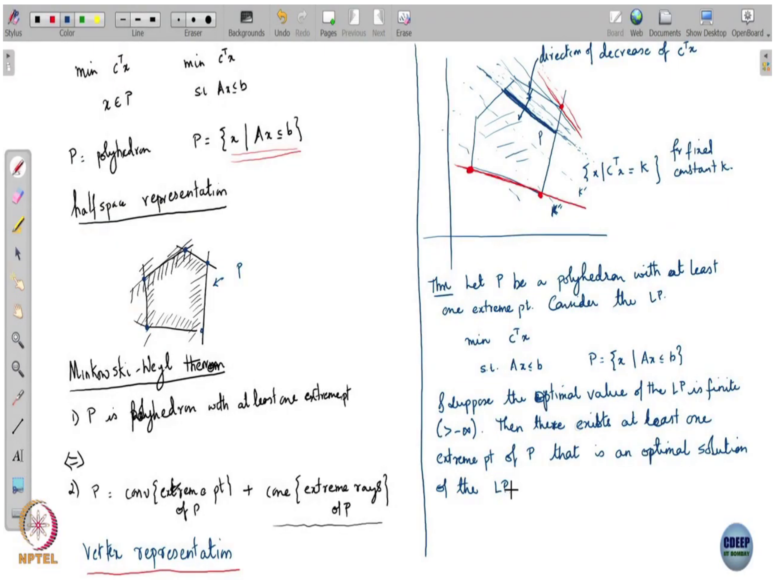
Optimization from Fundamentals
Prof. Ankur Kulkarni
Department of Systems and Control Engineering
Indian Institute of Technology, Bombay

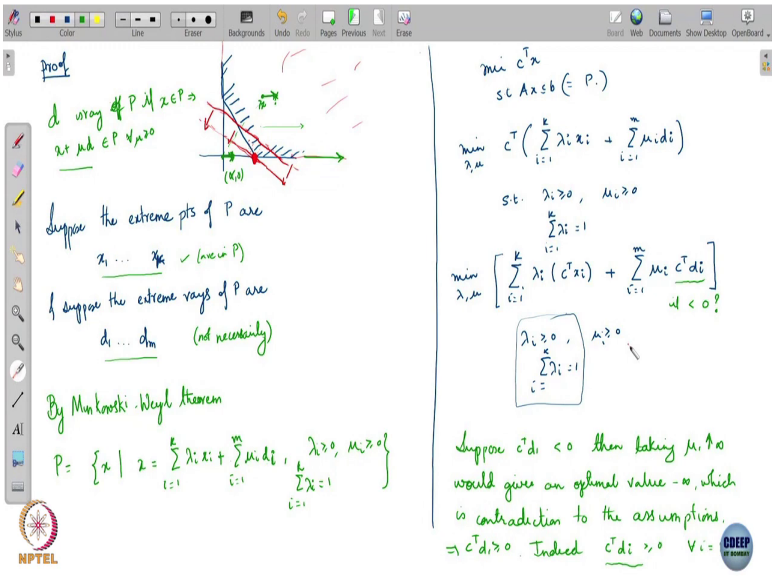
Lecture - 10B
Extreme points and optimal solution of an LP (continued)

(Refer Slide Time: 00:17)



So, the claim is that it does not matter with polyhedron is bounded or not. If it has at least one extreme point and your linear program has a finite optimal value, then it must be that there is at least one extreme point that is the solution alright ok. So, let us see why this is the case. Not necessarily, it depends on how the; so, the question here is it possible that your feasible region is unbounded ok.

(Refer Slide Time: 00:45)



So, that if the polyhedron is unbounded and yet you can get a finite optimal value; is that is that possible? And the answer is yes, so for so it is possible. For example, you look at suppose, if I take a feasible region like this just in r two, I am talking of a feasible region that comprises something like this and the objective function that I am trying to optimize is as this sort of contour which decreases in this direction. So, I am looking to minimize this function. This is the direction of decrease of the contour.

Now, what will happen is this sort of contour the one I have drawn here is cannot be optimal; so, but you can move it in this direction to the point till to the extent that eventually it is it just passes, just through this corner point, here and that sort of thing can will be obtained. It is irrelevant for this that the set is unbounded in this direction out here.

Because that is not a direction in which you are getting a better value right. So, what matters in getting an optimal value finite in a even in an unbounded feasible region, unbounded polyhedron like that is how the objective function, the direction in which the objective function changes, relative to the direction in which the set itself recedes right.

So, there would be extreme rays and there would be a raise of this set or the direction in which the set recedes, how is that relative to the direction in which the objective function decreases. Because we are talking of minimization, so decreases right. That is what determines whether you will get a finite value or not; in fact, that will be a evident as we do this proof also.

So, let us let just go through this proof ok. So, suppose the extreme points of P are say x 1 dot dot dot x k, these are extreme points of P and suppose, the extreme rays of P are say d 1 till d m suppose ok. Now, the extreme points of P are answer this. So, the extreme points of P are they in the set P? Are they elements of the set P? Yes of course, they are points in their they are points in the set P. What about extreme rays?

Student: (Refer Time: 04:18).

Yes, any point plus an extreme ray, it is a point in the set. Any point in the set P plus an extreme ray would be a point in the set P. What about the extreme ray itself? The extreme ray is remember just a vector ok; is that an element of the set P? Just to give you an example, let us look at this diagram that I drew here. I will just try to draw this with a different color.

See this is for example, an extreme direction in which the set recedes right. You can start from any point in the set and keep going along this direction and you will remain in the set. Now, I can think of this direction as this vector, this little vector here right, this vector. So, 0 comma sorry alpha comma 0 for example, where alpha is some small positive number.

Now, this vector, the point alpha comma 0 ok, I am imagining that as a direction on which is to be scaled and along which I can keep travelling etcetera; but the point alpha comma 0 is not in the set P. But anything any point in the set P, say x plus alpha comma 0 which gives you this point, this sort of point is in the set P ok. Is this clear?

So, these are not necessarily elements of P, d 1 to d m; x 1 to x k are elements of P ok. These not necessarily ok. So, these are suppose the extreme points of P and these the extreme rays of P. Now, what does the Minkowski Weyl theorem tell us? So, by Minkowski Weyl theorem, what do we get? We get that P is equal to x such that x can be written as convex combinations of the x i's; i equals 1 to k, all the lambda i s a greater than or equal to 0, the lambda is summed to 1. I am sorry.

So, by the Minkowski Weyl theorem, the polyhedron can be written as x such that x equal x can be written as sum lambda x i from 1 to k plus mu i d i; i going from 1 to m. And the lambdas are all non-negative, mu’s are also non-negative; but the lambdas summed to 1. Apology, this should be k here. This is clear?

So, this is saying that the polyhedron can be written as a convex combination plus a conic combination of these fixed points, you know these k and points when you are taking the convex combination, and these m points there when you are taking these m points, when you are taking the conic combination, is this clear?

This is not saying that we are not allowing in the definition here, we are not allowing the xi's and the di's to vary. There are being fixed outside the set, outside the step. Is this clear? What is varying here is are only the weights; the weights lambdas that summed to 1 and mu is which are non, which are just non negative. Yes.

Student: (Refer Time: 08:44).

So, the extreme ray is a direction. So, firstly, remember what a ray is? Ray is a directions.

Student: (Refer Time: 08:56).

No, no, no. Ray is a vector such that if I add any; so, a ray is a vector. So, d is a ray, if x plus mu d belongs to is a ray of P, if x plus mu d if x belongs to P implies x plus mu d belongs to P, for all mu greater than equal to 0. This is what it means for d to be a array.

So, I can take d multiplied by any non-negative scalar, add that to any vector in the set P and the result is going to be a point that lies in the set P. Once if I have a vector like that, then that vector is a ray. So, here for example as I said, alpha comma 0 that is a ray; but alpha by 2 comma 0 is also a ray because I can multiply by twice the scalar and I will again get the same property right.

So, the point here is rays really have only a direction, they do not have a length because I can scale them, you know the very definition is such that I can, I am supposed to be scaling them by any scalar mu outside right. So, we do not really look at. So, as a result, it does not make sense to talk of a ray being in the set. Because I can always scale it down and make it out of the set or scale it up and.

Student: (Refer Time: 10:30).

No, no, no. The ray is simply the ray is eventually a vector; it is a point in r 2. So, the way when I say this is the ray, really what I am doing is that actually I am sort of taking this here and moving its origin down to it. I am representing x plus mu d when I am writing that. Is this clear? Yeah, ok.

For clarity, make sure that you understand the non-triviality of the Minkowski Weyl theorem which is which we will use here is that these extreme points and extreme rays are fixed for this set. They are not in you know you give me a fixed set of points and I can generate the entire set using the convex combination and conic combinations of those fixed points only.

Yeah, I am not changing those points as I get; whereas, I try to get a different point in the set alright ok. So, if this is the Minkowski, this is my Minkowski Weyl theorem. So, now, let us come back to our linear program. The linear program says minimizing c transpose x subject to A x less than equal to b and that is my set. That is your set P ok.

So, by Minkowski Weyl theorem, P can be written in this, the P every point in P can be expressed as convex combination of the x i's and conic combination of the plus the conic combination of the d's or the di's which means that I can write this optimization in this equivalent form. I just do a change of variable. I in place of x, I will substitute this.

This form, this expression that I have and then, what I will get is an optimization over. So, my new variables will now become the lambdas and the mu's right. Because as I vary the lambdas and mu's, I get an x and for every x, there is a lambda and mu right. So, it is the standard change of variable that I am doing. So, I in place of x, I am going to put this.

Now, what are my constraints? Well, my constraints are precisely these, these constraints. So, I need I have to put in the constraints are that lambda is are all greater than equal to 0; mu i's are all greater than equal to 0 and the lambda is summed to 1 ok. So, let us look let us. Everyone is following this? Now, let us take this objective and work with rather.

So, I. So, this is the same as we minimizing it over lambda and mu, the objective this. So, what I will do is write this in the following form summation i from 1 till k lambda i times c transpose x i plus summation mu i c transpose d i; i sums from 1 to m and still lambda greater than equal to 0; mu greater than mu i is greater than equal to 0 and lambda is summed to 1 alright.

So, now let us look at this objective little more closely. So, you have the objective comprises of two terms the this here, this term which has all the lambdas in it and none of the mu's and there is this term which has only the mu's and no lambdas right. Now, the lambdas are to be selected so that they collectively satisfy these inequality, these constraints right. They have to be greater than equal to 0 and they have to summed to 1. What about the mu?

If you look at the mu's, they are all to be selected to be greater than equal to 0; but there is no constraint that connects one mu with the other mu right. The mu i's and the mu j's are not interconnected with each other through a constraint. So, they can be effectively chosen independently. So, long as there each greater than equal to 0, your constraints are satisfied right.

So, now look at this objective function. What would happens ok, what would happen if tell me what would happen if one of these is say less than 0? If one of these is less than 0, what would happen? If c transpose d i is less than 0 for some one of these i's from 1 to m?

So, precisely, so if what you can see here is if one of these c transpose di's is strictly less than 0, there being multiplied by a mu outside and the mu is for you to choose right? Say suppose, c transpose d 1, c transpose d 1 was suppose less than 0, if this was less than 0, then what kind what the optimal value of mu that I should choose; mu 1 that I should choose?

I am I can just make it as large as I want and I keep getting a lower and lower value right. Because since this is less than 0, say suppose this is minus 1 for example, I been my I am allowed to multiply it by a mu 1 and I am my constraints are telling me, I can make my mu 1 as large as I want, there is no nothing is been violated by doing that.

So, I can by increasing mu 1 and taking it all the way till eventually till plus infinity, I can bring the optimal value down to minus infinity. Is this clear? So, that then comes in contradiction with what we assumed in the theorem.

Remember, we said suppose the optimal value of the LP is finite right. So, if the optimal value of the LP is finite, it cannot happen that this sort of c transpose d 1 is negative; strictly negative because the optimal value is finite, it ensures that this cannot be the case. So, the kind of polyhedron, we are dealing with is the one, where c transpose d 1 is greater than equal to 0 right. So, let me write this down.

So, suppose c transpose d 1 is strictly less than 0, then taking mu 1, making mu 1 large. So, taking it to plus infinity would give an optimal value minus infinity which is a contradiction to the assumptions and what is true for c transpose d 1? I just said c transpose d 1 has to be non-negative. So, this means c transpose plus c transpose d 1 is greater than equal to 0 and what is true for c transpose d 1 is also true for c transpose d 2 and c transpose d 3 and so on.

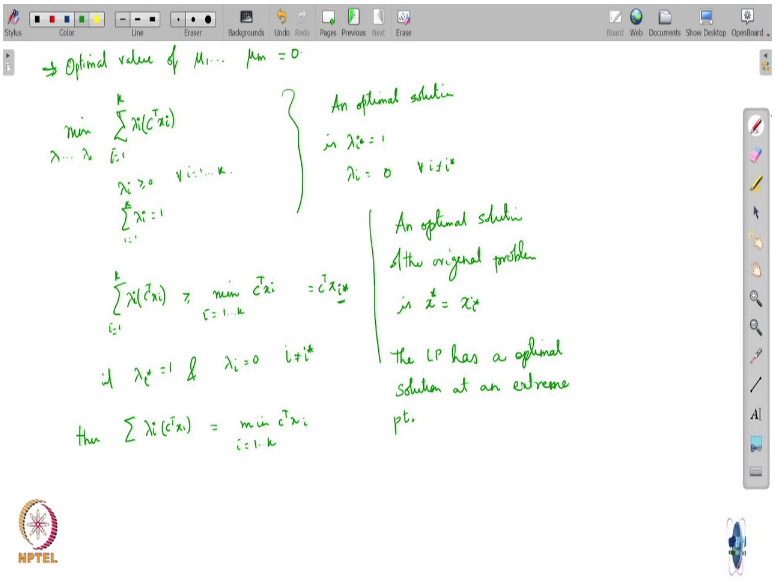
In short, all the c transpose di's have to be greater than equal to 0. So, your c has to be aligned in such a way that it makes an acute angle with all the di's, all the extreme rays. Now, this answers also the kind of the question you had asked at the earlier, is it possible that your objective function that your optimal value is finite even when this polyhedron itself is unbounded. Well, it is possible; it is possible for that to happen ok.

But for that, it is necessary condition is that for that to happen is that see these the extreme rays and theirs vector c make acute angles. Acute angle because we are talking of minimization. Is this clear? Ok. So, that is that this is absolutely necessary for them ok.

So, this has to be these have to be greater than equal to 0 ok. So, now, let us come back here. So, if this has; so, the question then is well if these are all greater than equal to 0 ok, then what is their right value of mu to choose for these? 0, because I am looking to minimize right, I want to get as smaller value as possible. I can choose any value from mu's which is greater than equal to 0, all of them can be chosen independently.

So, long as any of them they are all greater than equal to 0, I do not care. The optimal value for me to choose is therefore 0 because if any of these is if they are 0 if any of these c transpose di's are actually equal to 0 themselves, then they drop out of the equation. The once that are positive, I can scale them down to 0 by taking mu equal to 0 right.

(Refer Slide Time: 21:25)



So, in short, this gives us for free; the optimal value of mu 1 till mu m is equal to 0. So, by just inspection of the linear program and by working with the assumptions that we have which is that the LP must have a finite optimal solution, optimal value, we have got that the mu's are 0 ok.

So, what is that done, that has reduced our LP to this problem. Now, we have minimizing you are minimizing summation lambda i c transpose x i; (Refer Time: 22:12) 1 k, the mu's are all been said to equal to 0. So, my variables are only are only my lambdas. Write them explicitly here; lambda 1 to lambda k, they are constraint to be greater than equal to 0 and they have to be summed to 1 ok.

So, now what can we say about this optimization? So, what is the objective here? The objective if you look at the objective function, it is a weighted linear combination of the lambdas, with what weights? With weights c transpose x i right. I am taking a weighted linear combination of the lambdas with weight c transpose x i.

So, if I want this weighted combination to be the largest possible value, what should I take my lambdas as? Sorry, I wanted to be the least possible value, what should I take my lambdas as? So, this is simply an average weighted average with weighted average of the c transpose x i with the weights as lambda right and the lambda is summed to 1.

So, this being a weighted average will be greater than equal to the minimum of the individual terms. So, this is always so you look at this term c, this c transpose x i times lambda i summed over 1 to k, this is always greater than equal to the minimum of the c transpose xi's; the smallest value of c transpose x i as i ranges from 1 to k and moreover, if I take if you take if you take lambda.

So, suppose, this is the minimum of this is suppose say c transpose x i star; suppose, i star is the index that gives you the least possible value, then and if i take lambda i star as 1 and lambda i equal to 0 for not equal to i star, then actually I can get this equal to right.

So, what this means is in this linear program if I take all I have to do is look at the I that look at these coefficients here, c transpose x i and look for the i star which is where the c transpose x i is least; i equal to i star for which the c transpose x i is least, put all my weight of lambdas on to that one and put everything else equal to 0, that is that sort of point is feasible because it is it satisfies the constraints and it achieves this least possible value, you can all you can take. Is this clear?

So, the optimal, what this means is the optimal solution. So, an optimal solution is lambda i star is equal to 1 and lambda i equal to 0, for all i naught equal to i star. This is an optimal solution. Now, what is this what is this mean then? What is your x? Let us, we did change of variable and got the problem in terms of mu and lambda.

Now, let us go back and recover the x. Remember, we said we will put x, x plus x in terms of mu and lambda in this sort of way right. So, now we found our mu's and we found our lambda, what is the corresponding x now? It is x i star alright because all the mu's are 0, lambdas are also 0 except for i star; i equal to i star. So, my x which means the optimal an optimal solution of the original problem is x equal to x i star now, is let us say x star equal to x i star ok.

Now, what is x i star? x i star, go back was one of these points. It was one of these guys right and what were these? These were the extreme points; these were the extreme points of the polyhedron right. So, what have we got for free? We got that right. So, these are your extreme points. So, what have we got? We got therefore, that the LP has a solution, is optimal solution at an extreme point.