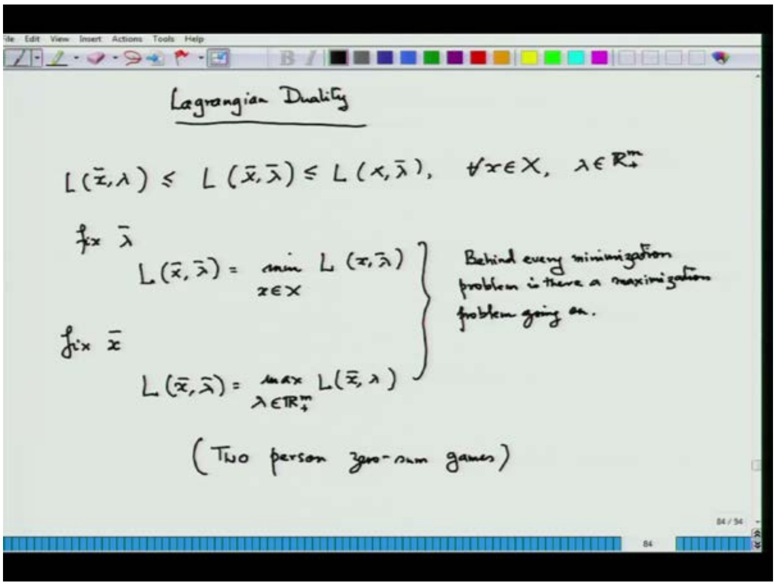
Convex Optimization

Prof. Joydeep Dutta
Department of Mathematics and Statistics
Indian Institute of Technology, Kanpur

Lecture No. # 16

We had spoken about saddle points in the last lecture, where you saw how nicely without much information about the problem, only knowing that the problem is convex and an additional condition like sadder condition holding true. You can get a neat optimality condition express to the Lagrangian function. Now, let us look at the Lagrangian functions slightly more carefully and the saddle point condition.

(Refer Slide Time: 00:45)

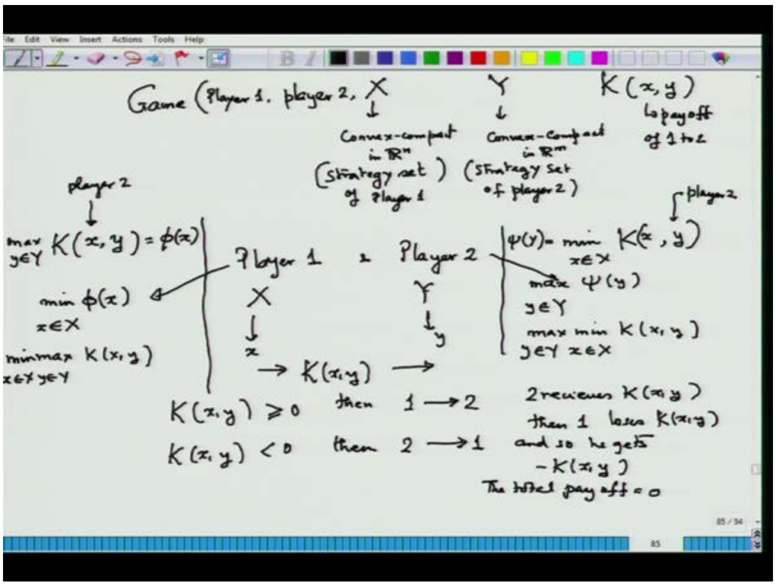


So, the saddle point condition says that if if you give me a minimum, then I can find a lambda bar, if x bar is the minimum of the convex problems such that this occurs for all x in some capital X and lambda element of R m plus. Now, if you look at it there are two aspects of this that if I fix lambda bar then L x bar lambda bar is the minimum value over x of L x lambda bar. Now, fix x bar and L x bar lambda bar is the supremum value, we take it as the maximum value, because x bar lambda bar where would lambda equal to lambda bar their equal.

So, maximum of lambda element of R m plus L x bar lambda; so, here x bar is fixed, a lambda bar is fixed. So, we were actually solving a convex minimization problem. And now here the Lagrangian function which by which we converted an constraint problem into an unconstraint one. Someone constraint, because we have not been able to handle this abstract constraint directly, if f is equal to r n then this problem is truly unconstraint problem. So, if we were actually doing minimization problem, how does a maximization problem always arise, it may, how did how did a maximization problem come to the scene. So, the question is behind every minimization problem is there are a maximization problem going on; that is the question. Now, what is the answer to this?

Let us look at search for answer in a much more interesting scenario; in a more real life realistic scenario of games. So, now we are going to discuss what is called (( )) not really very deeply, we are going to discuss what is called two person zero sum game. Any interaction between two human beings for example, is a game, it could be interaction between two human beings, it could be interaction between two armies fighting each other, it could be interaction between two football teams playing football or two cricket team playing cricket. So, any interaction between conglomerate, group of people or individually between few people such an interaction is called a game and of course, they are rules laid out.

(Refer Slide Time: 04:38)



So, in a two person game, staff of the games are two players which I call as player 1 and player 2, and everybody has certain strategy to play the game. You know chess players of course have an extensive strategy to play. So, they know the opponent, possibly this is the way the opponent is going to fight me, so this would be my strategy. So, mathematically speaking let there be a convex compact set in R n which is the strategy of player 1, just take it R n for simplicity. So, this is strategy set of player 1. So, all is moves in the game. So, what happens in a game is player 1 makes the move and player 2 makes another move. So, all is moves are recorded, he he has stored in this set colex; strategy set of player 1.

Now, let y be the strategy set of player 2 and let us assume this be convex compact R m. So, this could be mixed strategies you know, so those who know some game theories, so it could be a simplex actually, because it holds the probabilities of the moves. So, you can we are doing it in slightly abstract fashion; the strategy set of player 2. Now, of course, it is a win and loss situation means somebody is winning some body is losing the game. So, how do we write down the game, what what more we need? It is like playing, I mean it is like gambling you know that every giving out lot of this mathematics as come out of gambling.

So, here when x makes the move and y makes the move, and depending on the two moves a player 1 has to play some money to the player 2, if that money is positive then he actually has to pay, if not he returns I mean he gets a similar amount back. So, given a strategy x which is in x and a strategy y of player 2 which is in y, this is the payoff of 1 to 2. So, basically I can write this as player 1, player 2. So, it has a strategy set x, it has strategy set y, if you pushes in x and if you pushes in y then K(x,y) is the amount of money in paid by player 1 to player 2. Now, if K(x,y) is greater than equal to 0 then 1 is paying to 2, if K(x,y) does not equal to is strictly less than 0. 0 is an ambiguous not ambiguous when you can just decide then basically no one pays anything. Then 2 pays 1.

Now, what happens is that if K(x,y) if if 2 receives K(x,y) then 1 loses, the player 1, this 2 and 1 are marking the players and 1 loses K(x,y), and so he gets minus K(x,y). So, the total sum is 0 - total payoff is 0. That is exactly 0 sum game. So, is basically like this, for example, I can device a game standing here, so I play a game with you, a game is that I will tell a country (( )) is make a capital city. So, you on you know strategies here strategy set of all capital cities and another set is the country which I have. Now, I make a move country and if you have a capital city. Now, if the country and capital city matches then I pay you 1 rupee, and if the capital city and country does not match you pay me back 1 rupee. This is an example of a two persons 0 sum game on.

Now, what what can we look into from here; what can we conclude from here. So, how do the player knows that what is actually the end result. That is how do I know that there would be a strategy which would possibly with best for both, there would be a strategy which is optimal for both. So, if I move off from that strategy both are not in a good position; a one is improved position and other is not. So, what is the strategy which tells me that… This strategy is the best strategy to take when… So, let us see what happens what does player 1 do. Player 1 let us see let us look at let us look at an argument player 1 will take. See he has to pay money. So, he knows that suppose this guy - player 2 has given input y, he has played y, now he knows that given this y I have all elements x in x capital X that my disposal. So, I can plug in any x I want corresponding the every x he has to pay K(x,y).

Now, he would like to know what… So, he given me once the y is given to me, I know what player 2 has given, he wants to know what is the maximum loss he can make. So, this is the maximum amount of money he has to give - the maximum value of K(x,y). This is the maximum loss, because because of I assume the continuity and all nice property, so they will be in x naught in x which this minimum and maximum would actually occur. So, if we choose by mistake that is the strategy that would be his maximum loss. So, but nobody wants sorry I am making a little bit of mistake, I am telling from point of view player 2, I should tell from the point of view of player 1. Just just just rub it. It happens you know, you get carried away in your thoughts. Let us go back again.

So, as a player 1, I am pushing in a strategy x right and you now know that y can give in any strategy. I have given x, but I do not know what strategy y will play, I was just telling the reversing, I should have written it here. So, I do not know what is the strategy player y will play. He can play any y at his disposure. So, keeps on giving a y. So, how do I know? What would be my loss? How much money I have to pay? So, the maximum money I have to pay, my maximum loss is max of… So, there is a strategy y - small y if that is played by player 2 corresponding to my given x - player 1's x then that would be this would give the maximum loss means, because the function is nice, because of the compactness of capital Y. There exists of y naught for which K(x,y) naught is the maximum value of this.

Suppose for my this particular x, the guy puts y naught, then I lose a large amount of money. So, I want to know what is the maximum money I am actually loosing. If the player 2 is putting in the… If I put the strategy x and player 2 has the right to put any strategy 1's, what is the maximum amount of money I am loosing, this is the maximum amount of money I can lose. If I put the strategy x, so player 1 knows the (( )) puts in a strategy x; this is the maximum amount of money you will lose.

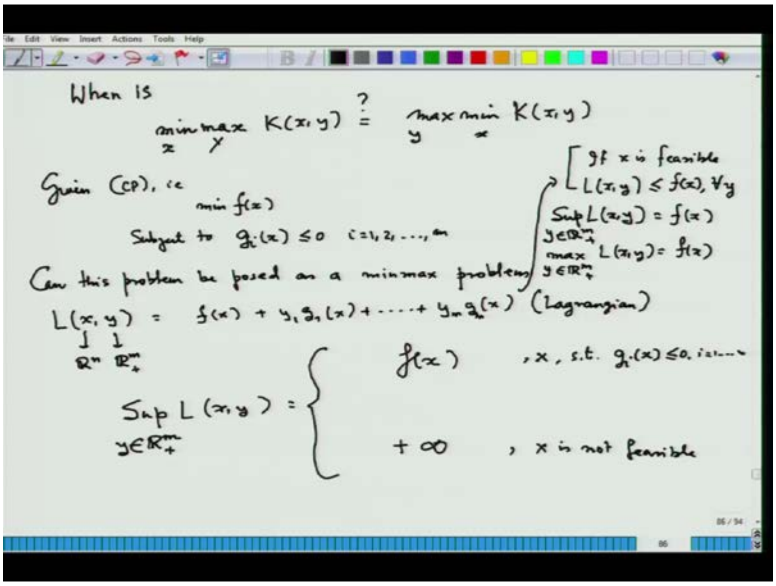
Now, what is this m? He wants to play such a strategy x which will minimize this loss - minimize the maximum loss. So, this mean minimized over y so which is the function of x, because you change the x, the value of the maximum value will change. So, what player 1 intense to do is to minimize over this, is to minimize this function phi. So, basically player 1's problem is what is called the min max problem - minima over x maxima over y of K(x,y). So, player 1 is actually running a min min problem, this is this final problem. Let us see what player 2 does. Player 2 can put in whatever strategy he likes. Suppose he chooses a strategy y, corresponding to this input y of player 2, he knows that player 1 has at his disposal any strategy from x.

Now, once he has this thing in his disposal, the interesting thing that comes out as a following. That you have x in your… This guy player 1 has any strategy likes, but given his strategies, he might choose such a strategy x for my given y, he might choose an x naught so such that for all possible K(x,y) values for this fixed y that would give me a minimum, that will give me the minimum value that I can get by playing this strategy.

So, he wants to know what is the minimum amount he can make that is his goal; he does not bother about what is the maximum amount he is making, he is not greedy at the beginning. This is the I just want to know that if I play y what is the least amount of money I am going to get. This is the least amount of money your he is going to get; which, because I am your minimizing over y, it is a function of minimizing over x it is a function of y. So, you see now this is the minimum amount I will get, but my aim would be to play such a strategy y, so that I can have as much as I can from that minimum amount. So, I will have, what I will do is max of and you see (( )) so this is absolutely fabulous. So, I what I have is a max min problem.

So, the player 2 is actually playing a maximization way game. Basically his problem is to maximize the function player 1's game is to minimize the function. So, this is exactly what player 1 is doing and this is exactly what player 2 is doing. So, you see even when you are talking about games there are two problems that is going on that if there is a minimization problem there is a maximization problem of course, one of the major problems of game theory is to know when is min max over max over y min over x and max over y and min over x when is this equal; that is the question; that is the central question of game theory.

(Refer Slide Time: 18:07)



If there is a pair x bar y bar or which this to value there are equal then such a pair x bar y bar is called the value of the game. That is the equilibrium means if you deviate from this one of you are in the down side. So, we will not go into discussion of game theory at this moment. So, there is another reverse question I want to ask. Can the usual optimization problem say convex optimization problem, minimizing a function, convex function over some convex inequalities say can be posed as a min max problem; that is the question.

So, then correspondingly we can think that there is some max min problem somewhere associated with it. So, this is what I will call the primer problem, this is what we will be call the dual problem. So, my question is this given c p that is minimize f x, take this simple problem do not bother much about (( )). Can this problem be posed as a min max problem? So, this problem is called a min max problem. Can it be posed as min max problem? In the sense that can you define first sum K(x,y) that is to define a min max problem, there must be function of two variables, two vectors x and y, because you are maximizing our y and minimizing our x.

So, if I give you a min max, give you this particular problem, if I ask you can you pose it has a min max problem, your first question would be how do I prove the min max problem I just have a one variable x here and where is my K(x,y) then you think a bit no no no no no no no no there is something. The something is ok, let me write it like this. L(x,y) as f(x) plus y 1, so I can from this problem I am creating a function of two variables go on. So, where I call this is in R n and this I take in R m plus your strategy set, just like your strategy set. This is your K(x,y), your replace K(x,y) with what is known as this is familiar to us, this is the Lagrangian. Now, it is associated with a convex problem c p, is it is Lagrangian function. So, can the Lagrangian function be used that is the interesting question that comes anyway this is the way naturally people would argue.

Now, if I have been able to associate a a function of two vectors x and y associated with the problem c p. Can this be used to pose c p has a minimum min max problem. Answer surprisingly easiest, but here we will not write min max, but instead we will write in sup, because we do not have compactness here. So, if the engineers do not understand what I am telling by in sup and all the thing do not bother. You can just replace whatever I am writing as in sup in as min max. So, it will be lose in some sense, but it will give you some understanding.

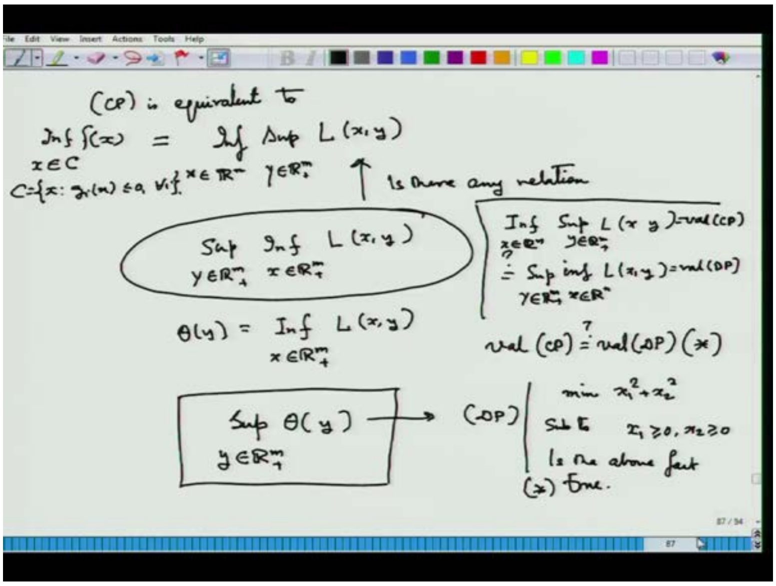
Let us look at the supremum of L(x,y) and y is greater than equal to 0, in the sense that y is in R m plus if you do not like the symbol which remains your real numbers, so I will just go and write it like this. This is exactly this first step; a first step here in K(x,y) that K(x,y) max of K(x,y) this is this is exactly the step and this step is what we are doing at this moment. Now, what does this give me? This give me something gives me something interesting. Now, if I pose that every y 1, y 2, y m are greater than equal to 0, and I take x to be feasible that is x such that if this is what I have that - I have taken x which is feasible then y 1 into g 1 x, y 2 into g 2 x, y m into g m x, all these are less than equal to 0.

So, when this happens let me do the calculation on the side, if x is feasible L(x,y) is less than equal to f(x) for all y. Now, if I had choose an all the y 1, y 2, y m such 0 then

L(x,y) would be exactly equal to f(x). So f(x) is the value taken up here. So, the supremum y in fact the maximum actually of L(x,y) y element of R m plus is equal to f(x), if you have not understood this argument, note that if I choose y is equal to 0 here then L(x,0) is exactly equal to f(x). So, in that sense f(x) is this 2 for all y, so L(x,0) is less than equal to f(x). So f(x) is less than equal to f(x) right. So, which means since f(x) value is one of these values, so this is f this supremum is attained actually. Now, I can also write, so when x is feasible this is this is what it is happening, so this is what I write down.

And now, suppose x is not feasible, if x is not feasible… So, there can be two cases only; x is feasible, x is not feasible so x is not feasible then there must be some i that among the 1, 2, 3, 4, m, m constants that g i x would be strictly greater than 0. Assume the g 1 x for this particular x for which g 1 x is strictly bigger than 0. So, what I will do is I will put the y is equal to 0 for all the rest, and keep on increasing the value of y 1 and I can keep on increasing, keep on increasing, keep on increasing and it will blow up which means what I have is the following, this is exactly our effective function - effective objective function the you know that you can pose optimization problem as extended valued on a optimization unconstraint problems with extended objective function, extended valued objective function that is function that takes both plus that can take plus infinity value also.

(Refer Slide Time: 26:59)



So, you get back this. So, what is my problem actually at the end my problem c p is equivalent to to the infimum of over R n supremum. So, this is my primal problem, the usual the problem c p is now looking like like this is. I have posed it as a min max problem in sup problem, what this exactly what question I have asked you. Now, you might now question going back to the min max problem and max min problem. The remembering what player 2 does; is there any problem which I can set like this, like I can change the position sup of y, if I write a problem like this, because I am just imitating what I have done for player 2 in this case you see, what I have done for player 2 is this problem max then min of the payoff.

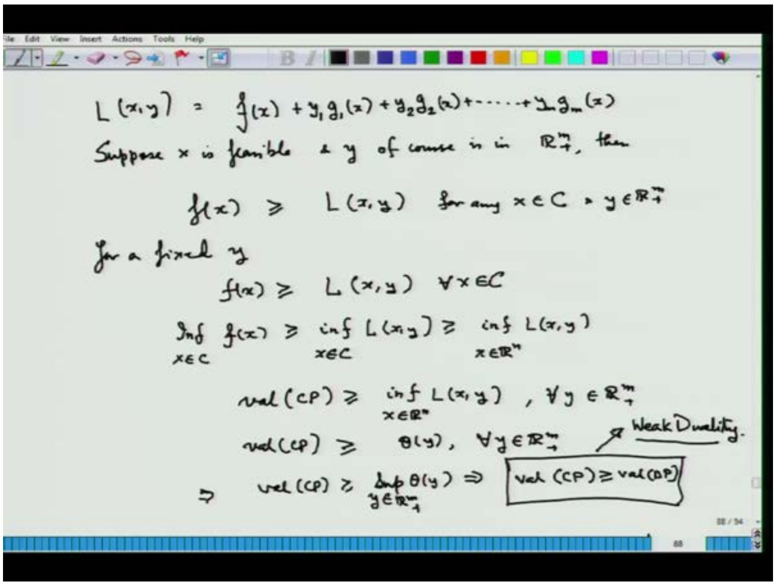
So, suppose I write this problem; is there any relation? That is the important question. Is there any relation? Now, if I breakup this problem a bit then I can write theta of y which is your psi of y here, so theta of y, I can write as infimum. Now, the dual problem is supremum this problem, which I call the dual problem. So, given the primal optimization problem by following the things that we have just learnt from 2 person 0 sum games, I can mark this problem as Dp or the dual problem. What is the relationship between these two problems? Just like the game theory guys did, I can ask this question again. That is if inf of sup… So, this is the minimum value. So, if I want to find the minimum value of f(x) or infimum value of f(x) does not matter, x element of the feasible set c is this.

Now, in sup, look at in sup here, just like the game theory, can I ask this question. Is in sup L(x,y) this is our y, this is y element of R m plus and this is x element of R n is this then equal to… Like the game theory just I am asking the question, your problem is convex, so you might have some interior feeling that convexity had been good for, so long may be it would not behave badly right now mind as you know keeps on wondering. So, wondering as well as wondering, so you can forgive me for this little hang-ups. The question is, is this true if you have a convex problem; that is that question. If this is true then we said strong duality holds. Now, question is this always true that is that is now interesting question, you might take it has I will not give you this question as homework. It is a thing that you need to founder upon, because this is one of the central facts and a beautiful fact of optimization theory that is why that is that is what makes it an elegant subject.

Now, let us see what we can do out of this. So, this is sometimes called value of c p, minimum value of f over those and this value, and this is called the value of Dp dual problem. So, what question - the major question lies is, this is quite a tough question you just cannot figure out immediately that this will be equal or not. But most little bit of example that you can try out at home, take a simple convex problem, I give a convex problem.

So, is above fact which I can write a star now true, just check it out you will find it is actually a most convex nice problems it is. Now, instead of asking such very, very stringent question equalities are stringent things for many people equality might be something very beautiful precise, because it exciting the two functional value objective values of two different problem, one is the maximum ,one is the minimum is actually equal is a very big statement. But it is a something slightly loose I can say for example, any functional value for any feasible element in c for any x in c f f(x) is definitely bigger than equal to the value of Dp. So, if you take any feasible x, if suppose this is true and suppose this is true then f(x) any for any x, f(x) is always bigger than equal to c p, bigger than equal to Dp. So, is the something slightly loose means get a c. I can get something not so strong there, I was something loose which can just work through.

(Refer Slide Time: 34:29)



Now, that look at f(x) plus… So, this is my Lagrangian function just to recollect from here and now let me figure out some relationship between them. Now, suppose x is feasible and y of course is in R n plus then this whole portion is negative in less than equal to 0 non positive to (( )). Then this portion is obviously less than f(x), because if added. So, negative quantity that something, so that value of that quantity decreases. So f(x) is bigger than equal to L(x,y) for any x in c which is the feasible set and y element of R m plus.

Now, let me fix up the y in R m plus, let me fix up sorry let me fix up the x that I have taken in c. And I keep on varying the y, so whatever y I take this is the story for this particular x. So for a fixed x, f(x) then you think what can I do if. I fix x I cannot move. Let us on a fix y fix nothing can be done, I now you see, I cannot, if I fix the x, then I have to operate something on y, but on y I have told I am always maximizing on y. So, I just cannot take this step first so let me just change it. So, for a fixed y this is true for all x in c; so, I am see. This is how would one would argue. Now, once I get this. This would immediately mean the following that now what I can do is my y is fixed, but my x is varying I can write inf of x over c of f(x) is bigger than inf of x over c L(x,y).

And now, once y is fixed the infimum over smaller set would always be bigger than infimum over larger set. This is the nothing but the value that is what I see value of the problem c p which is bigger than infimum. Now, I have done it for a fixed y, but this is true for every y, if I take 1 y this argument hold take another y the same thing will hold. So, this is true for all y which tells me that value of c p is always bigger than theta y for all y. This would imply that value of c p. Now, value of c p is a fixed number. So, I can operate supremum of y on both side, but the supremum will have no effect on this side, because independent of y. So, this would finally, boiled down to the following fact this is value of Dp.

This will say that value of c p is bigger than equal to value of Dp and this is something which we got in a very straight forward way and this is called weak duality. Now, how to under what condition will equality hold, answer is surprisingly beautiful. If slater condition holds then equality holds, if slater condition fails equality may fail. But then you ask a question is there any sub class of convex optimization problems, where, strong duality will always hold and that class of optimization problems is called a linear programming problem. If my objective and constraints are linear then strong duality will always hold.

So, tomorrow in our next class, I would prove the strong duality theorem under slater condition for the convex case. We would then construct for certain very special class of convex problems, what is the Lagrangian dual say for a liner problem, for a semi definite programming problem and then knowing that linear programing is very special, because without any condition the strong duality holds. We will get into the pleasures of linear programming. See, what does this result tells me? It tells me that if my value of Dp is finite that is if p Dp is feasible which it is always, because y element of R m plus. Then if this is feasible and this value is finite then value of c p and if c p is feasible. I know what is the lower bound to c p; I know that this problem has a solution in the sense that it will have an infimum, but I do not know whether there will be an x where that infimum would be achieved.

So, with this little facts we stop our lecture today here and tomorrow we will get into the beautiful world of strong duality, compute towards for certain class of functions certain class of problems, important convex problems, conic problems semi definite problems and then we will start taking a journey into looking into the special problems. And the first journey will take is into linear programming and not just linear programming, as I tell you we will take a journey into the pleasures of linear programming. Thank you very much.