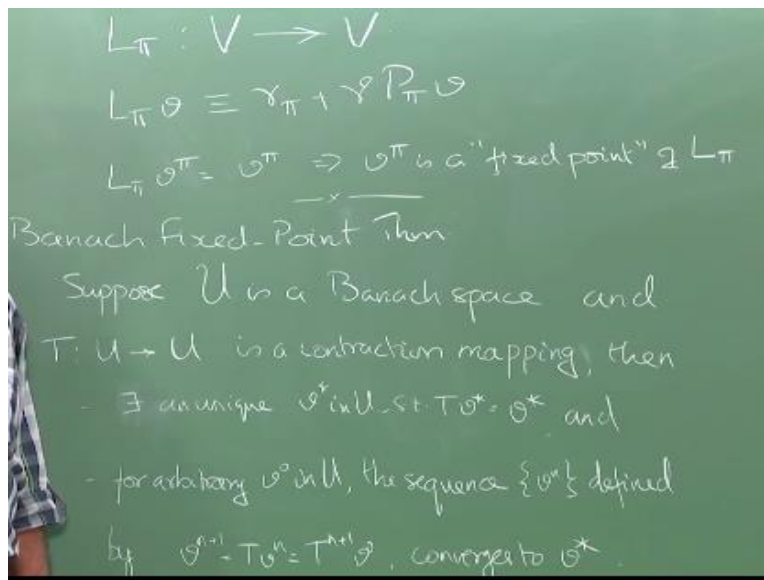
## NPTEL ONLINE CERTIFICATION COURSE

### REINFORCEMENT LEARNING

#### Banach Fixed Point Theorem

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

(Refer Slide Time: 00:16)



So I am going to define an operator a linear operator called  $L_{\pi}$ , right so  $L_{\pi}$  is something that is going to take a vector from the space  $V$  where capital  $v$  is the space of all value functions right so capital  $v$  is a complete norm vector space of all value functions right so I am going I am going to define an operator  $L_{\pi}$  that takes an element in  $v$  to another element in  $B$  right so how would define  $L_{\pi}$ ?

$L_{\pi}$  operating on  $V$  will give me back a function that is right so  $v$  some element in capital  $V$  which is some value function okay or rather some element in capital  $v$  okay, so what is the difference between saying there is some value function and then saying some element in capital

$v$  and he says a value function I am kind of implicitly saying that there is a policy for which that is the expected return right.

But when I say it is a point in the space  $V$  that need not necessarily be true I mean it is just some point in the  $s$  dimensional space that need not necessarily exist in sister policy corresponding to that function okay, so I am just saying is some function  $V$  right the dents in exists in that space capital  $V$ , between what other space  $V^2V$  does that make sense means confusing to people less everybody's on board right good right.

So what does my bellman equation tell me, if you think about it the bellman equation tells me right that is what my bellman equations essentially is saying that I take this point  $v \in V$  in the space capital  $v$  I apply the operator  $L$  on it whatever it gives me back will be  $Lv$  right so likewise if I can define an operator for the optimality equation it will tell me that  $v^*$  is a point in space  $V$  such that if I apply the operator let us call it  $T$  if I apply  $T$  on it will give me back  $v^*$  so this implies that  $v^*$  is what is known as a fixed point of  $T$  right side so if you think about it right so normally I take some vector  $v \in V$  I apply  $T$  on it okay.

It will go to some other point then I will apply  $T$  on it will go to some other point then I will apply  $T$  on it will keep moving around right in space right if I apply  $T$  on it and it stays there then I call it a fixed point because I am using the defining the operator through  $R$  and  $P$  which themselves are defined through  $\Pi$  right, I mean if you looked at if you remember the definition Frenchman no it is not because the mapping is  $T$  I mean mapping uses  $R$  and  $P$  is integral components of it right.

It is not independent so the other one is independent that there we will drop the dependence the optimality equation has no dependence on  $\Pi$  there will rock the dependence on fine but the bellman equation has a  $\Pi$  in it great so introduce you to something called the banner fixed point theorem which is very basic I mean believe it or not whatever we are doing is very fundamental stuff right.

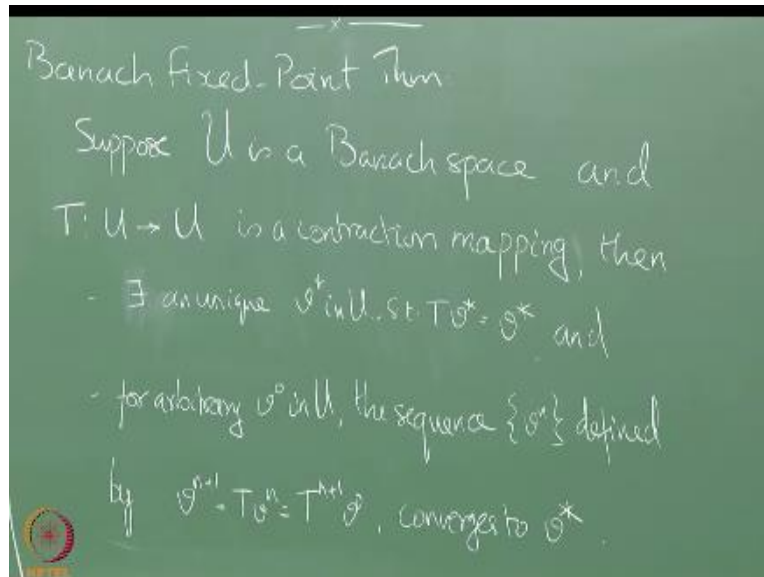
Very minimal math in terms of trying to prove convergence and things like that because I have restricted you to finite MD piece okay, so the more hairy the more interesting math comes when you start operating in the infinite space state spaces right where I mean you cannot write down simple results okay you have to have a lot of qualifications if your MDP says that A satisfies this and that and that and A satisfies this and that and that.

Then something will hold right but here you can say some ways we can write very easy results so I am going to use something called the Banach fixed point theorem okay which will which makes our life a lot easier, right. I have sounds like a very fancy term right Banach space so what is a Banach space we already know what a Banach space complete norm vector space is called a Banach space I just do not get put off by terminology and we understand this right.

So we know what a Banach space is a complete norm vector space is a Banach space, so one thing which I forgot I will introduce it let this so some mapping  $T$  that maps points in you to another point in you right we assume it is a contraction mapping I will tell you what is a contraction mapping in a minute okay but you can think in kind of case water contraction mapping is right it takes two points in space then I take you and be I take  $T(u)$  and I take  $T(v)$  right.

So  $T(u)$  and  $T(v)$   $T$  will be closer to each other than you and we were all right so that is essentially contraction mapping so I will write it down formally in a minute right, then guess what we are.

Refer Slide Time: 08:38)



saying suppose use a Banach space and  $t$  is a contraction mapping essentially things that brings points together then that exists a unique  $v$  star in you says that  $Tv$  star equal to  $v$  Star so that essentially means that  $T$  has a fixed point.

$T$  has a fixed point and it is a unique fixed point okay and the second thing is for arbitrary starting point  $v_0$  if I keep repeatedly applying  $T$  on  $v_0$  right I will converge to be star right.

So showing 1 does not imply to just because a fixed point exists right repeatedly applying it did not necessarily converge to the fixed point in fact there are some results for certain algorithms where we can show that that algorithm itself has a fixed point then I mean or rather  $v$  the system of equations you set up itself as a fixed point but the iterative process that you are using for solving it did not necessarily converge to it okay.

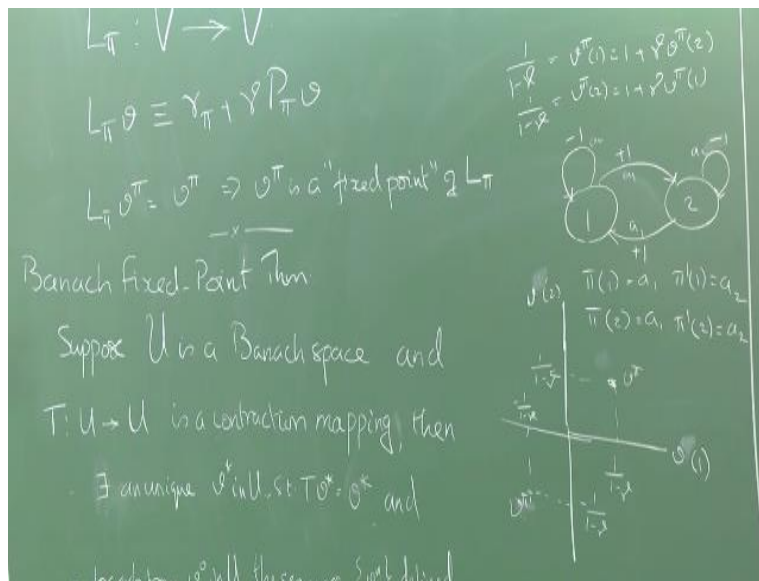
There are cases that will happen but in this case it says turns out that repeatedly applying  $T$  that is essentially what the sequence is at I start off with  $v$  not apply than it once I get  $v_1$  I apply  $t$  on it again I get  $v_2$  I apply  $t$  on I get  $v_3$  and so on so forth and I keep doing this then often enough

great I will end up in v star is how many times is often enough something we said even we look at actual algorithms for solving these problems.

But these are the things that exists a unique v star and this is a couple of things that we need to show one is that the use of Barack space here again well it is a plan it is a point in the league space is eigenvector fine but I came as a point as a nice place the point is that is what we have been saying so what is v  $\Pi$ ,  $v\Pi$  is a vector that the first component is v  $\Pi(1)$  v $\Pi(2)$  v $\Pi(3)$  v $\Pi(4)$

This is a point I mean in S dimensional vector space right cannot you see let us think of a case where there are two states alone let us state 1 state 2 right I define some policy let us I am just going.

(Refer Slide Time: 11:17)



to cook up something let us look at how it works one I have too some arbitrary thing have written and FDP for UK right so I take action one from state 1 I will go to state 2 and get a reward of 1 I will take action one in state 2 I will get go to state 1 to the robot of one I take action to in either state 1 or state to I am going to stay there with a robot of minus one, okay.

Now let me define a policy so what is the policy  $\pi$  of one we see a 1 and  $\Pi$  of two is a to determine am sorry a one I of 20 let us just define a 1 a 1 for both right so what is the value of state 100 and you can do it easily enough but for some value of  $\gamma$  you can solve it or you can just write the summation write it just 1 plus  $\gamma$  plus  $\gamma^2$  plus  $\gamma^3$  I mean you can solve the system of equation or you can just tell me what is the limit of the summation.

Somebody tell me now if I think of this in a two dimensional space yes I have a two dimensional space this dimension is  $v \Pi(1)$  this dimension is  $v \pi$  of two just a point, so this point is whether the servants if I envy of one now we can think of defining a  $\Pi$  prime which is like shape- of them right so for any policy I choose like this right  $\Pi(1)$  and  $\Pi(A1)$  to whatever things I choose I am going to end up with some point in this space right.

So this is what I mean by the space capital  $V$  so this is my space capital  $V$  and each value function is a point in this space right I mean I was hoping people have actually this in their mind when they are nodding with whatever saying earlier right so this is what I mean by saying all of these things which we have talked about right this being a complete norm vector space everything we have talked about is with this picture in mind so okay.

Does it make sense any questions about that of course I just chose deterministic policies to be easy right you could choose to Casting policies you just have to write out this more terms in this equation and this equation will have the sum over  $\pi$  is a right and then you solve it and then you get it right so essentially the tricky part here turns out showing that whatever mapping  $T$  you are considering is a contraction mapping.

so if you can show whatever is the map that you consider in this case we are going to be considering  $\Pi$  if you can show that  $L\Pi$  is a contraction mapping right then you get the rest of it for free yeah so that's it essentially that is what you have to do to show us a contraction it is a contraction mapping and just tell you in a minute I formally write down what a contraction mapping is right.

(Refer Slide Time: 16:12)

$T$  is a contraction.  $\|Tu - Tv\| \leq \lambda \|u - v\|$ ,  $0 \leq \lambda < 1$ ,  $\forall u, v \in U$

Proof:  $\|T^m u - T^m v\| \leq \sum_{k=0}^{m-1} \|T^{n+k+1} u - T^{n+k+1} v\|$

$$= \sum_{k=0}^{m-1} \|T^{n+k} Tu - T^{n+k} Tv\|$$
$$\leq \sum_{k=0}^{m-1} \lambda^{n+k} \|u - v\|$$

$\|x\| = 0 \iff x = 0$   
 $\|ax\| = a\|x\|$   
 $\|x+y\| \leq \|x\| + \|y\|$

So for sure okay it is it moves away so this is what a contraction mapping is so  $U$  and  $B$  where some distance apart after I applied  $t$  on both of them they will be closer than what they were earlier by a factor of at least  $\lambda$  where  $\lambda$  lies between 0 and 1 a 1 excluded so the smaller the value of  $\lambda$  what does it mean the faster the points are motor together the more the contraction is at the closer the value of  $\lambda$  is 1.

The lesser the contractions obviously  $\lambda$  equal to one does not make sense right if  $\lambda$  is 1 then there is no contraction if  $\lambda$  1 is  $T$  identity need not be and it is, now just telling you this the distance have to be preserved right I can swap you and we also just remember that okay, so now let us look at the proof of the Banach explain it so why is it called the Banach fixed point theorem.

Because it applies in Maroc space right so for different kinds of spaces that you are starting off with you have different kinds of fixed point there is a whole family of fixed point theorems okay depending on the underlying assumptions you made so this is a fixed point theorem that we will show in the barrack space it is one of the simplest of the fixed point theorems okay, so if you

have lesser restrictions on the underlying space then the showing fixed points becomes more tricky okay.

So here is the trick that will be using repeatedly I am just taking some iterates in that sequence right I have taken some  $VN$  and some  $VN + N$  plus so  $VN + N$  occurs after  $VN$ . So I am going to use the triangle in equality right so I can pick some arbitrary point let us say in between in the sequence right so we say  $VN + N/2$  I can pick some arbitrary point in the sequence which is  $VN + N/2$  and then i can say this is less than  $VN + N - VN + N/2 + VN + N/2 - VN$  I can write it like this some time.

Being forget about the fight they are coming from a sequence I just think of them as three points in space let us say  $V$  and this some point in that space  $VN + N$  is some other point in that space and this case is a third point in that space right so once I give you three points I can use a triangle inequality right I just chosen the third point to be somewhere in between in that sequence okay so that is all I mean nothing to do with the fact that these three points or the sequence.

I can use a triangle inequality given any three points because I can do it this way right but there is nothing sent no sanctity about doing this in fact I can split this further I can say  $VN + N/2$  and then  $N/4$  I can split that and do that in fact I can do this for every step along the way so I am going to replace it with this right, so basically  $N+1, N+2, N+3$  so it keeps doing this repeatedly applied the triangle inequality to get this right.

So these pick from do to you do to you takes you and then maps another point in you so two different points right so I take two arbitrary points you and be this for all points  $U$  and  $V$  and  $v$  let us any arbitrary point I take I apply  $T$  on  $U$  and I apply  $t$  on  $v$  right so I will get two resulting points right so I will take the distance between them they should be less than the original distance between  $u$  and  $v$ .

So  $t$  is a mapping from  $U$  to  $U$  it is not  $U \times U$  right it is mapping from  $U$  okay, so now we have reduced it to the same  $T$  applying multiple times right so I can apply  $t$  use this multiple times right so what I am going to get using this contraction operator here is a fact was it the second



stage yeah let the operating team because we found is decreasing the distance between the two subsequent it had I apply the operator  $T$ .

By definition  $\|T^{n+K} - T^n\|$  is  $T$  operating on  $V_N + K$  times, if it makes sense now this quantity is independent of  $K$  so I can take it out and sum over  $K$  so what is this sum I mean very hard keep getting calls for a I do not know why like my third call during the class, what can you say about the sequence as  $n$  and  $n$  becomes large it becomes smaller and smaller and smaller write as  $n$  and then becomes larger and larger.

That sequence is going to become smaller and smaller so this is a constant so it does not matter so  $N$  and then become larger and larger this is going to become smaller and smaller so what can you say the sequence  $V$  and this Kashi okay, we will get that anything else we can say we assume that is a Banach space right that is what I said Banach space fixed point theorems are easy we assume it is a Banach space therefore  $V_N$  is convergent we already know that right so what needs to be shown now.

**IIT Madras Production**

Funded by

Department of Higher Education

Ministry of Human Resource Development

Government of India

[www.nptel.ac.in](http://www.nptel.ac.in)

Copyrights Reserved