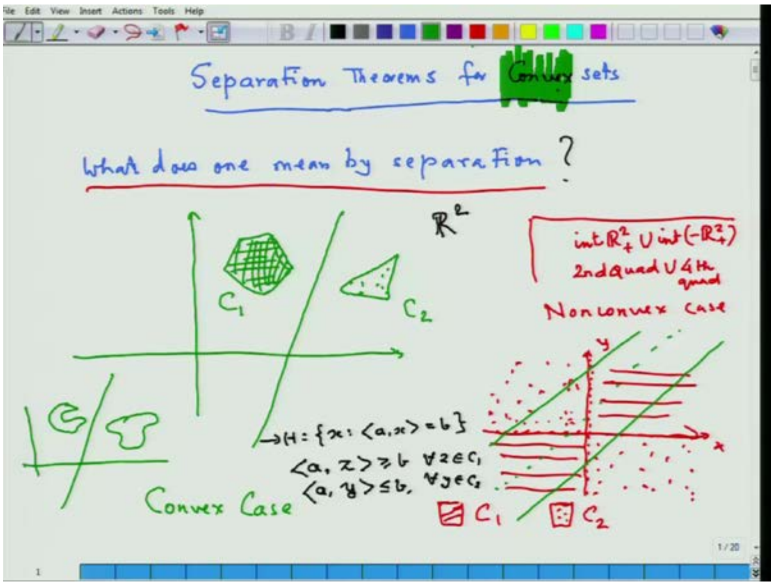
Convex Optimization

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

Lecture No. # 05

(Refer Slide Time: 00:26)



Welcome again to the (( )) talks on convex optimization and today, we are going to start a very important topic called separation theorem for convex sets. As I had already told you, that separation theorem for convex sets are important because optimality itself can be expressed as, can be expressed to the separation of two disjoint convex sets. So, this topic is of extreme importance to our study. And observe that with the term separation the term convex set is highlighted here because in general, for a non-convex case, this thing called separation, need not work.

So, what does one mean by separation? What one means is the following, that if you take two convex sets, which are disjoint from each other with this one and this one, they are convex sets, but they are disjoint from each other in R 3. Then, in R 3, you can pass a plane, which separates this two, that is, you basically put a border or a wall between this two. So, this thing lies on one side of the wall and this thing lies on other sides of the wall.

So, in R 2, which is our playground, you see here, I have a convex set C 1, another convex set C 2 and they are disjoint, then I can draw a line separating this two. So, this line, which is a hyper plane in R 2 divides, as you know, the whole space into two parts. So, C 1 lies in one half space, C 2 lies in other half space, that is exactly what is meant by separation.

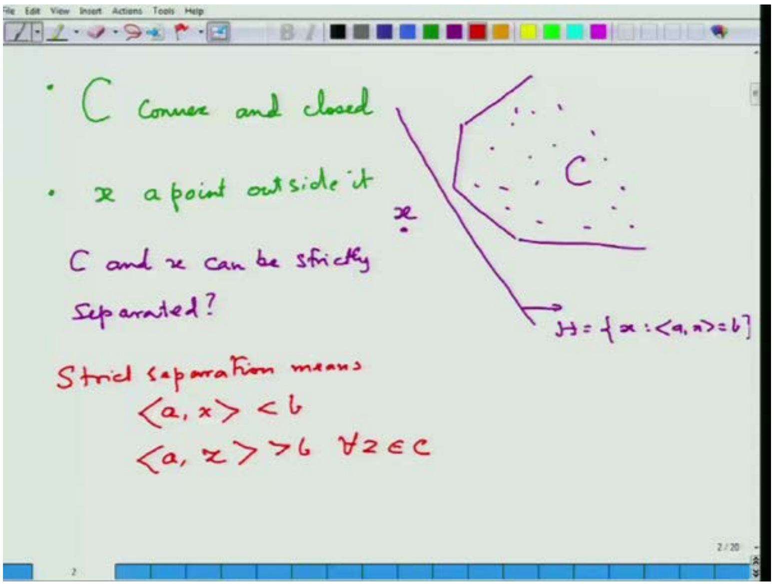
I will mathematically describe this thing. Suppose, this hyper plane H is given as set of all x, I am writing in the general sense of R n though everything is in R 2. Suppose, this is, this is what the hyper plane is. Then, a of z, where z belongs to C 1 is greater than equal to b for all z in C 1 and a of y is less than equal to b for all y in C 2, which is in the lower half space. So, that is exactly what is the meaning of separation, that the whole set of C 1 is contained in the upper half space here and the whole set of C 2 is contained in the lower half space, there cannot be an intermingling. That is, a part of C 1 in the lower half space and part in upper half space, and part in upper half part of C 2 is in upper half space and part in lower half space, such a thing would never happen, but means, you can always draw a boundary.

But look at this non-convex case. This non-convex case, you know, you can draw out. What I have done is I have taken the interior of the 1st quadrant and interior of the 3rd quadrant and is, writing more technically, I have taken interior of R 2 plus union interior of minus R 2 plus.

Now, look at the union of the 2nd quadrant with the 4th quadrant, then we look at these two sets. This 1st one and the 2nd one, they are having empty intersection, but can you draw a line such that the one set, this set is in one set of the line, this set is in other set of the line. Let us try to draw a line, whatever line you draw, you cannot put one of these sets into one part and another of these sets into the other part, that is, one in one half space and other in other half space, whatever be a line it really does not matter; whatever we have line, it really does not matter.

This thing need not hold for the general non-convex case. Of course, you can say no, never mind, I can just draw a non-convex set like this, another non-convex set like this, then I can say this. Yeah, of course, there can be instances where this can hold this separation, but it will never hold. In general, for every pair of disjoint non-convex sets, but for every pair of disjoint convex sets, this thing will always hold and that is what is our goal. So, we would see, we will now prove that such a thing can be mathematically established or demonstrated.

(Refer Slide Time: 05:22)

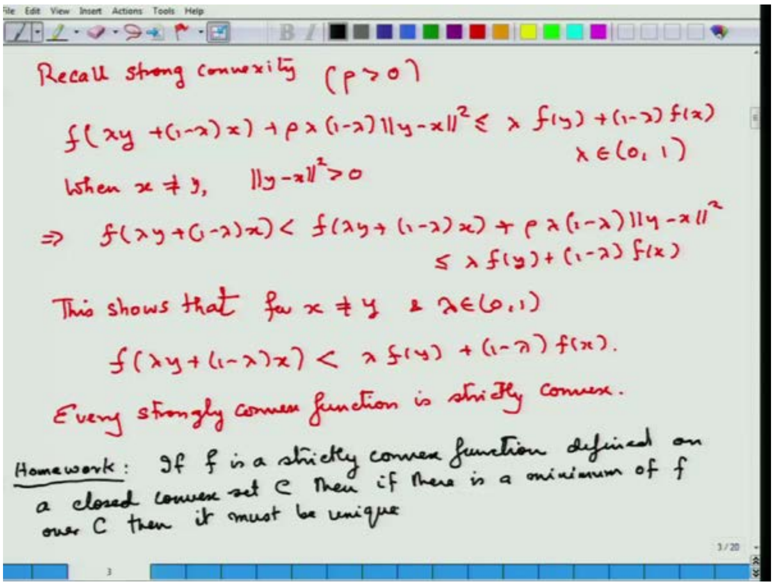


Our 1st case would be the simplest one. C Convex set; x, a point outside. So, what I am having here is something like this, a set C and a point x outside it. What we intend to say, that this set C and this point x can be strictly separated; so, C and x can be strictly separated. So, what I mean by this? What I mean is the following, that x would lie in one of the strict half spaces and C would lie in one of the strict half spaces, that this line would neither touch C nor x. At the same time, dividing these two into two different segments, that is, putting x in one half space and C in the other half space, it would never touch either of them.

So, it means, that if this is written as H, that x equal to b, then what I have is a following, strict separation means the following. So, it should not touch, which means, in this case x is in the lower half space, C is in the upper half space. So, a of x is strictly less than b and a of z is strictly greater than b for all z in C and that is what we would like to establish. For any closed set C, a set C need not be bounded, it could be unbounded. Also, it does not matter, but x has to be point outside C and that, that is all.

How do I establish this? The establishment of this thing calls in to the fore the notions of a strongly convex function, and so we have to learn something about strongly convex functions, which we are already, discussed in the last lecture. We would learn something with more and then go in and talk about this separation procedure.

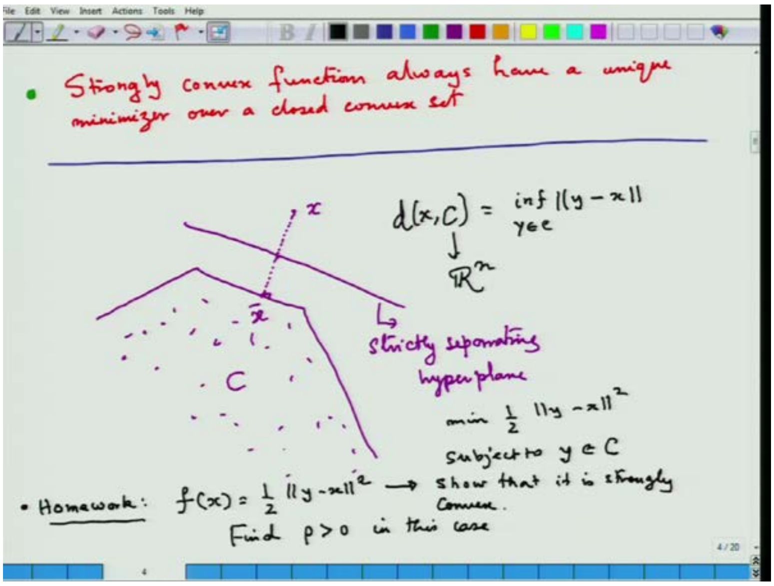
(Refer Slide Time: 08:29)



So, just recall strong convexity. The strong convexity means, that there would exist a number rho, such that for any x, y you give me, I can always write any lambda between 0 and 1. So, this is less than lambda times, therefore y, now this is what is strong convexity, rho being greater than 0. Now, when x is not equal to y, norm of y minus x whole square is strictly bigger than 0 and this would imply, that f of lambda y plus 1 minus lambda x is strictly less than f of lambda y plus 1 minus lambda x plus. So, you have added a positive number to this, so it gets a strict reduction. So, this is again, this shows, that for x naught equal to y and lambda in 0 and 1. So, then, this whole thing is positive actually. Of course, here I have to write lambda in 0, 1; f of lambda y plus 1 minus lambda x is strictly less than lambda f y plus 1 minus lambda f x. So, this means, that every strongly convex function is strictly convex.

So, again, I start with a very simple problem, but it is homework for you. So, if f is a strictly convex function defined on a convex set C, on a closed convex set C, then if there is a minimum, minimum, there is a minimum of f over C, then it must be unique, but for the strongly convex case we have a better result.

(Refer Slide Time: 12:34)



The result, is at, is as follows. Strongly convex functions always have, strongly convex functions always have a unique minimizer over a close convex set. Those who have some knowledge of optimization or those who are experts, who are just watching this possibly for fun or trying to see how a fellow researcher is speaking about the subject, for them, of course, it is obvious and you know how to do, it is all about coercively of the function f.

So, but for others I would not ask you to get into the details, especially those from the engineering background, I would just like them to keep this in mind. So, that is why, strongly convex functions are so important because you are always guaranteed a unique minimize, over a closed convex set there will be a minimizer. It does not say, that the set has to be bounded, the set need not be bounded, but still it will give you a unique minimizer and that is a very, very, very, very, rather important result, which we separate out.

How this result is important for us is what we are going to see. So, we are now going back to this situation or a convex set and a point x outside it. So, how do I am supposed to draw separating hyper plane, which is strictly separating in the sense, that we have just defined? In order to do so we just do the following. Let us drop a perpendicular from x to the set C and then you know, that except this point, say, which I call as x bar, there is no other point on this line segment, which is lying on C.

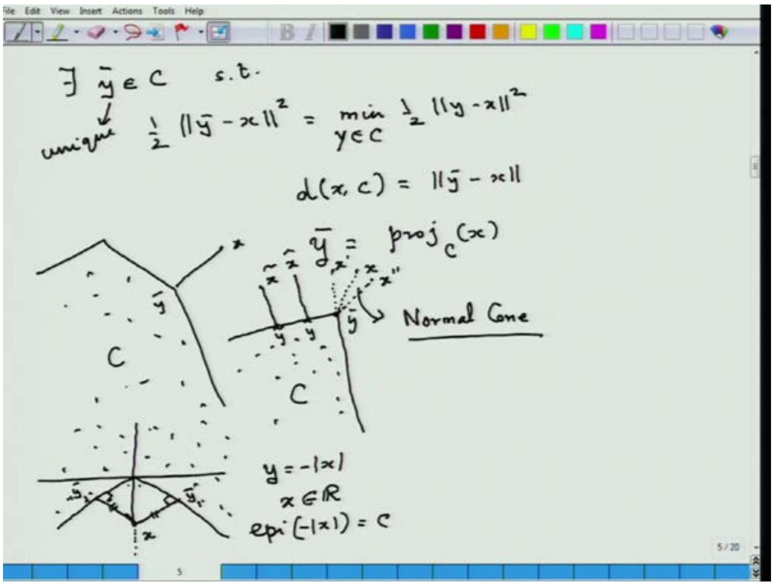
So, through any such line, right, through any such line, any such point, just stroke it at a midpoint of this, draw a line perpendicular to this, perpendicular line to x minus x bar and this could be the separating hyper plane or strictly separating hyper plane. How do we guarantee, that if I just drop perpendicular, there would exist an x bar? It will go and hit and of course, from this picture is clear, that you draw a perpendicular on x, it will go and hit a point x bar on the set C because set C is closed for the boundary of the set, is also included in the set. But how can I be mathematically sure, that such an x bar would always exist, such an x bar will also be there and this would lead us to this result? We have to; we will just use this result.

So, we will now first talk about when there is a point x and a set C, what we are trying to measure by the perpendicular? We are trying to measure the distance between x and the distance between, distance between the x and set C. So, what is the meaning of distance of the point x from the set C? So, what you do is with every point in C you calculate the distance of x, basically you calculate the length of x, every point y in C you calculate the length y minus x.

So, basically what you do? You calculate for every y in C, you do that and by the distance is very clear, should be the minimum one, that is, the distance from the point x. So, you take the infimum of these values when y belongs to the set C. So, how do I guarantee, that this would happen? So, finding the distance is same as trying to find the square of, square of the distance, in the sense, that I can look now at the problem of minimizing. So, any solution of this problem, it is, it is obvious solution of this problem and I am not going to tell you how to do that. This is so clear, that you can just take up a pen and a paper and just finish off in a minute.

So, now observe, that this one is a strongly convex function, this function is a strongly convex function. So, homework is as follows. Of course, in our general writing, y and x, this all are in R n, so the set C is in R n. In our general writing, everything in terms of general spaces though explanations would be in R 2. So, f of x is equal to… Show that it is strongly convex; find rho in this case. Now, this set C is a closed convex set and this is a strongly convex function. So, what we have said, that strongly convex function always, has, have a unique minimizer over a closed convex set. So, there must exist some y in C such that…

(Refer Slide Time: 19:35)



So, what would happen is that there would exist a y (( )) in C such that y bar minus x whole square is minimum half. So, there x bar, what we are calling here as, actually what is the y bar, we have just written same thing. So, I will just draw again for you the picture. Now, this simply means, the distance of x from C is nothing, but norm of y bar minus x.

So, what I am always guaranteeing? That there exists a unique point y bar because a minimum, minimizer is unique, of course, this y bar is unique, there cannot be more than one. There exists a unique y bar, such that at that point, norm of y minus x bar, this is the distance of x from C, that is the minimizer achieved at y 1 and is done uniquely.

So, more generally, y bar is called the projection, you know, when you have the x, y coordinates. So, if you have x, y, so what you do? If you drop a perpendicular on the x-axis, you get the x coordinate; if you drop a perpendicular on the y-axis, you get the y coordinate. So, this is the same idea. So, it is, so y bar is called the projection of x on the set C.

Now, the crux of the matrix are as follows. Take this convex set C and take this x, this y bar of course, is a nearest point from x, nearest point from x to C. But you see, if I take x bar, this another, say x dash here, this is the same y bar, which is the nearest. This is not the nearest because then it will not be perpendicular, it will be slanted. If you take another x, x double dash here, then you see the same y bar is the nearest one. So, a single point y bar could be the nearest point of projection, points to many, many such x bars.

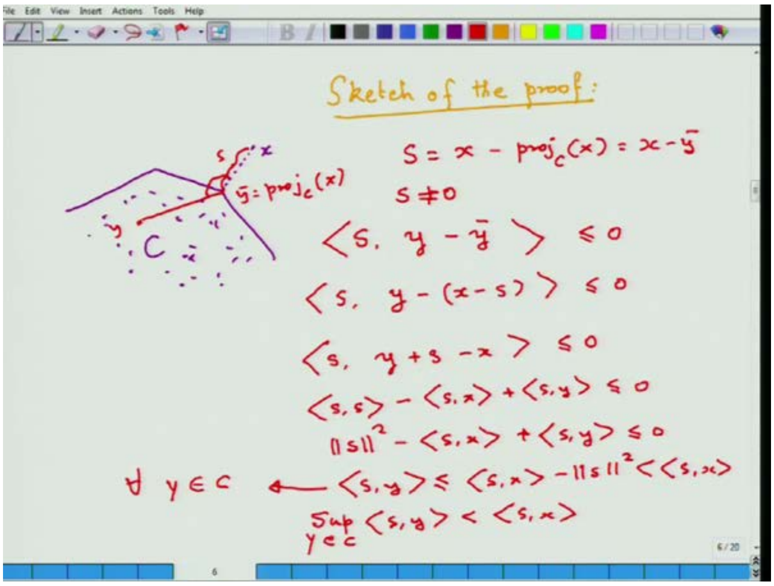
So, for any, any effector on this line, on this line, on this line and anything in between the projection is the same point y bar. For example, if you take it on this side, suppose you take some x tilde, then the projection here is y tilde, but y tilde cannot be a projection to this point, say x hat y tilde cannot be a projection to x hat, here the projection would be y hat.

So, at these rough corners of convex sets these sort of situations can arise and this thing actually forms a cone, that is, the set of all xs will actually form a cone, which will be called a normal cone. In the future this is, this we call the normal cone, but that is not exactly what we need at this moment. We will go ahead and discuss this thing slightly later, we will learn about them slightly later, but what I want to recollect at this point is that at this points, that is, at for all this x dash, x. x tilde, x double dash, y bar is the only projection point.

Now, you see if I do not have a convex set, this unique projection business does not work because for example, if you take the set, which is epigraph of y equal to minus mod x, so you take the epigraph, the set C has an epigraph of minus mod x. So, take any point on the line here, drop a perpendicular this is a nearest one drop, a perpendicular. Thus, by symmetry or from simple Euclidean geometry you can prove, that this side is equal to this side, right, because what would happen is that if this, because if this is equal to this because these are 45 degree angles, so this is equal to this, so this is also equal to this. If this is 45, this is 45, this is 45. So, what does it mean? It means this is equal to this, this is equal to this; 45, 45, the same angles. This is an isosceles triangle, this is an isosceles triangle, so this is equal to this, this is equal to this and so I have this side is equal to this side.

Now, once I know this, you show these. For this point x this y bar 1 and this y bar 2, two different points, both of them are projection points, but because this is the non-convex and so the epigraph here, you see, is a non-convex set, for a non-convex function point outside the set can have two different projection points. Such a thing cannot occur just because of the use of the factor, use of this idea on strong convexity optimization is a great and beautiful subject because these are proved in a very simple and very beautiful manner, but we want to get into all these things. So, now we will start proving what we had just discussed.

(Refer Slide Time: 26:11)



So, we will go back draw the picture again. C, So, there is a point, say, x outside and then I want to describe the hyper plane, which will strictly separate the two such matrix. Let me mention, that what I am going to give you here is a sketch of the proof, not absolutely hand waving, but not extremely rigorous either.

So, what we really need to look here is the following, that if you take any point, (( )), y inside it and this is y bar, which is the projection of x on C. If we join y with y bar, this angle is always obtuse, whatever y you take in this convex set, this will always be true. Now, this proof you are avoiding at the moment because it needs certain calculations. We should go on, so we are just avoiding, but from the geometry you can really make it clear to yourself, that this is exactly what is happening.

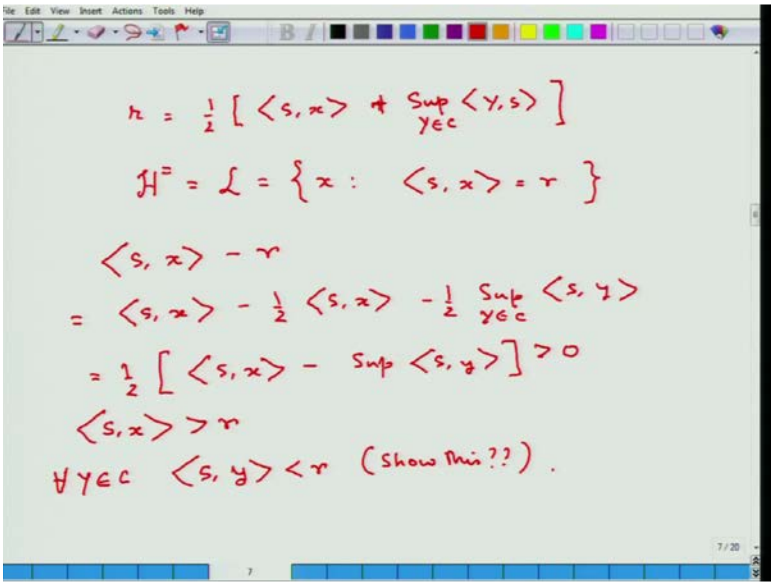
Now, let us look at something more interesting is that you see, let me denote by s, x minus the projection of C, projection of x 1 C. There is x minus y bar. Now, this s cannot be equal to 0, so obviously, x and y bar are not the same point because x is the point outside C and y bar is the point in C, because y bar is a solution of a constant minimization problem. So, y (( )) not be feasible and hence y bar is in C.

Now, we will employ this obtuse angle business that we just spoke about. In order to do so, what we do here is the following. So, s is this one, s minus y bar. Now, take any other y here, y minus y bar, this inner product. So, this is your s, this s vector, this inner product must be negative or non-positive because the angle is obtuse. You know how to do dot products? So, I can now write s of y minus, what is y bar, y bar is x minus s, so you have (s, y) plus s minus x or (s, s) minus (s, x) plus (s, y) to be less than or equal to 0. So, this can be written as norm square of s minus (s, x) plus (s, y). This simply means, that s of y is less than or equal to s of x minus norm s square. We take this thing to other side. Now, observe, that norm s is not equal to 0 because s is not equal to 0, so this is strictly less than s of x. So, this is what we actually have obtained, that now you can say. I have got a strict separation take any number, which is between (s, x), which is lying strictly between (s, x) and norm s square and that will do the job, (s, x) minus norm square and (s, x).

So, take any real number line between this real number and this real number strictly in between them and that will do the job, but we can really figure out an exact one. See, you can write this as following. Supremum, so this is true for every y in C, this thing is true for all y in C, so you can write the supremum over y in C of (s, y). This is obviously, less than equal to this, which is strictly less than this and hence, the whole thing is strictly less than s of x.

Now, this is the representation of, from this projection business, but how do I really speak about the finding the hyper plane. Let me do that.

(Refer Slide Time: 32:22)

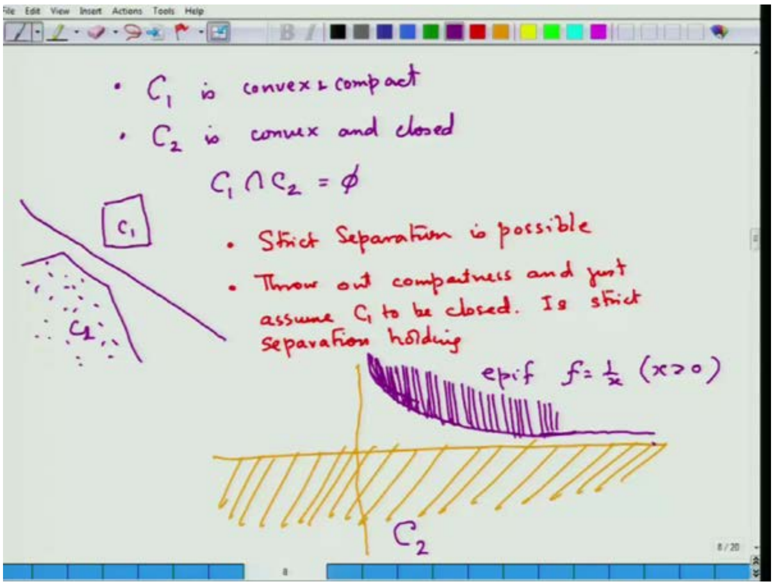


Let us find an r, which I can write as half of (s, x). So, x is given to me, s I have found, so I am writing this. Thus, the supremum, of course the supremum here has a finite value because it is bounded by (s, x) as we have seen, supremum s of y, where y is in C, this is the plus sign, of course this r is depending on your choice of s, of course, because it depends on s, is again on it; s is again depending on x, the choice of, you, the choice of the point x, so you have taken this r.

Now, this r is what we really want, that is, hyper plane or we can try to write like this or we can write as, L is a set of all x such that s or x is equal to r. Observe, that (s, x) minus r, what does it give me? It gives me (s, x) minus half (s, x) minus half supremum y element of C (s, y). So, this will become half of (s, x) minus supremum of, so this whole thing. Obviously, you know from this fact, that this is strictly bigger than 0, so you get s of x is strictly bigger than this r. So, this is lying in the, x is lying in the upper half space. Similarly, you can show, that for all y in C, s of y is strictly less than r, show this. So, this is my required hyper plane, which strictly separates. So, we have mathematically demonstrated this procedure.

Now, comes a second question. We have taken the small simplistic scenario of a point and a set, how do I extend this scenario to two convex sets because we were speaking about two convex sets always. There they have been non-empty, of course just a point x is itself a convex set and it been outside a set immediately tells, that these two convex sets are disjoint.

(Refer Slide Time: 35:45)



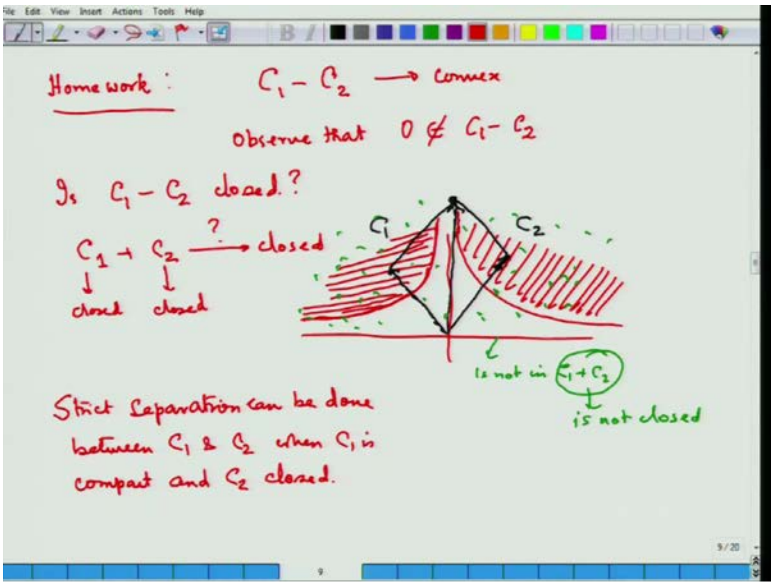
Now, let me consider this case. Now, C 1 is convex and compact, that is, it is convex, closed and bounded and C 2 is convex and closed obviously. Be aware, that in this, all these statements of separation theorem, this set C was always closed though need not be bounded, but this set x, single point is, obviously, closed, bounded and convex, so it is a compact set. So, I am making the first generalization of this by replacing this x by some compact set. You will see why such a, such first step of, first step generalization to include compactness is useful.

What I am trying to say is that if this happens and C 1 intersection C 2 is empty, that is the disjoint, so you have a convex set like this C, which is closed C 1, sorry, C 2, which is closed, but need not be bounded and you have another like this C 1, which is closed and bounded. What we are trying to say is that then strict separation is possible. So, the thing is or conclusion is strict separation is possible. Now, you want to ask the question if I, throughout compactness from C 1, strict separation is that possible or not? So, our question is, throughout compactness from C 1, compactness and just assume C 1 to be closed is strict separation holding? The answer is of course negative, otherwise we would not get into all this issues.

Let us see why that answer is negative. Take the x-axis and y-axis, take this plane, this half space along with this line x-axis, consider this to be C 2. For C 1 consider the epigraph of the convex function y equal to 1 by x for x greater than 0. Of course, the domain is x greater than 0. So, this is the epigraph of f, where f is equal to 1 by x, where x is greater than 0. Now, you see, this 1 by x is asymptotically going towards the x-axis, asymptotically going towards 0, that is, as you go further and further towards infinity, more and more towards infinity, this distance between this x-axis and this curve keeps on decreasing, but they never touch each other.

So, now, if I draw any line other than the x-axis to separate this epi f and C 2, then no matter what line I draw, there will be a point where it will go and cut the epigraph of f because it is asymptotically going towards 0, as the result of which even though these two sets are closed, then they are having empty intersection. You cannot have a strict separation here; the separating line is this x-axis, the yellow line. So, that is why compactness is so very, very important. So, how do you go above trying to prove this fact? So, this would be a partial homework for you, not a complete homework.

(Refer Slide Time: 40:55)



Now, observe, consider the convex set C 1 minus C 2. Now, of course it is convex, you can prove it, I do not need to do that. Now, observe, that 0 does not belong to C 1 minus C 2, why? If 0 belongs to C 1 minus C 2, there is a common point between C 1 and C 2 because then there exist x in C 1 and y in C 2 such that 0 is equal to x minus y and hence x is equal to y. So, both are in x, is in, so these are common elements of C 1 and C 2, but C 1 and C 2 are disjoint, so this cannot happen.

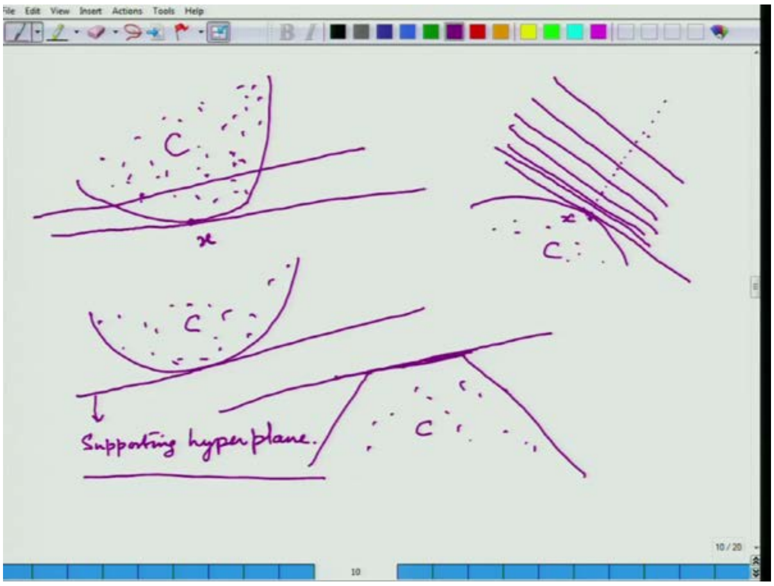
Now, you observe, here is a convex set and there is a point outside it and if this set is closed I can apply what I have learned just now. The question is, is C 1 minus C 2 closed? So, this brings us to this, to the following question. So, if C 1 and C 2 are two convex sets and both are closed, the question is, is this closed? Is C 1 plus C 2 is closed? This answer is not true in generalness, in general unless one of them is compared.

Here, we have assumed C 1 is compact. So, C 1 minus C 2 is finally closed, but here C 1 plus C 2, in general minus is nothing, just we can place a plus of minus C 2. So, C 1 plus C 2 is not closed unless one of them is compact because you take again, an example of this type where you take a convex set like this y is equal to 1 by x, this y equal to minus 1 by x.

Look at these two, so take any two points, so take a point here and take a point, take a point here. So, this is my C 1 and this is my C 2, both are closed and both are disjoint from each other. In fact, you see, of course disjoint is not needed. I want to construct C 1 plus C 2. So, I take a point here, I take a point here, I do the vector addition, let this, take this point here. So, I do the vector addition I get the point here. So, what I am doing is, that if I keep on doing all points here, like this sort of combinations, I will basically get this set, this whole half space, but I will not get this line red line. So, this red line is not in C 1 plus C 2. So, but this, hence this is part of the closure of C 1 plus C 2, closure of the half space, so as a result, C 1 plus C 2. In this case, this one is not closed, but in our case, because we have assumed C 1 to be compact, so C 1 minus C 2 is closed and of course, now you can go ahead and apply the result, that you already know. So, that application has to be done as homework, so finish this thing as homework.

So, strict separation can be done between C 1 and C 2, C 2 and C 1 is compact and C 2 closed. There is another very important conclusion from this; all this is the study of the supporting hyper-plane theorem.

(Refer Slide Time: 45:54)



So, if you take a convex set of this form, take a point here on the boundary, you can always draw a tangential line through the boundary so that this hyper-plane creates a half space and the whole set C is in one side of the half space. So, here is the convex set and now here is the point x on the boundary. I can always prove, that there is a hyper-plane passing through this point x, so that the whole set C is in one half space, it does not have two half spaces. For example, if I take point inside and if I try to draw a line, again draw a line, then it will set C would be covering both the half spaces. But here, in this case, where I have just the set C and a point here and I draw a line here for example, tangential in this case, as it looks like in the picture, then C is in one half space, it does not come in spread over to this half space. This whole thing, this is called a supporting hyper-plane.

So, what does it say is, that if you give me a closed set, then I can always have a supporting hyper-plane. So, what is the idea behind this fact? The idea is as follows, that if you have a closed set, so you know, this supporting hyper plane does not mean, that it will touch only at one point, for example, here I can have this itself. So, this is a convex set, so this whole side can be a part of that supporting hyper plane.

The idea essentially is as follows. Suppose, you have a set C and because you take any point x in the boundary, so given any point x in the boundary there is a supporting hyper-plane passing through that point, that is exactly what the theorem says. So, what it say is that if, because this is in the boundary point, any neighborhood you take, there is a point in C and a point outside C, that is a meaning of a boundary point and as a result, you can have a sequence of points coming and converging to x.

Keeping in view, mind, keeping in view the audience, I would not like to go into the details of the proof because it would require more sophistication and would really be a proof for mathematic students. So, I am assuming engineers are also viewing this.

So, what you can observe? If you keep on drawing this sort of balls around x, there will be a point inside x and there will be a point outside x. So, those xs, those points, sequences, if you think of it as something you are hopping one after another, you can hop on, on this point and come on to x. For each of this point x is actually a point outside the set C, in C complement and from there you can have, strictly for each of them you can have a strictly separating hyper-plane and finally, you can come in like this and come and touch the boundary at x and that would give you. So, basically what would happen is that, that would generate a sequence of planes, which can come and touch in x and that would give you the separating hyper-plane or the, sorry, not the separating hyper-plane, the supporting hyper-plane to C at x.

So, with these very basic ideas about separation we close our discussion and go into more of the optimization issues in the next class.