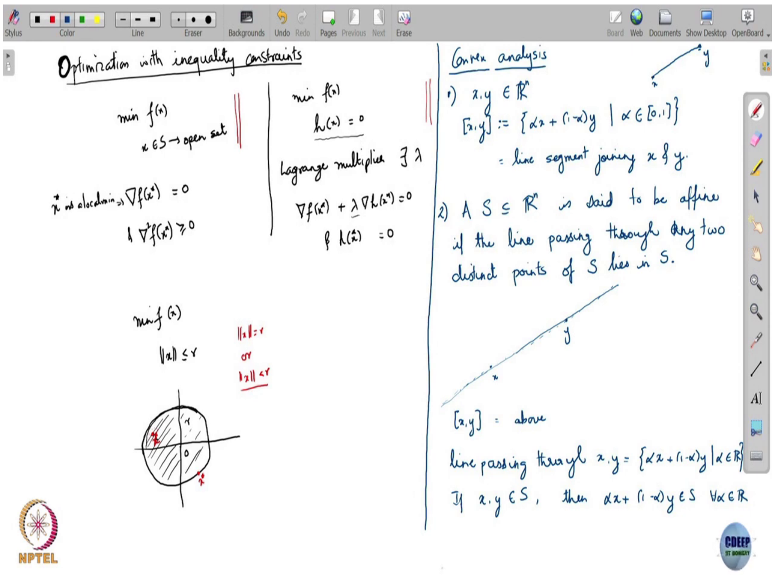
Optimization from Fundamentals
Prof. Ankur Kulkarni
Department of Systems and Control Engineering
Indian Institute of Technology, Bombay

Lecture – 8A
Convex Analysis - I

(Refer Slide Time: 00:25)



Let us move on to the to a harder topic in Optimization which is which involves Optimization with inequality constraints. Now, remember what we the two observations we have made so far, one is that if you have an optimization problem, where the constraint is an open set. If this is an open set then a necessary condition for optimization which means just to.

So, long which means a condition that any minimum minimizer must satisfy is simply that the gradient is. So, if x star is a local minimum, then the gradient of f at x star should be equal to 0 and if you have second order information the hessian should be positive semi definite right.

Now, if you have instead equality constraints, suppose you have minimizing this subject to equality constraint let us say for simplicity a single equality constraint, h x equal to 0.

Then in that case we found that in the necessary condition takes a slightly different form. That we you need to introduce another variable another variable into the calculation which is called the Lagrange multiplier corresponding to this constraint. So, we for every constraint that is a Lagrange multiplier.

And so, the condition then becomes something like this. Gradient of f x star plus lambda times gradient of h evaluated at x star should be equal to 0. So, Lagrange multiplier there exists a lambda such that this is and of course, that h of x should be equal to 0; h of x star is equal to 0, right. So, these the so the point here is that, when you are in an open set the problem has a looks very different from where you are on the surface.

Of course, when you are on the surface also we were able to somehow reduce it to a problem of optimizing over an open set, but that was a reduction, but then that reduction also brought in this new quantity, which is called the Lagrange multiplier right.

Now, the problem with inequality constraints is that it combines the features of both of these problems. So, what happens when you are in if you have an inequality constraint like this? Suppose you want to suppose minimize this subject to the norm of this less than equal to r.

So, what sort of region is this? So, this is suppose these are suppose my axis this is my origin this is a radius r and I am looking I am optimizing the function f over a ball of radius r. So, the ball contains all the points on the surface of so on the shell of the ball on the surface of the sphere as well as ball all these points that are inside, right. So, if I consider a point x star here that is inside then and this is suppose a local minimum, ok suppose write this in red.

Suppose, I consider a point x star here that is in the interior of this ball and it is suppose a local minimum of this optimization problem then it is also what it means is it is a local minimum over this a small neighborhood around it, which lies completely inside the ball. So, I can take a small enough neighborhood around x star, which lies completely inside the ball and x star is a local minimum over that is a minimum over that neighborhood.

Then what that would do is that would then in that case the problem would sort of take the character of this problem of optimizing a function over an open set, right. Whereas, if my x star is here on the boundary or on the shell of this region, then the problem would take the form of character of this sort of thing, because you would be satisfying h x equal to 0 or in this case norm x equal to r. So, what happens is in once you have inequality constraints like this.

Inequality constraints are basically capturing two conditions together either saying norm x is either equal to r or norm x is strictly less than r either of these is ok with r. Now, strictly less than r is basically the interior of this, of the sphere and that is like this problem that is like optimizing over an open set. The one that you have equal to r is the shell of the sphere, that is like optimizing over inequality constraints, right.

So, all the complications in optimization appear because of this this sort of character of this problem, which is sort of changes. There is a phase change that happens once you go from interior onto the boundary. The nature of the problem the way you think about it all of it changes ok. And so, we what we will now build towards is ways by which you can again come up with conditions that a local minimums of such a problem must satisfy ok.

And then using those one can build the algorithms for finding a local minimum ok. What this also entails is that we need to learn a little bit about what is called convex analysis ok, convex analysis. And so, I will tell you what that I will introduce that subject today. So, the so, convexity is or convex analysis is basically the study of overall shape of a region or of a body ok.

It is not just about its local shape, but about the overall shape of it and so, for that let us start introducing a few concepts. So, suppose if I give you two points x comma y in R n the, this here we will just denote the line segment joining x and y.

So, what is that? What is the line segment joining x and y? How do I express that? So, this is a say a point alpha x plus 1 minus alpha y, where alpha can range over 0 to 1. So, as I range alpha from 0 to 1 if here is my point x here is my point y as y range alpha from 0 to 1. At 0 I am at y, at 1 I am at x I cover this entire segment. So, this is the line segment joining x and y, ok.

Now, a set S of set of R n is said to be affine, if the line passing through any two distinct points of S lies in S, ok. So, for so, what this means is. So, if a set is said to be affine if the line passing through any two distinct points of S lies in S.

Inside you should you can take any two distinct points say point x and y. So, these are points in S. You look at the line that passes through them. Now, can you explain can you tell me what is the line that passes through them? What should be the expression for this? The segment joining x and y? I have defined above. So, what is the line passing through x and y? Yeah. So, if I take the so, the line passing through x and y is all points like this.

Alpha x plus 1 minus alpha y, where alpha is any real number. Alpha can be positive negative greater than 1 less than 1 does not work. So, it of course, contains the segment, but it also extends beyond the beyond x and beyond y ok. So, that is this line. Now, what does this what is the, what is the definition of an affine set saying? It is saying that set is said to be affine if the lines passing through any two distinct points of S lies in S.

So, take any two distinct points like this x and y, the line passing through them which is which is this set its the set of all alpha x plus 1 minus alpha y where alpha ranges over R. This this line should be in S right. So, which means that this here. So, if x and y belong to S, then alpha x plus 1 minus alpha y belong to S, for all alpha in R. This is what this is what is this is what it means for S to be an affine set ok.

Now, can you give me examples of an affine set? So, the entire ambient space R n is an affine space clearly. Any other? A plane a plane in R n. So, we have to define for you what a plane is first we have not done that, but if you are just willing to just let us look at r let us look at R 3 for example, and look at a plane in R 3.

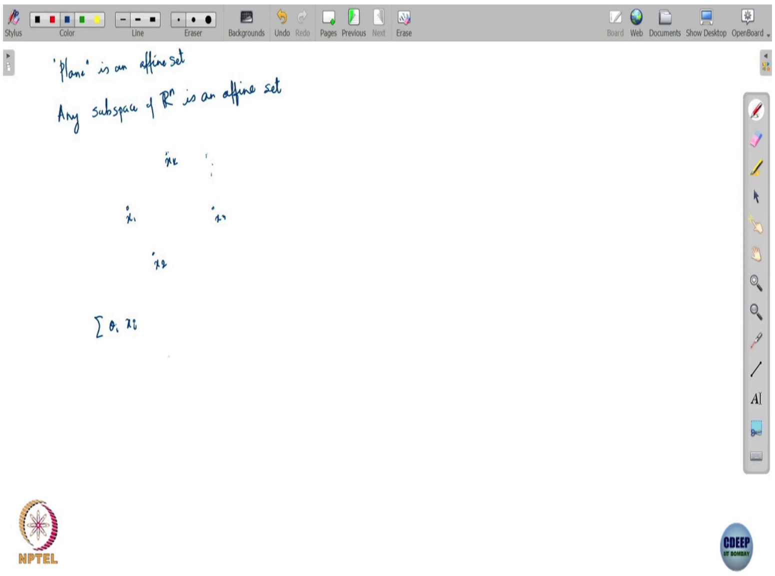
R 3 is the space around us plane and R 3 would be say the wall here that the entire that wall you just extend from it to all the way to the infinity up and down in all directions. That will define for your plane in R 3 ok.

Now is that an affine set? Why is that an affine set? I need to if I take any two points that lie on the wall look at the line passing through those points, that entire line will still be on the wall ok. So, that is that that is clearly an affine set. Now suppose I ask you is ok, so, the wall is an affine set what about the wall and all the air on this side of it ok? Why is that not an affine set?

Student: (Refer Time: 13:18).

Precisely. So, if I take two points like this that here take the line passing through this, this will go hit the wall puncture the wall and go on to the other side of the wall right. And so, it will not necessarily be it is not necessarily that its such a line will be always in this set ok.

(Refer Slide Time: 13:44)

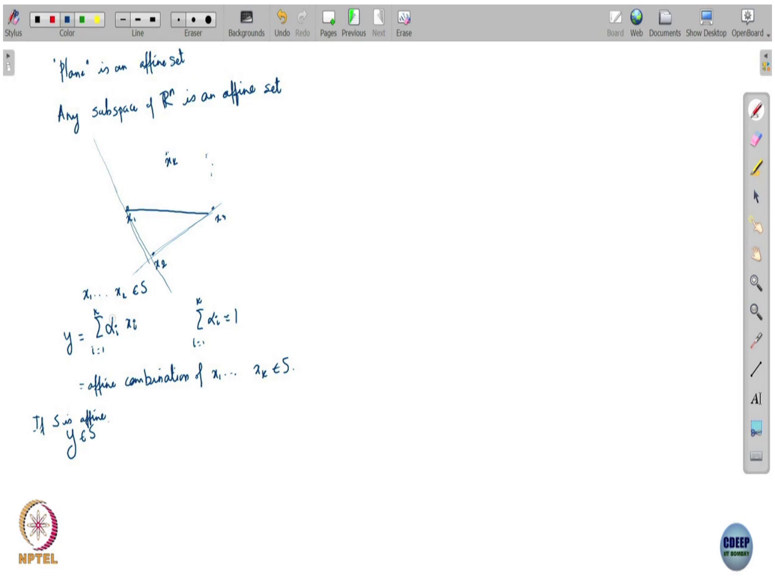


So, a plane the so, a plane that you just mentioned is an affine set. I will explain what a plane as we have to define it, but a plane as an affine set. More generally actually every subspace is an affine set right, any subspace of rn that you take is an affine set. Any subspace of rn is an affine set.

Because, the subspace is always closed under scalar multiplication and linear combinations. So, consequently subspaces is an affine set. Suppose if we take k suppose we take k points in an affine set, ok. Suppose you take not just to take k points in an affine set.

How do we define? So, what one let me ask this question in the following. Suppose I take k points in an affine set and you say x 1, x 2, x 3 dot dot say x k ok. Now, look at this sort of point. This point is summation theta i x i. Why is the earlier notation; let us use alpha.

(Refer Slide Time: 15:12)



Summation alpha i x i, where alphas are such that they sum to 1. So, you look at this so, you have taken these k points, which are in the air in an affine set. So, x 1, x 2 x k all of them in S and consider this point let us call this point y ok y is defined as summation alpha i x i I mean from 1 to k, where alpha i's have this are chosen such a way that this sum to 1..

Now, what can you say about this point? Now, let us take 2, let us let us suppose I make alpha 1 and alpha 2 positive and the alpha 1 and alpha 2 non-zero and the other's 0. So, this will that will then capture as I vary my alphas, that will then capture all the the entire line passing through x 1 and x 2, right.

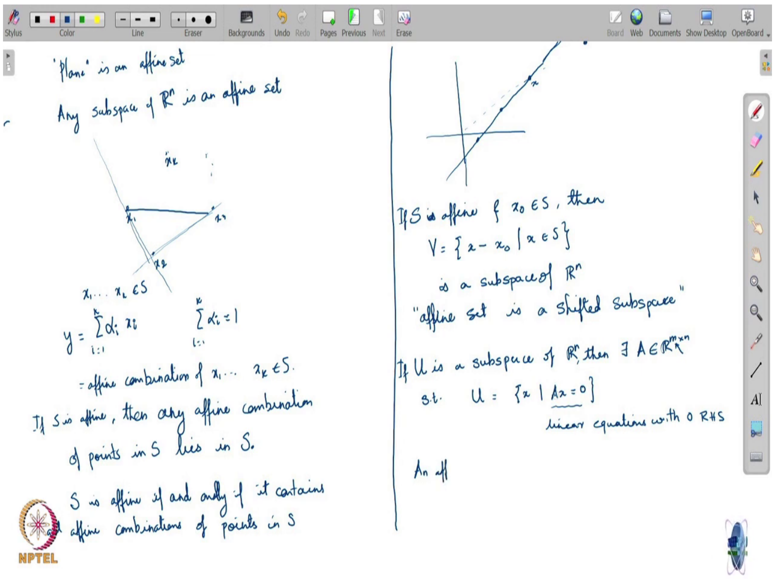
If I take similarly alpha 2 and alpha 3 as non-zero and make everything else 0 that will capture for me the entire line passing through x 2 and x 3. What if I took alpha 1, alpha 2 and alpha 3 as non-zero? If you think about it for a moment you will realize that this actually captures for you any point, that lies in the same plane as alpha as x 1, x 2 and x 3. The entire plane defined by x 1, x 2, x 3 is being captured by you.

And as you keep doing this you what you are doing is you are you are building your creating, what you are creating is is something what we what is called a hyperplane. As you you include more and more points like this, eventually what you are going to create is something that is akin to a subspace. So, what is going to be the property of that? It is not exactly a subspace I will be clear about that what is what would be the property of that?

It would be such that you so, what it would be such that if you take any two points that lie in that set the line passing through that those two points will be in the set. So, the my the point I want to make is that if your set is affine and then you consider any k points in it and consider a point y any k points like this x 1 till x k and consider a point y that is defined like this by taking a linear combination which sums to 1, ok.

This is what is called an affine combination of the points x 1 to x k which are in S, ok. So, you take any affine combination of this point, this point this affine combination also lies in S.

(Refer Slide Time: 18:35)



So, if S is is affine then any affine combination of points in S lies in S that is the definition of an affine combination. So, it generalizes the previous definition in the previous case.

So, in the previous case we had my alpha 1 was this was this alpha and alpha 2 was the one was 1 minus alpha. So, the alpha 1 plus alpha 2 is still 1 got it. So, yeah. So, this this this condition that the alphas should sum to 1 is simply is just generalizing the is generalizing the previous condition. So, what we are so, here what we are the claim that is being made here is that if S is affine.

And not only that that the line passing through any two points lies completely in S, you can take any affine combination of any point not just 2 any k point ok. And k can be anything, take any affine combination that entire affine combination lies in S ok. In fact, the in fact, you can define you can actually say the following that S is affine if and only if it contains all affine combinations of points in S.

Now, if you start thinking about it this way you will start realizing that an affine set has many of the properties that you would think subspace should have because if you take linear combinations of points in a subspace that also lies in the same subspace right. The only thing that is missing is the the property of scalar multiplication. If you scale the point in this in a subspace you have you are always in the subspace, but an affine set need not be like that.

Take for example, let us draw this in R 2. Say for example, here is an affine set. This is just a line take any two points on it its affine combination is the line itself. It lies completely it is the line itself. So, this is an affine set. The problem is that if I; however, if I scale any point say take a point vector x that lies here and scale it. Say suppose I double the length of it. The double the length of it that will end up here somewhere outside that is not on this line, right.

So, the what is the issue? The issue is that the line need not pass through the origin. So, an affine set is actually a subspace that has been shifted ok. So, this can actually be shown. So, if S is affine and say x 0 belongs to S and look at the set V. V is just say the. So, V depends is this x minus x 0 as x ranges through S. So, what I am doing is I have taken the set S taken some point x 0 in it and subtracted that from all the points in the set.

Now, V will have will pass through the origin. V contain 0 why because x 0 is is in S right. So, V contains the origin. So, but it is essentially just the shift of S by all vectors have been you have subtracted x 0 from every vector in S. So, you have shifted it shifted the entire set as, but now that you have shifted it will pass through the origin and it will now have the properties of scalar multiplication of being closed under scalar multiplication also. So, V is actually a subspace.

So, if you take if you; so, an affine set is this sort of flat set which is not necessarily passing through the origin I shift it it will now be a it becomes actually a subspace after shift ok. Does not matter and it does not matter what I shifted by. They can choose any x 0 and S and that

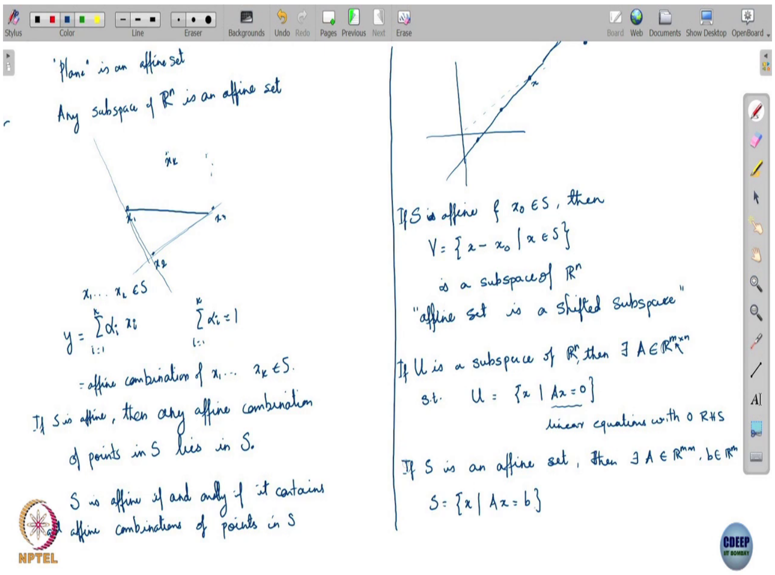
that is up to a; every x 0 works just as well ok. When affine set you should remember is the just the shift in subspace.

Now, if it is a shifted subspace that also tells you something. Now, a subspace is always can always be given by this and by the null space of some matrix, right. So, if you have any subspace, so, suppose if, so, if u is a subspace of R n, then you can always find then there they are all there exists some matrix like this; a in some m cross n such that u is the same as the null space of this matrix. So, u is the set of solutions of the equation a x equals 0, right.

You can always find such a matrix A. The m , the m depends on the dimension of u. With the point the, point is that you can get you can always find an matrix A such that if you like (Refer Time: 26:13) the null space of A that is actually the this the subspace u.

Now, which means any subspace is the solution set of linear system of equations with a 0 right hand side right. So, these are linear equations with 0 on the RHS. So, now what would happen if I shift this? If I shift this instead of having a 0 on the RHS what I would get is some other vector on the RHS right.

(Refer Slide Time: 27:05)



So, then an affine set if S is an affine set then there exist then there exist some A like this because m and b in some R n such that x is just the set of solutions of this system of equation ok. So, an affine set is actually is nothing but a solution set of a linear system of equations. This written in a x just expressed in a geometric way rather than in an algebraic way right.