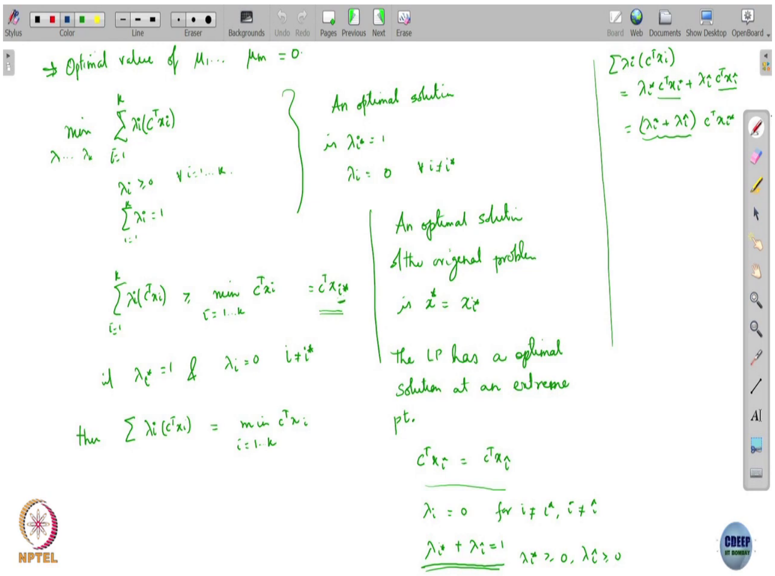
Optimization from Fundamentals

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

Lecture - 10C
Extreme points and basic feasible solutions

(Refer Slide Time: 00:16)



Now, let us make a few other observations here. I just said that an optimal solution is this lambda i star equal is 1 and all the other lambdas equal to 0. Are there other optimal solutions? Are there other optimal solutions? Yeah, precisely, so, if there are, there could be other optima solutions.

Say for example, here I said well x i star is the one that is giving you the least here, least of these, right x i star was the one that is giving there could have potentially been another one, right. Say for example, this could also, there could also be another one. Say let me write it here suppose c transpose x i star is equal to also c transpose x i hat another index which another extreme point which happens to have the same objective value as c i star.

Then, both, I could have taken i hat here and I would have I could have continued with the same argument, I would have got that one as the solution. In fact, in that case I could have done even more. Instead of putting lambda i hat equal to 1 and everything else as 0, ok what I could have done is this. I could have constructed solution like this, lambda i equal to 0 for i not equal to i star and i not equal to i hat, ok.

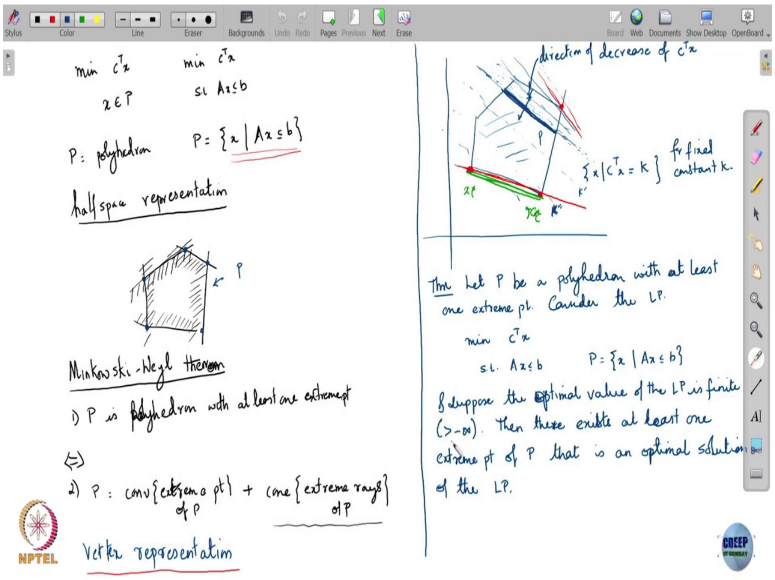
It is not i hat or i star then its 0, and the remaining two, the other two; that means, i star and i hat are chosen to be greater than equal to 0 and sum to 1.

All that matters is that they sum to 1, because then in that case if you look at this if you look at this term here all the what is this going to reduce to. In that case if you look at the objective, the objective is going to reduce when I when you look at this summation lambda i c transpose x i this because all the lambdas are 0 what is left is with lambda i star, c transpose x i star plus lambda i hat, c transpose x i hat, right.

But when c transpose x i star is equal to c transpose x i hat, they are both extreme points have the same value, then what you are left with is just lambda i star plus lambda i hat times c transpose x i star. And so, long as these sums to one I have I still get the same value, right.

So, what the what this means is if I not only is x i star as an optimal solution, x i hat is also an optimal solution, and the segment joining them, right, the segment joining them, which is lambda lambda i star x i star plus lambda i hat x i hat, where lambda is summed to 1, that entire segment joining them is also an optimal solution. Take any point on that segment that is an optimal solution, right.

(Refer Slide Time: 04:07)

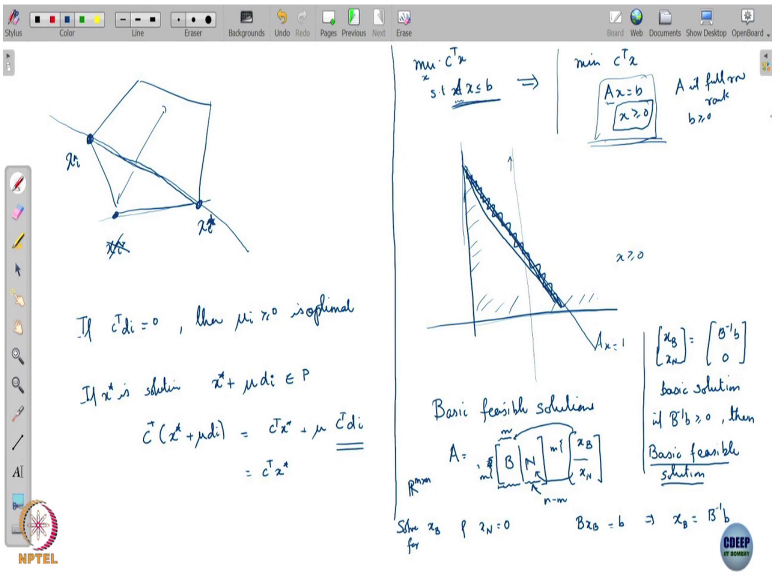


So, that that is seen in this figure that I had drawn here. You have, this is like your point x i star this is your point x i hat and this is the entire segment joining these two points.

Student: When the line segment joining (Refer Time: 04:21)?

Yeah. So, it is a good question. So, is it possible; so, the question he is asking me let me draw new open new diagram here will see, very nice question.

(Refer Slide Time: 04:35)



So, the question he is asking me is suppose let us take this sort of polyhedron. Here suppose your x i star, i star and the other point that I, so the way I illustrated this was I said let us take this as x i hat, right, x i hat here and x i star here and these two are both optimal and the entire segment joining them was optimal.

What the question that, that is been asked is what if you did not have this as x i hat, but rather say this as x i hat, say it was what he said was diagonally across. Say suppose this is x i hat in that case this would the segment joining them which is passing through this would that be optimal. So, then where is the fallacy? You are saying it cannot be optimal, then where is the fallacy.

Student: (Refer Time: 05:27)

The point here is that you it will never happen that such two, these two points are both optimal. Because the very fact that these two points are optimal means that they are on the same contour and then your contour can be moved this way or this way to get a better answer, better value, right, that can that sort of thing cannot happen, right.

So, eventually you it has to be that the extreme points are what is there on the same face of the polyhedron, ok. The extreme points that are optimal around the same field in which case the entire face becomes optimal, yes, yes. So, we are exploiting to the hilt here that the problem is linear, in the sense that the objective is linear, the constraints are linear, the shape of the polyhedron all of that is been exploited completely, ok.

If I have if I had some other non-linear objective then the lot of these arguments will not go through as elegantly as they go to here. And of course, if we did not have a polyhedron then you do not have Minkowski-Weyl theorem to begin with. So, that is another problem.

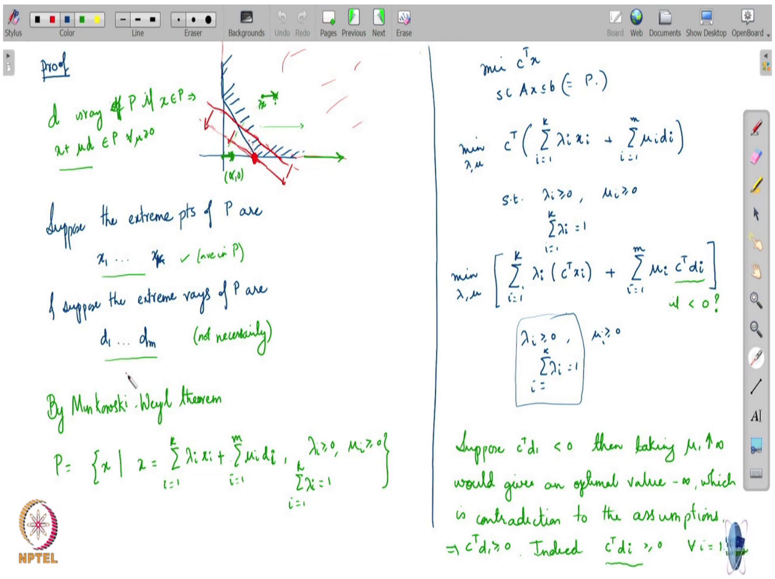
So, you this says that; so, let me come back to what I was saying. So, this just ensures tells us that that is at least one extreme point that is a solution does not mean that there are non-extreme points cannot be solutions, ok. There can be that depends on the, depends on the object, ok.

Now, let me ask you can there be a ray of solutions. So, of course, there can be a segment when you have two extreme points that are solutions, the entire segment joining them as a solution, but can you have a entire ray means that point and direction along which you keep going and you keep yeah and you keep getting all those points as solutions. Is that possible?

Student: (Refer Time: 07:19) polyhedron.

Yes. So, it needs a polyhedron to begin with for you to have a ray of solutions, you need to have a ray of feasible points, right. So, the polyhedron as to be unbounded, ok. So, it is interesting to ask this question only if you have an unbounded polyhedron, which means that you are in the regime where all these d's are these d's the d i's are not 0, ok. There is at least one nonzero d i here, non-trivial extreme ray, ok.

(Refer Slide Time: 07:36)



Then in, but in that case, but that does that ensure that you have a ray of solutions. It can be and then what must be the case for that situation, right. So, remember we said that we argued that c transpose d i's have to be greater than equal to 0, right. And therefore, we said we should put all the coefficient mu corresponding to it as 0, all the mu's were set as 0.

But, what if one of the c transpose d i's was exactly equal to 0? In that case, the mu would be immaterial; it would not appear in the objective at all, right. So, you look at the objective sorry it is here. In that case, the mu would drop out because c transpose d i is equal to 0, right. So, if you have any if c transpose d i is equal to 0 for some i, then what can I take the mu corresponding to that i has, I can take it as anything, right.

So, then mu then mu i greater than equal to 0 is optimal, any mu greater than equal to 0 is ok, right, which means then I can start with any solution. So, if x star is a solution, then I know that x star plus mu d i belongs to P this is feasible it belongs to P because after all d i s an extreme ray, right. So, I can keep going along d i from x star and remain in P. The x star plus mu d i is in P.

And what is the objective at x star plus mu d i? Well, the objective is c transpose x star plus mu d i is equal to c transpose x star plus mu times c transpose d i. And what is c transpose d i? c transpose d i is 0, right. So, this means this is equal to c transpose x star. So, what this means if you; if x star is a solution and d i is a direction d is such is a direction such that c transpose d i is equal to 0, then x star plus mu d i is also a solution for any mu greater than equal to 0, right.

So, I can keep to adding this mu d i to my point x star and I will keep generating solutions. So, it is not just x star that is a solution, the entire ray starting from x star going all the way along the direction all the way to infinity along the direction d i that entire thing is a solution, ok. So, this way you can get a ray of solutions also, right.

But remember it require a fine coincidence your c transpose d i has to be exactly 0, right. So, they have; so, your objective function the direction in which it is decreasing has to be exactly orthogonal to the direction in which the set recedes, ok. You get, but, you get a ray of solutions, but remember still that their optimal value is still finite because the optimal values not changing because of this orthogonality between c and d i the optimal value just does not change as you go along that way, that is what this, right, ok.

So, what has this taught us, what this has taught us is in Minkowski-Weyl theorem is super powerful. It basically has reduced optimization over linear optimization, linear programming to just a matter of enumerating c transpose x i's over the extreme points, right.

So, but the catch in all of this is that we did not know what the extreme points are, to find, to given the x less than equal to b type of representation of a polyhedron your half space representation from there finding the corresponding x 1 till x k and find and the d 1 till d m that is where the challenge is, right; and that is what.

So, all algorithm design would have to in some way or the other deal with this situation, trying to get to the (Refer Time: 12:26), ok. At least if it is aiming to find an extreme point solution, it has to somehow has to get to this, ok. So, this is the qualitative theory of linear programming.

Let us since we have a few minutes let us just do a few examples, ok. Before I do an example let me I want to also mention one more point. Yeah, this is important one. Go back to the statement of the theorem again. So, we saw how we made use of this point that the optimal value of the LP is finite we made use of this, right that the optimal value of the LP is finite that is how we got that the mu's should be 0, right.

Where did we use that the polyhedron must have at least one extreme point? Yes, we use that in invoking Minkowski-Weyl theorem, right. In Minkowski-Weyl theorem also requires that the set as at least one extreme point. That is when you can write it in this sort of form, right.

So, the Minkowski-Weyl theorem, we used it in Minkowski-Weyl theorem. Now, can you give me an example of a polyhedron that does not have an extreme point? Yeah. So, if I take any hyper plane for example, that is a polyhedron that has no extreme points, right. So, in that case, we cannot say that the extreme point is a solution, ok.

Now, what the; so, comeback also now to what I had done in the previous lecture; we said we start with a problem like this, which is simply x belongs to some polyhedron which is it is written in this sort of form, x minimize c transpose x subject to A x less than equal to b.

But then, we converted it to this sort of problem. We said minimizing c transpose x subject to A x equal to b and x greater than equal to 0, alright. We converted it into this form remember. So, again emphasis is on form these, A's are not necessarily equal, b's are not necessarily the same, the c's are also necessarily same.

In fact, the x is also possibly change. Just that it means that there is a you can write this in this form and recover back the solution of your original, alright, ok. Now, focusing on this form again now a polyhedron like this when I write it as A x less than equal to b this sort of polyhedron need not have an extreme point, right.

So, this could be for example, just a hyper plane two opposing inequalities before that will define a hyper plane that has no extreme point. What about this sort of polyhedron? What about the polyhedron on the right? So, we were; let me be more specific here we said that A is full row rank, and we said that b can be taken as greater than equal to 0.

So, this sort of polyhedron now is a polyhedron, which expressed in this form A x equal to b and x greater than equal to 0 for all the constraints, all the variables, right. This sort of; what can you say about, this is a much more specific polyhedron, right. It is not a generic polyhedron like the one on the left.

So, generic polyhedron may not have an extreme point, but what about the specific one? So, it turns out this sort of this specific polyhedron actually has an extreme, always has at least one extreme points. Yeah, see I will give you a pictorial view of this right now, we can prove it in the next class.

See the way this sort of polyhedron looks A x equal to b x greater than equal to 0 the way this sort of polyhedron looks is your from the two rays to n possible orthants of space r n, the x greater than equal to 0 is one of the orthants, right. So, the signs of x will distribute the entire space into two rays to n possible orthants, x greater than equal to 0 just stops out brings the one of them. This is suppose that orthant.

So, this is just talking of, this is just expressing x greater than equal to 0. What is a, what is A x equal to b? A x equal to b is the intersection of several hyper planes in this space, right. So, it is actually a hyper plane, ok, a lower dimensional hyper plane, it is intersection of several hyper planes. So, it is as you keep taking intersections you gets lower and lower dimensional sort of affine sets.

So, it is some affine sets A x equal to b. It is an affine set that is be that is being intersected with the non-negative orthant, this orthant, right. So, an affine set remember is like a gust of shifted subspace. So, I will draw it like this. This is my A x equal to b, and A x equal to b intersected with x greater than equal to 0 is this region, is basically just this, right, just that segment, that part there.

Here because we are in r 2 it looks like a segment, but in general it will be some it will be some polyhedral set. This sort of set can also be unbounded, ok. It is possible for example, it can be that this A x equal to b is aligned like this with the axis in which case it will be unbounded in this direction. It can be unbounded, it is not necessary; it is not a bounded set.

But one thing is for sure it cannot escape, eventually cannot, it cannot happen that it does not hit any of these axis. They will be, it will either hit this axis or this axis, one of those n different axis is going to be hit. So, it will intersect one of them. And when it intersects one of those axis, one or more of those axis, right, ok.

This is this is all in happening in higher dimensions, so remember when several of its coordinate when it intersects several of the coordinates of x will end up being 0, and the others will not be 0. When you when you get, so that sort of point that sort of point will end up been an extreme point of the set.

So, the kind of polyhedra they do not have extreme points and those that are sort of you know say hyper planes, etcetera, which you cannot which you are not able to sort of confine. So, here the x greater than equal to 0 forces you effectively to end up having an extreme point. So, I am giving you an intuitive picture here, but basically what you have what is the claim the thing to notice is a set like this A x equal to b and x greater than equal to 0, this sort of polyhedron always has an extreme point.

So, what this does is this actually helps us in designing algorithms for linear programs, implicitly. A set like this, an optimization problem like this it did not you did not even have an extreme point. So, we could not even say were, that we did not even have a target saying, ok, we are looking for an extreme point to solve the problem.

Now, we this once we reduce it to standard form not only is it standard and easy to program and so on in addition it is guaranteeing us that there is an extreme point. So, the search for a solution also becomes easy, ok. So, you get. So, this sort of set must have an extreme point. Is this clear? Ok. So, yeah so, this is so, the extreme points of this sort of set are what are called basic feasible solutions.

I will talk about this later basic feasible solutions. So, since we have about 10 minutes I can just define for you for what this is. So, if I take the set the matrix A, and I can, now what I can do is I can write it in the following form. So, remember A is full row rank. So, if A has m rows and n columns, then I can write it in the form in this sort of form, where the vectors B, the vectors B are such that; the they columns of A that are linearly independent, ok.

And I can, how many can I put in? How many such columns I can add to B? m of them, right, because that the rank of A. So, I will take B to have m columns, ok. So, B will now have m rows and m columns. And all the remaining n minus m I am putting in n, ok, n minus m columns.

So, I can chose such a such m columns from my from this from the n columns that I have from for A, then that will give me a way of writing A in this sort of form as B and N. And then if I solve using this what I will do is to solve for x, what I what I will do is, I will. So, I am going to multiply this by x let me write this as x B and x N. So, this x B is of length m because what multiplies with B, x N multiples with N, right.

So, what I will do is, I will put I will solve for x B and put x N as 0. So, putting x N equal to 0 and solving for x solving for x B, what does this give me? This give me the equation B x B equal to small b, right which is which and because B is now m cross n and the it has linearly independent columns, right, so B is invertible. So, x B will be equal to B inverse b, ok.

So, a point like this x B comma N equal to B inverse b comma 0 this sort of point is what is called a basic solution. A basic solution is a point is a point like this. So, what you need to do is you create you look for columns of A that are linearly independent, you will m of them you will should be able to find m of them. So, take m of these linearly independent columns.

Solve for the x's corresponding to those columns and put all the remaining x axis as 0, ok. So, see both the remaining x's as 0 and solve for the x's that corresponding to this column that will give you a vector like this B inverse b comma comma 0. Now, this sort of vector may not satisfy x greater than equal to 0, right.

So, if B inverse b is also greater than equal to 0, then we call it a basic feasible solution. So, what I will argue next time is that actually basic feasible solutions are the same as extreme points of this sort of set, ok. A x equal to B and x greater than equal to 0 solving for extreme points is effectively, solving for finding points in x by putting some of the coefficients equal to 0 and solving for the rest, right.

And that is like what I was saying you have these you have these affine set, it intersects, it hits one of the axis by; when you say it hits the axis it means that some of the coordinates becomes 0, you need the other ones need to be found and that that then gives you the that the other needs to be found through the equations that you have, ok. So, that is how we get to extreme work. We will do this in more detail next time, ok.