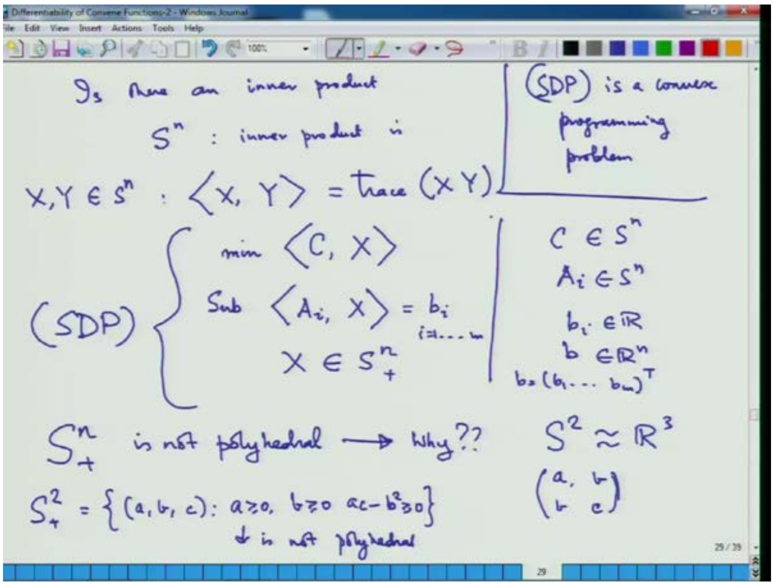
Convex Optimization

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

Lecture No. # 34

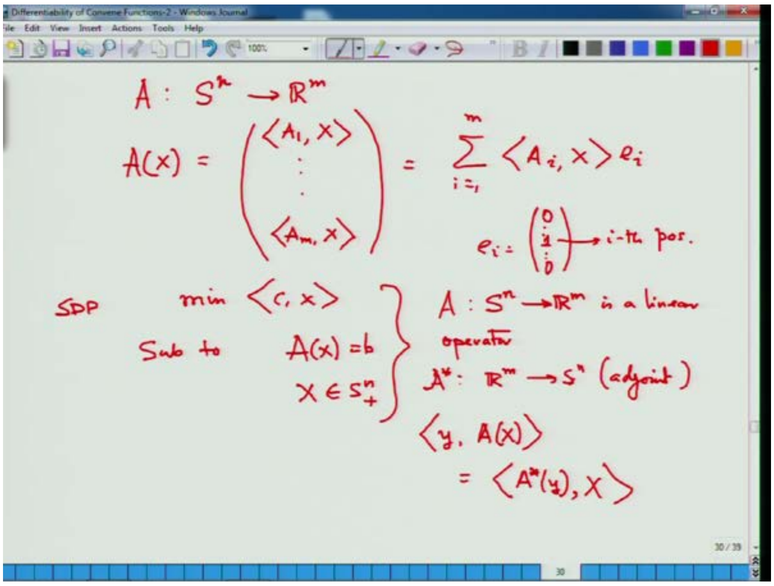
Welcome back to the lecture on convex optimization, and we are on the last leg of our lecture, so we are covering a very very important area, and very hot area of convex optimization called semi definite programming; we are a we are see not going to cover each and everything that is there in this vast and beautiful subject, but we are going to touch on some main things like optimality, duality, and we are going to also mention the major results rather than getting into details. And we will provide some examples and give a brief idea of how interior point methods can be developed for them. So, this is in general our plan for semi definite programming, and we will show that many important important and interesting problems can be posed as a semi definite programming problem, and SDP can be actually solved out.

(Refer Slide Time: 01:15)



So here is a semi definite programming problem as we have seen in the last class, and where S n is the space of symmetric matrices, and the inner product is nothing but trace of X,Y, and these would lead us to define a problem, which looks almost like a linear programming problem in matrices, but since S n plus is not polyhedral as we have discussed, this is not exactly a linear programming problem in matrices, but a convex programming problem in matrices, so basically a solving a linear matrix function over a convex set in S n, in S n plus rather. So, if you look at this expression, these constraints A i X equal to b i, let me consider the following linear operator.

(Refer Slide Time: 01:56)



The following linear operator A, which takes an S n and gives puts in R m. So, it is A of X.

(No audio from 02:08 to 02:22)

So if I look at this, then what is my semi definite programming problem; if you look at this, I can write this also as…

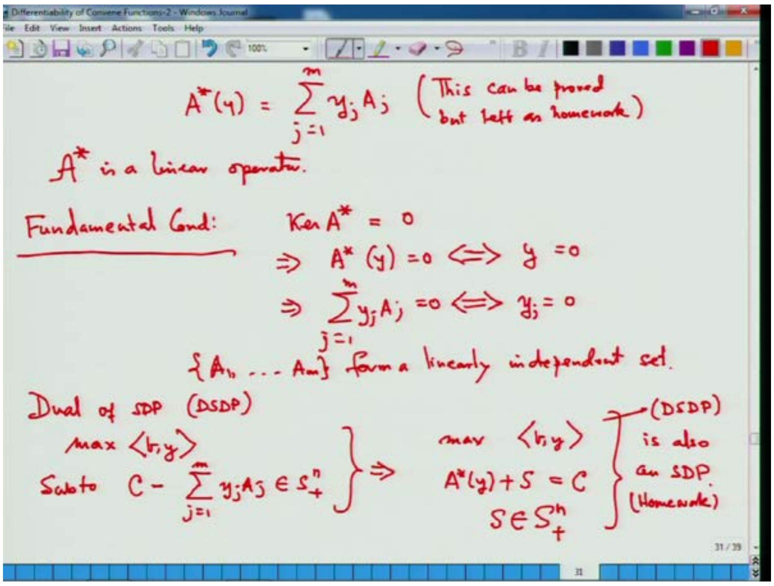
(No audio from 02:31 to 02:41)

See here you know is the vector, whose i eth position is 1 and rest are 0. Now, once I know this, I can write my SDP problem as to minimize C, X subject to the linear operator A (X) equal to b, while X is in S n plus; now this A be in a linear operator, this is A, this A is a linear operator.

(No audio from 03:27 to 03:40)

This is a linear operator, so we can define the adjoint to this operator A star, this is obviously this adjoint operator is unique, so which is defined as follows. So this inner product A (X) is in R n is same as A star (y), so A star (y) takes an element in R m and max it in S n and that you so this is nothing but trace of A star (y), X this thing; so they here through this adjoint, we are linking the ordinary dot product in fundamental spaces with the dot product of or the inner product of the spaces.

(Refer Slide Time: 04:47)



Now, it can be shown that A star (y) can be actually written as this is what A star (y) can be written as, this can be proved, which we do not proved proved, but left here as homework. Now, what is happening is if you look at this thing, this A star (y), A star is again a linear operator, because it is a adjoint of another linear operator; a fundamental condition that we will use, which is (( )) full rank condition a linear programming is that kernel, which implies A star (y) is equal to 0, if and only if y is equal to 0, this implies summation y j A j j is equal to 1 to m is equal to 0, you know, basically A j is form a linearly independent set, of this result would become crucial in establishing some duality and I would say duality and linear optimality of this linear programming problem.

So, what about the dual of this problem that is very crucial and dual is something, which we have already shown earlier.

(No audio from 07:07 to 07:15)

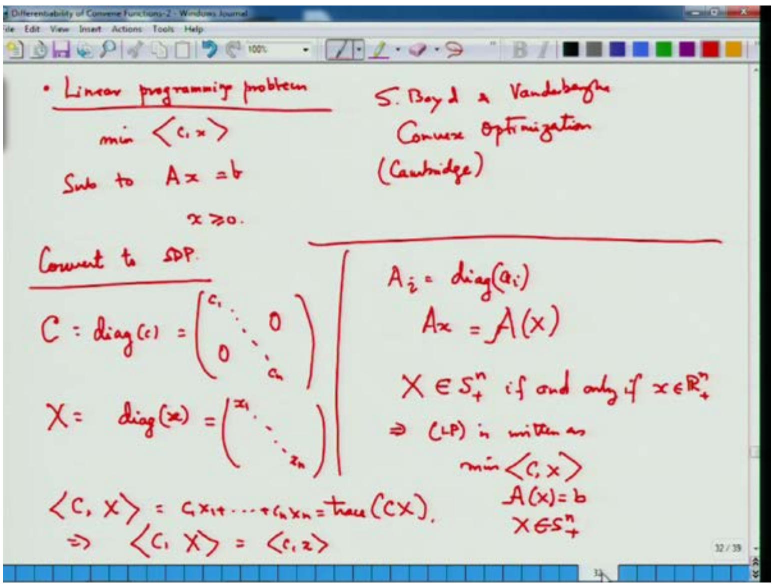
The dual, which we denote as dual of SDP is max b, y C minus j equal to 1 to m y j A j, this is element of S n plus, so another equivalent way of writing is so there would be an extra semi definite positive semi definite matrix, so I can write this as max of such that this can be treated in slightly better way. So, as a huge amount of mimicking the a standard linear programming problem, but at the end, this is not really linear programming problem, and needs very different techniques to solve them.

Now let us look at some selected special cases, which can be converted to semi definite programming problem; for example, linear programming problem.

(No audio from 08:37 to 08:50)

This is actually an SDP problem; let us look at this problem, here you see, I am I am not yet expressed this in the standard SDP form, what we can do or rather show that as a homework you show that DSDP is also an SDP; DSDP is also an SDP, so this is something, I leave as homework which will be just little bit of fun to prove. Now, what I have not shown here there is that what is the S vec type operation, so I will here, what will not do here, we have not shown you, what is the is you know, what is the wake operation of converting a matrix into a vector, so there is something called symmetric way K, which converts a symmetric matrix into a vector, we have not said anything about that, but we will possibly say it later, but now let us look at the linear programming problem.

(Refer Slide Time: 10:01)

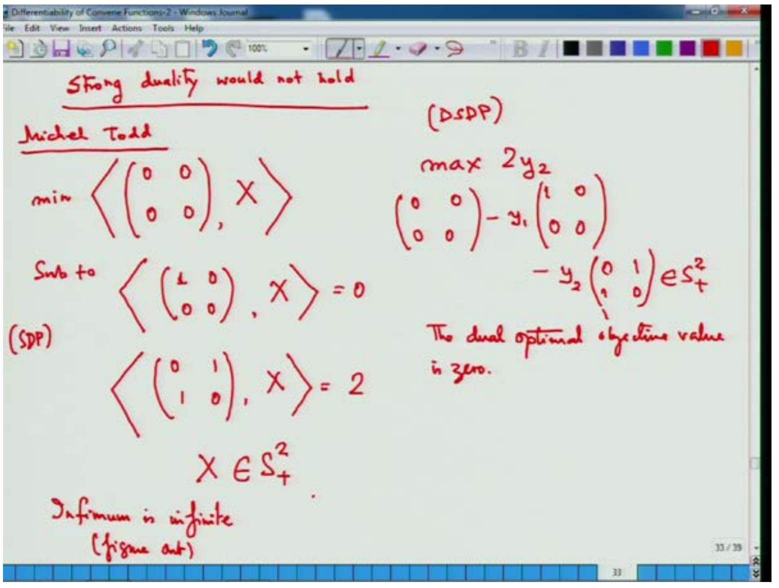


So again have the standard LP problem, minimize subject to A X equal to b, and x greater than equal to 0. So we will convert this to an SDP; how do we convert to SDP that would the case. So now, if we just said that we want to convert this linear programming problem into an SDP problem, so we will write C is equal to diagonal matrix of C, which is nothing but the matrix C n, and capital X diagonal of this vector X, so which, so in this sense C of X is nothing but C 1 X 1 plus C n X n, because if you multiply this vector C X is multiply this of this is same of trace of C X, so you multiply C with X matrix multiplication, this you will get C 1 x 1, C 2 x 2, C n x n, and this is trace of X; but this so this would imply that C of X is C x, which is here; now we are left to conclude about the remaining part that is write A i is diagonal of A i. So you take the i eth row of the vector and make a diagonal matrix.

So A of x would be same as a capital A of X of del A of X; so basically then we will have so capital X is element of X n plus if and only if small x is element of R n plus, this is capital clear; when X is equal to diag (x), so this means I have been able to convert my problem, so L P is written as… so if so if you solve this SDP problem with X is this, C is this, and you have actually converted a linear programming problem into a semi definite programming problem. So this is (( )) even coordinate optimization problems can be also converted to semi definite programming problems, and they are very helpful in many, many cases.

Now, I just want to mention a book by Stephen Boyd and Vandenberghe book; name of this book is convex optimization, of this by Cambridge and so very fabulous book, which deals with application of SDP. Though it looks like a linear programming problem in matrices, we have already mentioned that this is nothing but a convex programming problem; it is not in general a linear programming problem in matrices. So what would happen is that strong duality theorem that is the duality between this dual of SDP and the original SDP, there need not be strong duality; so unlike linear programming, see you see there will be little difference strong duality would not hold.

(Refer Slide Time: 14:43)



(No audio from 14:40 to 14:53)

So, we are going to prove some… we are going to put some examples, showing their strong duality does not hold, and this example is due to Michel Todd, a very famous optimization theory some kernel. So, we are going to give this example, which will show that the supremum and infimum of a semi definite our semi definite programming problem and its dual; the infimum of the SDP problem and the dual, they are having two different values; so here again, we establish the claim, that in really not linear programming, because if it was linear programming, then one could have actually caught strong duality.

(No audio from 15:44 to 15:58)

So you minimize this, so you can observe that these are all positive - symmetric positive semi definite matrices for example, this one, it has a greater than equal to 0 Eigen value…

(No audio from 16:18 to 16:32)

(( )) matrix…

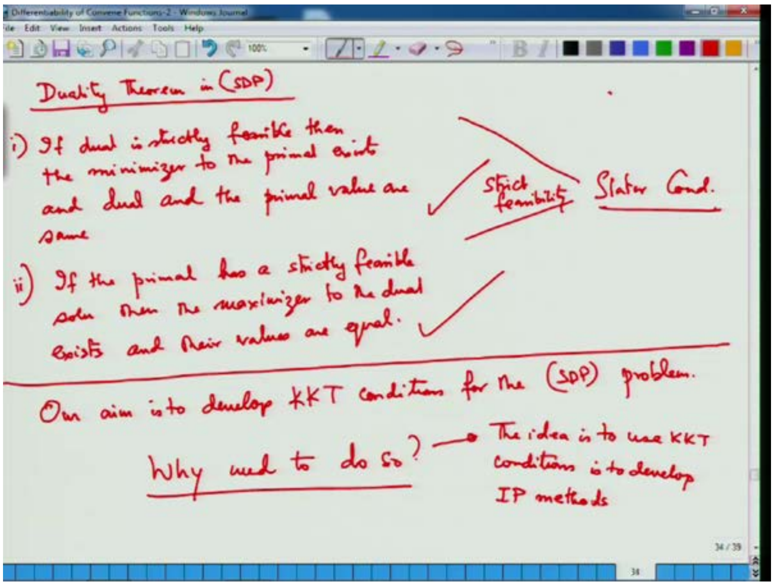
(No audio from 16:35 to 16:48)

So this is Michel Todd simple semi definite programming problem, and we are now trying to find the dual. So the dual of SDP, so this is my given SDP; I am writing the dual of the SDP, so you have to maximize twice y 2, you see there are two constraints, so there will be two dual variable, number of functional constraints and the number of dual variables are always same; just go back, and have a look at the top on Lagrangian multipliers KKT condition and all those things. So, here…

(No audio from 17:36 to 18:05)

y 1 a 1 y 2 a 2, this, and C minus summation y i this C and summation y i is element of X 2 plus. So, this is something important. Now, infimum of this problem is infinite is unbounded. So infinity is the infimum of this problem, so you have to figure this out and for this problem, the supremum the or the dual optimal objective value is 0.The dual optimal objective value is 0, so there is a duality cap, you see the cap cannot be reduced unless there are certain conditions; so we will now mention few duality results in SDP, and they will be mention through this book, which I can want to show these are very, very useful book at the undergraduate and beginning graduate level in optimization.

(Refer Slide Time: 19:42)



Now, duality theorem in SDP; so if the dual problem has the strict feasible solution, strict feasible point; can you understand what is the meaning of strict feasible point, in this case, you can translate what we have learnt for linear programming, where is C minus summation y i A i should be in the inter S n plus or S n plus plus that is the meaning of dual dual problem as the strict feasible point y, then the minimize of the primal point exists, and the dual and primal value are same, (( )) is the primal problem as the strict feasible solution. So in case of convex optimization, the corresponding notion is that of Slater condition; so if the primal problem has a strict feasible point X, then a maximize that to the dual problem exists, and their values are equal; this sort of results are also true in case of general convex programming problem or the second one for example, does really hold for the general convex programming problem.

If both problems have strict feasible solution, then both of optimizer and whose value is so and so, so if the dual has the strict feasible solution, then the primal achieves its minimum, which is very interesting; while if the primal has a strict feasible solution, the dual achieves its maximum, and that is the very interesting thing. So, if dual is strictly feasible, then solution minimizer to the primal is attained or you can guaranty the minimizer to the primal exists, that is what it says, and the dual and the primal value are same.

(No audio from 22:02 to 22:15)

So its dual is strictly feasible, then the minimizer to the primal exists, and the dual and primal value has same, this one the first result, so second you see we are talking about… So we are again going to tell the second one, if the primal is strictly feasible, and then the dual is attained and the values are same. If the primal has a strict feasible solution, there is X in X is element of S n plus plus, not just in S n plus; there is actually where strictness comes in, if the primal has a…

(No audio from 23:56 to 23:07)

So, if primal has a strict feasible solution, then the minimizer to the dual exists.

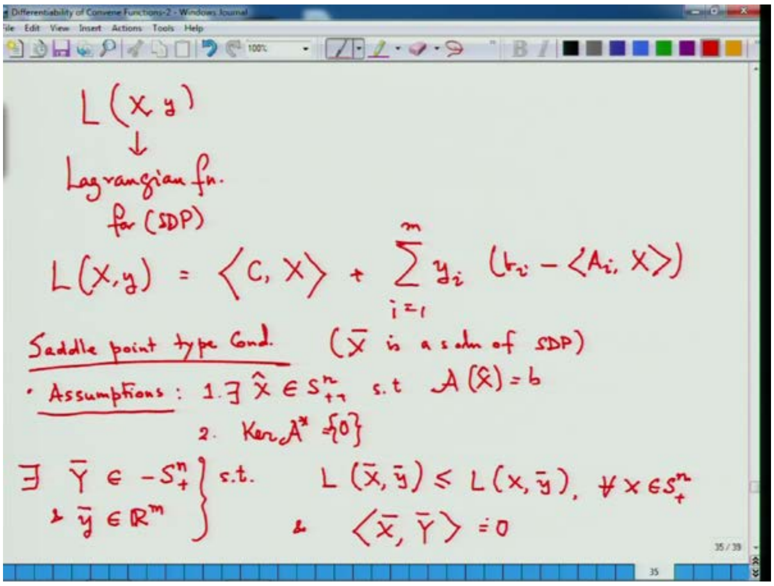
(No audio from 23:13 to 23:24)

So the maximizer to the dual exists, I made some mistake, so maximizer to the dual exists, and the values are equal; maximizer to the dual exists, and the values are equal. Third one, we will not want to write, which says that if both are strictly possible, then both have optimizer and so on so forth. So, these are two important conclusion when you look at SDP, so it is none none like a linear programming conclusion, where you state conclude, that if both the primal and dual problem have a non-empty feasible set, then their solution exists and the basically their solutions coincides.

And one has to observe that linear programming if both primal and dual have a feasible solution, the strong duality automatically holds; that is if you just if you want to detect the primal dual… if you want to detect the primal feasible point and a dual feasible point, then you have done, then basically you know that this linear programming has a solution, which is very, very important. So, you just have to know this little part of semi definite programming, not not in much detail; so now our aim is to develop KKT conditions, now aim to develop KKT condition for the SDP problem; before we proceed to do, so let us tell that why that this strict feasibility is here, actually strict feasibility is corresponding to slater condition.

Now, our aim to develop KKT conditions for the SDP problem; the question is why we need to do so; the answer is that if we can develop a KKT condition for the SDP problem, then if I can solve that KKT condition using interior point techniques, then I can solve the SDP problem. The idea is to use KKT conditions to develop interior point's techniques for SDP problem. Remember all the while, this is an important class of problem with lot of applications coming in you should see the book of Stephen Boyd to see more detailed applications, but you must be very, very careful that this is not a linear programming problem in matrices. So then you get a huge literature on the net on semi definite programming, so idea to use the KKT conditions is to develop IP or interior point methods.

(Refer Slide Time: 27:06)



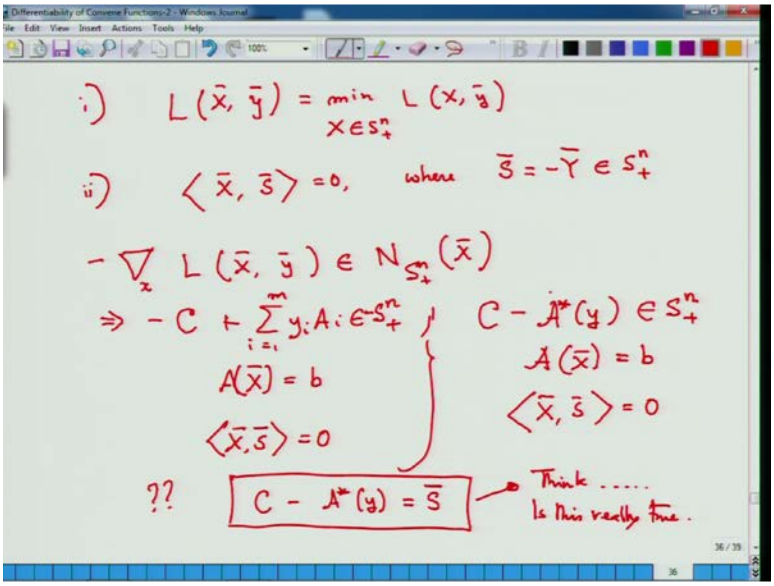
So, in order to do so, we will first construct the Lagrangian - Lagrangian function associated with the SDP problem; Lagrangian function for SDP; so this is we will denote like this, and L of X, y is C of X plus summation y i b i minus A i X; so if this is what I have, we have the following problem. So, we are first going to develop a saddle point condition, before we talk about a derivative in this particular case, what do I have is follows; you will see that we will be able to tell something more than general convex programming problem.

So let us do a assumptions; the assumptions are as follows; the assumption number 1 is that there exists X hat element of S n plus plus, which is not really required, which is true; so there could always be element of S n plus plus such that A of X hat is equal to b; number 2 kernel of A star is equal to 0 - 0 vector. So once you have this, we can prove that there exists of Y bar element of minus S n plus such that L of X bar, Y bar; so of course, we have assume that X bar is a solution of SDP that this…

(No audio from 29:50 to 30:04)

Now, this been some sort of a linear function, again in the set space of matrices; so you know here we have in general, the Y bar there exists Y bar element of X n plus, and I forget to tell you an Y bar element of R m, because this Lagrange multiplier associated with equality constraints will not have a sign. So there would exists these two quantities such that this will be true and X, bar Y bar this is an additional information is equal to 0.

(Refer Slide Time: 30:46)



Now what does it show that L of X bar, y bar is equal to the condition, I can now write them into this equivalent form minimum over all X element of S n plus L (capital X, y bar), this is what you have, and subject to X bar S bar is equal to 0, where S bar is equal to minus capital Y bar is element of S n plus; I can write it like this. Now the first condition obviously this Lagrangian is differentiable when linear, you would have gradient of X of L X bar, y bar is equal to 0. So if you have a gradient of that, then it will become C; if look at the gradient again, so it will be C plus your taking gradient with respect to X. So basically, it will become C, you have to do y i b i, and this will get into summation y i A i. So basically, what you will have is C minus summation y i A i; so if you take the gradient, this is what is coming.

An obviously, X bar has to satisfy… So this and you also have X bar, S bar is equal to 0; so these condition this can be written as C minus no this cannot be exactly equal to 0, its element of S n plus; so I would ask you to figure out, now this cannot be grad X equal to 0, it should be grad X is element to normal cone to S n plus at X bar, an negative of this; the negative of the gradient is this. So normal cone is minus S n plus, it will become minus C plus this, so it will become C minus A star y, this is now element of S n plus, and A X bar is equal to b; and we have this additional condition, we can X bar X bar equal to 0, but what is my and also I have this additional condition that X bar and S bar is equal to 0.

Now, how do I prove that this is actually equal to S bar, so my problem would be solved if I can show that this is exactly S bar. So this S bar, so if I can show this fact, then my problem actually is finished, then I have actually solved the problem. So let us leave at this stage, and ask you to think about this thing, is this really true? And then, we exactly have a similar looking KKT condition; of course, I have not done many details, I have just writing down certain things, so then you can form the basic conditions. Now you observe that this is what you will have, and the normal cone here is nothing but S n plus, a thing which you one might not able to understand, so those who are not (( )) mathematically involved, do not get too much bogged up with this particular issue. So y X bar into X bar would become 0 or whether this is, whether is this true? This is the question I want to ask, and tomorrow in the next class, we will address this question I will end for the day.