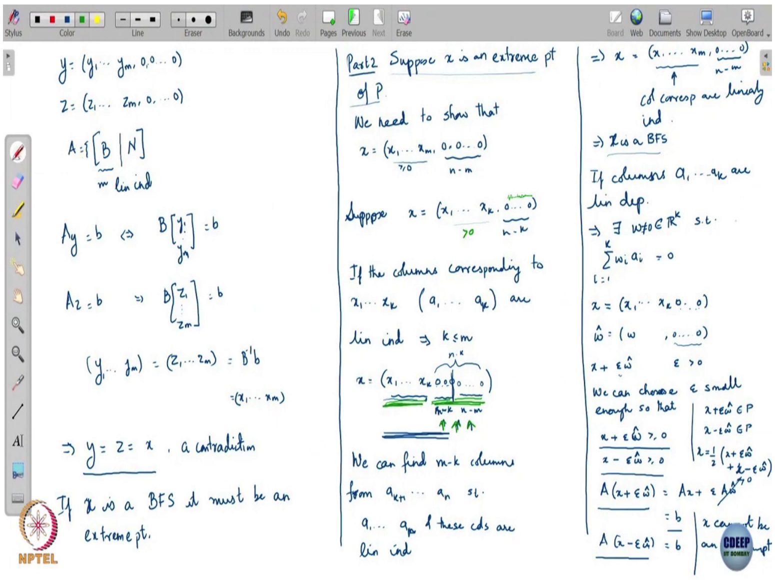
Optimization from Fundamentals

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

Lecture - 11B
Equivalence of extreme point and BFS (continued)

(Refer Slide Time: 00:17)



Alright, now let us look at it the other way. So, this is part 2 suppose x is an extreme point; extreme point of P. Now if x is an extreme point of P we want we need to show that x can be written and we need to show that x takes this sort of form x 1 till x m followed by n minus m 0s ok.

So, suppose I arrange the coordinates of x in this sort of way suppose I write x as x 1 till x k followed by some n minus k 0s. Now k could be n in which case there will be no 0s that is also fine. So, this is every x can be written in this kind of form ok this is without loss of generality.

Now, if the columns of a corresponding to x 1 till x k. So that means these are a 1 till a k right. These columns, if these columns are linearly independent then what can one say. If you have k linearly independent columns of a what must be the case, well the only way that can happen for you to be able to find k linearly independent columns of a is that k should be less than equal to m right.

So, then if these columns are linearly independent then it means that you have when k is less than equal to m right, if k is less than equal to; then what that means is therefore that x takes this sort of form x 1 till x k this is your first k, then you have some 0s here.

How many of these are m minus k 0s and then you have the remaining 0s which is n minus m 0s correct, because k is less than m less than equal to m. It means that m is some you have first you have these k then you have some you have some 0s here k minus m minus k 0s and then some more 0s right.

So, all of these together the number of 0s here is still n minus k just as I had written before. But what is the meaning of this? What I can do is now notice that well this is actually what I can do is. I can pick up, I can pick up m minus k columns from here, such that together with the columns of together with a 1 to a k they form a basis.

See a 1 to a k are linearly independent, so k has to be less than equal to m from the remaining columns I can pick up from the remaining columns of a I can choose m minus k. So, that together they form a basis right.

And then so using so we can find m minus k columns from a k plus 1 till a n, such that a 1 to a k and these columns are linearly independent. So, these are these should all be linearly independent in which case we can then say that this means that x can be written in this form x

1 till x m followed by x 1 till x m followed by n minus m 0s. Where the columns corresponding to these are columns corresponding to this are linearly independent.

So, if I wanted to show that x can be written in this form, for that I if k. If the columns that are I can write x as x 1 till x k followed by n minus k 0s and if the columns a 1 till a k corresponding to x 1 till x k. If they are already linearly independent all I need to do is pick m minus k columns from the remaining put that together that then makes creates for me a linearly independent set and then that becomes my that then becomes a basic feasible solution ok.

So, which means then that this is a x is a basic feasible solution. I can do this if under this condition which is if the columns corresponding to x 1 till x k which means they are the a 1 to a k, if they are linearly independent. If they are not linearly independent then I cannot say that k is even less than equal to m. If they are linearly if they are not linearly independent then k need not be less than m k could be even greater than m right. So, this logic will not work ok.

So, then that gives you the other case; so the second case is that if the columns a 1 till a k are linearly dependent. So, if they are linearly dependent what does it mean? This means that I can find there exist say a w such that so w in R k such that summation w i a i equals 0 going from 1 to k.

So, if these are linearly dependent then I can find a linear combination of these such that that linear combination is 0, there exists a w not equal to 0 in R k. Such that I can if I take this weighted sum of these columns that that gives me 0, this is what it means to be linearly dependent right.

So, since they are now what I do is the following look I have x which is written like this x 1 till x k followed by n minus k 0s. What I will do to it is the following; I will I add to it add this vector let us call this w hat, w hat is comprised of w which is what I have earlier my the w above and a bunch of 0s I have w and a bunch of 0s right.

So, now look at the vector x plus epsilon w hat. What can you say about x plus epsilon w hat, where epsilon is just some small positive quantity; x plus epsilon w hat what kind of components it has?

If you look at the last n minus k components all the n minus k components are all 0 the last n minus k components are 0. The first k components are getting perturbed a little bit. So, you have the components of x 1 you have the x 1 till x k, but then on top of that you are adding epsilon times the w i's right or. So, it is being perturbed a little bit.

But I can find we can choose; we can choose epsilon small enough. So, that x plus epsilon w hat remains greater than equal to 0 right. So, remember x was so here we started of saying that x is an extreme point. So, of P that means x is certainly feasible which means x these are all greater than equal to 0 these component; the last m guys were all last n minus m were 0 right.

So now, what I am saying is I can put a small enough perturbation ok. I can put a small enough perturbation, so that these are all greater than equal to 0. Because what I have done the form of x I have assumed is that the first the all the nonzero ones are here and the remaining guys are all 0 right.

So, we can choose epsilon small enough, so that this is greater than equal to x plus epsilon w hat is greater than equal to 0, not only that x minus epsilon w hat is can should also be greater than equal to 0 yeah. So, fixing A w and scaling by epsilon, so x plus epsilon w hat is greater than equal to 0 x minus epsilon w hat is also greater than equal to 0 and moreover look at this x A x plus epsilon w hat. What is this?

This is A x plus epsilon A w hat. What is this quantity? A x is b and epsilon times w hat is what? It is w it is so all the last n minus k are all 0 in w hat. So, what you are really doing is doing this sum here; the sum written here summation w i a i here and that is equal to 0 right. So, this term therefore is 0. So, what you are left with is just b.

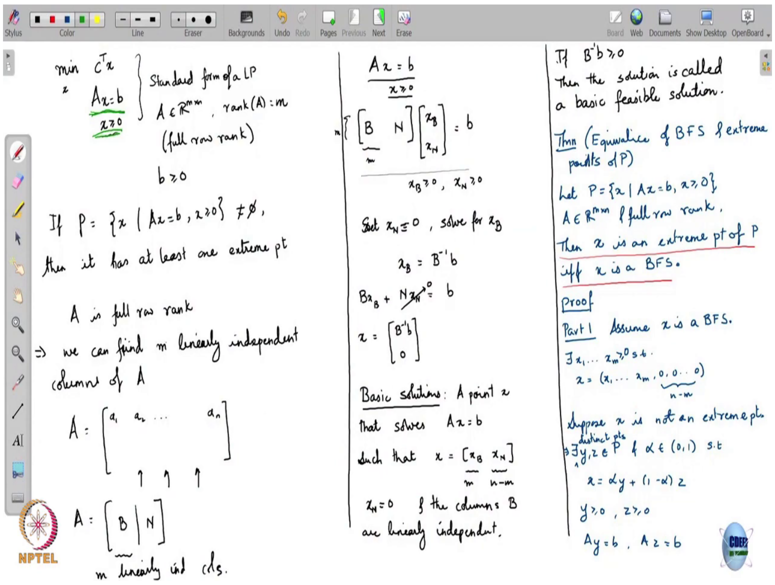
And similarly x minus epsilon w hat is also equal to b. Now what did we conclude as a result x plus epsilon w hat is greater than equal to 0, x minus epsilon w hat is greater than equal to 0, A times x plus epsilon w hat is great is equal to b, A minus A into x minus epsilon w hat is also equal to b right.

What this means is if you look at these two points x plus epsilon w hat that belongs to P, x minus epsilon w hat also belongs to P. Now if these two both belong to P, but then x is actually nothing but half this plus half plus the other one, x is half of x is simply the average the midpoint of these two, right.

So, what have we concluded? We have concluded that if the columns if it so happens that the columns a 1 to a k are linearly dependent. Then you can find at 2 points in P, x plus epsilon w hat and x minus epsilon w hat 2 point whose midpoint is x right; which means what which means x cannot be an extreme point right.

So, it means x cannot be an extreme point and that is a contradiction, because we started off assuming x is an extreme point of P started of assuming x is an extreme point and we found that well. If in this case a 1 till a k if they are linearly dependent then it cannot be an extreme point. So, the only possibility is that they are linearly independent and then if they are linearly independent. We said that it is actually possible to look at view x as a BFS by padding additional columns from the remaining n minus k columns is it clear.

(Refer Slide Time: 14:53)



So, what is what we concluded as a result is that essentially is go back to the statement of the theorem, we concluded just this that x is an extreme point of P if and only if x is a basic feasible solution. Now, these sort of characterization which basically take you from a geometry property to an algebraic property are extremely useful, because computers can only do algebra they can cannot do geometry right.

Extreme point is a geometric property that has no formula associated with it BFS has a formula associated with it right. So, there is an algorithm you can now try to design around this ok. So, what so the most the most popular algorithm for solving linear programs actually just basically does this.

It just goes from one basic feasible point to another looking for trying to see if that is a solution alright, that is what is called the Simplex method we will come to that in later if possible. So, but let me tell you a couple of corollaries that follows in a straightforward manner from here yes which part of the proof.

Student: (Refer Time: 15:57).

Yeah, so these were all assumed to be positive that was in the first case. So, going back here let me revise this see x can always be written in this form where you put all the non all the positive guys as x 1 till x k.

And the rest as rest are 0, it could be as I said k is equal to n in which case there will be no 0 elements ok, all elements will be positive. But ok now if the yeah so the so if you have if the columns corresponding to x 1 till x k are linearly independent, then I have fewer then I have I actually have fewer columns then I needed to make this a BFS.

So I have k of them they are linearly independent, all I have to do is borrow m minus k additional columns from these guys from the ones that are multiplying these elements. So, you have these k then you have m minus k and then you have the last n minus m.

So, all I am saying is I can find the sum I can find I can find m minus k columns from these remaining ones here, such that together with these a 1 till a k. They form a linearly they form a basis they form they should form they become m linearly independent columns.

Student: (Refer Time: 17:29).

No, we have not used that here.

Student: (Refer Time: 17:32).

So, x if x is an so if x is an a we just use that yeah. So, we basically said well if we did not use that x is an extreme point, we said that suppose it has these many if this is the form of it. Then all I need to do is I need to find m minus k and if these a 1 to a k are linearly independent. Then all I need to do is find these remaining additional columns and create that into a basic feasible solution right.

This, what you are saying is right this observation will work whenever k is less than m, whenever the a 1 to a k are linearly independent. This will work does not matter whether x is an extreme point or not. But that is not the only possibility the other possibility is that they are linearly dependent and that is where we need that x is an extreme point right correct.

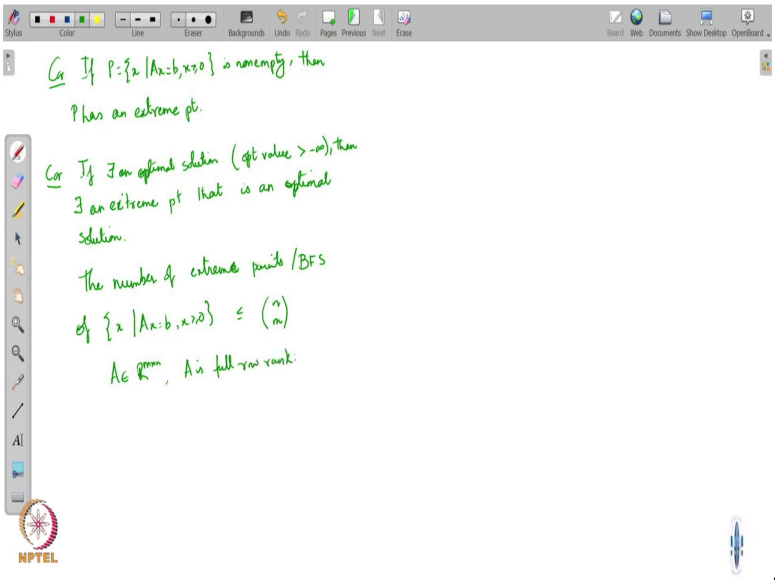
So, what he is asking is it possible that you can have a polyhedron is an extreme point of a polyhedron may not have some component 0 right that is a fair comment. But the reason why this theorem works is because this is not an arbitrary polyhedron anymore. It is an LP in standard form it once you have a polyhedron written in this form A x equals b and x greater than equal to 0, then extreme points must have this feature.

This is now this polyhedron now is in the non negative orphaned of your n dimensional space and there is an affine set that cuts through that.

So the part of the affine set the affine set is a is this A x equals b that the part of it which lies in the non negative orphaned that is your polyhedron. Such a the we are characterizing the extreme points of such polyhedron. What you are saying is absolutely right if I wrote out an arbitrary polyhedron drew an arbitrary polyhedron on a page there is nothing there is nothing to say that it should have you know the the extreme points must have some coordinates 0.

But these kind of polyhedra do have that right. So, this is one of the also an unwritten advantage of converting an LP to a standard form that you get. Firstly, you get two things one is that if the feasible you get a formula for the extreme points that is one thing. The other thing is that if the feasible region is non empty you have for free getting a guarantee that there is an extreme point right ok.

(Refer Slide Time: 20:23)



So, let me just write this out since we are on that topic. So, let me write this is a corollary. So, if P which is x such that A x equals b x is greater than x greater than equal to 0, if this is non empty then P has an extreme point. Why does it have an extreme point? Because it always has a BFS right, if it is non empty it must have a BFS right, because and once it has a BFS it has an extreme point.

And then combining with what we already know if there exists an optimal solution, which means optimal values what this means is optimal value is not minus infinity it is greater than minus infinity. Then there exists an extreme point that is an optimal solution. It has a simple question for you how many extreme points can you have.

Student: (Refer Time: 22:13).

So if I gave you a bunch of if I told you that I have a polyhedron in n dimensions, how many extreme points can it have? Is there a simple answer to this question? Say a polyhedron just in r 2, how many extreme points can it have?

Student: (Refer Time: 22:36).

Any number of extreme points right because you can just you can just keep making you know a polyhedron that looks or more or more like a circle and you can get as many extreme points as you like through them. So, thus the dimension does not automatically say how many extreme points you should have you can have.

But the beauty of this the standard form is that now that you, you are in the standard form you can actually say put a bound on how many extreme bounds points you can have? How many would that be?

See every extreme point is a BFS and how do you generate a BFS, you generate a BFS by looking at m linearly independent columns of A right. How many such choices do you have for linearly independent columns of A?

Student: (Refer Slide Time: 23:32).

You have n columns you want to choose m of them n choose m is the number of linearly possible choices not no more than that. And each such choice will give you potentially 1 BFS or 1 extreme point right. So, the number of the number of extreme points or basic feasible solutions of this set x, such that A x equals b x greater than equal to 0. This is the number of extreme points or face basically the solution no more than n choose m right. Where remember yeah where A is R m cross n and full row rank.

So what this is another thing that the that it that putting an LP in standard form, what it gives you is that well it tells you that at the most you need to search over these many points. This is the number of points that your solution lies in one of these, if it has a solution it lies in one of these.