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

Lecture No. # 19

So, welcome today to this (( )) a lectures on convex optimization; and today we are going to talk about a very, very important class of convex optimization problems is the linear programming problem. The linear programming, the growth of linear programming had been a phenomenon success story in the history of mathematical optimization. Linear programming came into the scene in the late forties and dominated much of the seen in the mathematical programming possibility in the need of weight is; when the interior point methods gave it not only a further boost, but also bought in to the four in lot of other problems, which were modeling applications quite much in a much better way than linear programming problem. But the influence in the techniques of linear programming in optimization continues to remains so, and possibly linear programming is this one of the single most important aspect of convex optimization, which is much well understood.

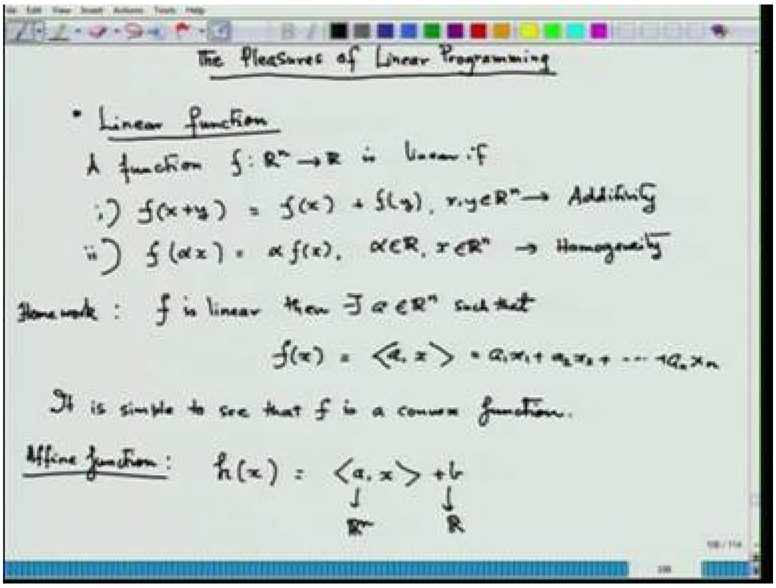
So, that is why we would call this series set of lectures? So, we though there is separate lecture on linear programming in this NPTEL series given by Professor. Prabha Sharma. She was my ex collegian. I just want to let you know that, I want to give this lectures in order the repetition of what should given rather give you completely different point of view more from of on the point of convex optimization rather than the separate point of view one has for linear programming, because linear functions are special class of functions; linear functions are special class of functions, and so for linear programming problems, there are separate techniques for solving them. Though will show one of the major techniques simplest method would be discussed here, but in a approach which is completely different from the standard approach that you see in the books and literature associated with linear programming, especially the other way under graduate level.

Our approach would be more from point of view non-linear optimization non-linear convex optimization, and we will see what impact one can have; when we bringing those methods to study in linear programming. So, will spend up part of what time discussing linear programming not the whole of time; rather I would say in optimization that two approaches, two types of optimization one is continuous optimization another is discreet or combinatorial optimization. What we had been discussing till now, she is essential continuous optimization and in discreet optimization you have up feasible set could be discreet, but having a large number of points.

So, point by point a numeration of the function value becomes very difficult. So, one is to finite easier way to go to the minimum value. So, that me exact subject much more harder continuity gives you lot of ground, because convexity gives you ground, because of its connectedness, if you point you neighborhood or grounded, and (( )) so that this is lot of plus point of and try to understand optimization by learning continuous optimization.

So, what we having to do in this set of lectures is; to give the continuous optimization point of view of linear programming. Though we will talk about simplest method, but it will be form of very, very different angle. So, but I am sure you will a fun throughout as you learn this part, and that is why call this series of like of set of lectures, this specific part for linear programming as pleasures of linear programming. So, this start with programming were the objective and constant functions are linear or affine; I would rather go head and first start you giving a little bit of idea of what is a liner function, affine function, and then so you have already haven idea about linear function and affine function, but just recollect and strongmen you understanding have just do that.

(Refer Slide Time: 04:44)



So, first we start with the definition of linear function.

(No audio from 04:43 to 04:57)

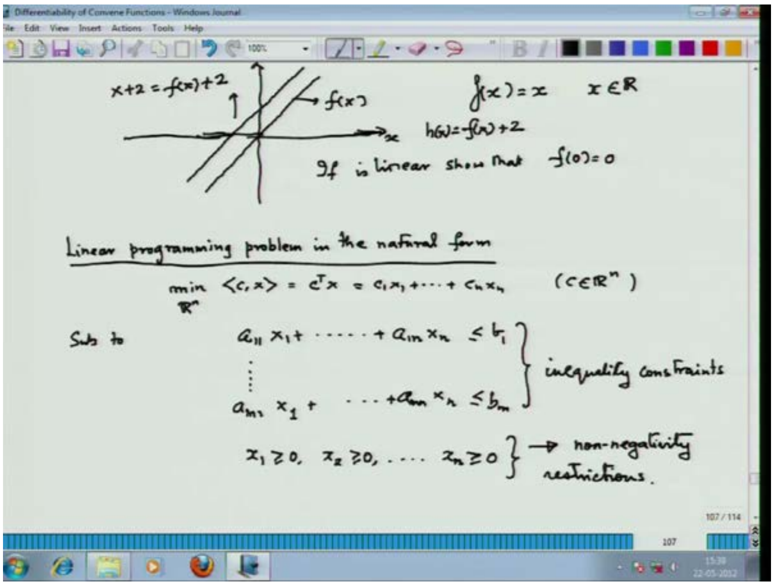
A function f: Which takes an element in a R n in a places in R is linear, if the following two conditions hold. First one is condition of additively that is vectorically add two elements in R n. Then the functional value at the new element x plus y is same as the functional value of f x added with functional value of y added as real numbers. See we could define this from R n to R n, R n to R n does not matter, but we are doing its R n to R, because that is exactly what we are interested in.

Second property is… If you take any alpha element R taken any x element R n… f of alpha x if you made take of scale of multiplication of x with alpha, then operate the function there is same as taking the function of value f x and then multiplying it with the alpha. So, this is true for any pear x alpha, in R n this true for any pear alpha any alpha in R n, any x in R n. So, this property is call the additively property. And this property is called the homogeneity property. Now, it is very important to understand that thus this function; you know have a particular representation is something much more specific then this sort of definition.

You can actually use very basics affects from linear algebra and live you as a home work to show that; if you give me f, if f is linear then there exists a in R n. Such that, f of x is nothing; if any evaluate the functional value f at x is nothing but; inner product of this a or dot product of this a with this vector x. Which is nothing but a 1 x 1 say a 2 x 2. So, is a 1 x 1 plus a 2 plus x 2 plus a n x n.

So, this is the standard definition of dot product; which all of you know. Now, can so, if it is simple to see (No audio from 07:56 to 08:05) that f is a convex function. So, now, we come to the definition of an affine function.(No audio from 08:20 to 08:30) So, in our definition of affine functions; affine functions simply means, that you have a linear functions affine function h x is a linear function f x which is a times x plus some number b. So, a is in R n and b is in a R.

(Refer Slide Time: 09:11)



Now, it is very, very important to note, the geometrical picture of this two things linear and affine function. Let become to r 2. Now, if I look at a function f x is equal to x function x is in R, then this represents a linear function. So, the line f x is equal to x is line through the origin x some this is your f sorry. So, this is a origin. So, this is your x equal to f x. Now, if I say f x is equal to… so, h x is equal to f of x plus 2, then you add two to all of this. And there are line whole line get shifted. So, this is your f x, and this is your f x plus 2, this is your f x and this is your f x plus 2.

So, this is a linear function, and this is an affine function. The linear function always process through 0. Infect if f is linear; show that so, homework, but f 0 need not be equal to 0, f 0 here is 2 for function f x plus 2, that is this is not this function nothing but x plus

2. So, geometrically there is a shift, so you translate or a linear function you get in a affine function.

So, now, let me write down the very basic, or canonical or natural model I would say; our linear programming problem will give some examples, of how what are linear programming in problems? First site let us write down the mathematical thing and then we will try to develop some examples. So, linear programming problems in that can natural form or canonical form.

So, linear programming problem (No audio from 11:24 to 11:38) in the standard or natural form; is to minimize c, x in a product of c n x, which is also return as c transpose x, c n x of vector this is same as c 1 x 1. Of course, c is element of R n and minimizing our R n, and though subject to constraints following constraints.

(No audio from 12:18 to 12:40)

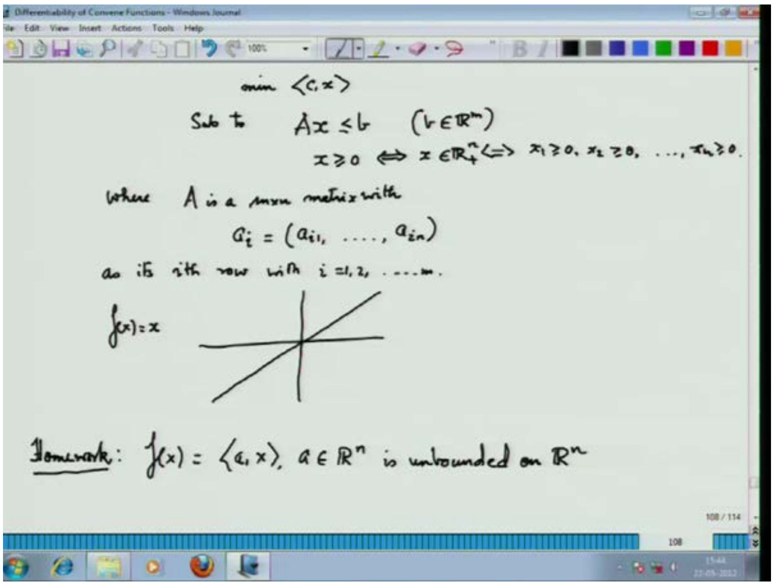
So, here are my m first m in equality constraints, and because usually linear programming comes in a applications were the decision variables are non-negative usually not always. So, this is quite of n standard to have this n, additional inequality constraints call the non-negativity restrictions.

So, usually by you are going by the lingo of programming, these are called in equality constraints. And these are called this x 1, x 2, x n these are called non negativity restriction. In fact, even if you problem has not got restriction from the variables, we can still post them as a problem like this will come to this later on.

(No audio from 13:50 to 14:08)

Now, there is a important thing, I can write this thing in a much more compact way, suppose I consider a metrics whose row is a first to rise a 1 1, a 1 2, dot, dot, dot, a 1; and m the row is a n dot, dot, dot a sorry a m 1 dot, dot, dot a m n. Then they form matrix, and this can be later express the matrix multiplication.

(Refer Slide Time: 14:54)



Now, so, I can write the much more compact forms, I can write this as minimize (no audio from 14:51 to 15:06), where A is a m cross n matrix; with a i is equal to a i 1, a i n as its i th row, row with i varying from 1 to n. Now, there is important question. What truth I mean, by this science? This science means, component wise greater than or lesser than. So, A x is a vector; taking a vector in R n to R m. So, that is, what is the m cross n matrix, and b is element in R m.

So, this b… Now, what I want to say; is that the first component of this vector A x is less than equal to b 1, once again component less than b 2 and so on. And x 1 greater than equal to 0, x 2 greater than equal to 0. It is written in a compact form. Same as writing x is in R n plus… Is same as writing… This is n component, it as a compact representation.

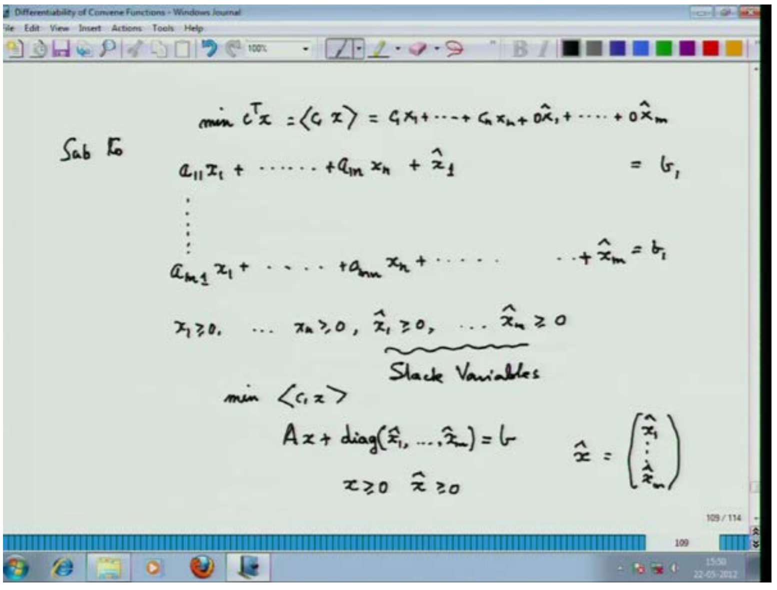
Now, the important question, why you need constrains to define a linear program or linear programming problem? This is some time see usually called LP. LP in the canonical form; so, I will just say LPc. So, why you need constraints to define this problem? Because if you have a linear just ordinary constraints problem, ordinary convex problem; say x square. You can obviously minimize our whole R. But why cannot you do it; what happens? Why you need to dissertation thing for linear programming problem?

The question is a very simple, and answer is also simple. The answer is the following, if you look at the function f x equal to the x. Let us do a Slater experiment, whichever you go this way or that way, whichever you where you go, if I go made here x; go along x infinity. My function value becomes larger and larger. If I make x go to minus infinity, my function value become lower and lower.

What you can prove, which is again your homework. f x equal to say a x; a element R n is unbounded an R n. (no audio from 18:24 to 18:33) So, this is your homework. f x equal to a x is unbounded an R n. Now, I will show you that you can still this in a much more simplified to it, because; if you observe that solving inequalities is the very difficult wall, if it is not easy; so, any solve inequalities.

So, we would rather solve equalities rather than solving inequalities, it is because equalities are much solve easier to solve; you know that Gaussian elimination, and some take other techniques for which can do something. So now, can I write this at least this part in the form of equality, it is unanswered is s. So, we go back and try to do that. So, now, I can write this as follows.

(Refer Slide Time: 19:19)



(No audio from 19:18 to 19:33)

So, these just we complete the writing, 0 into… And then do the following, subject to…

(No audio from 19:50 to 20:18)

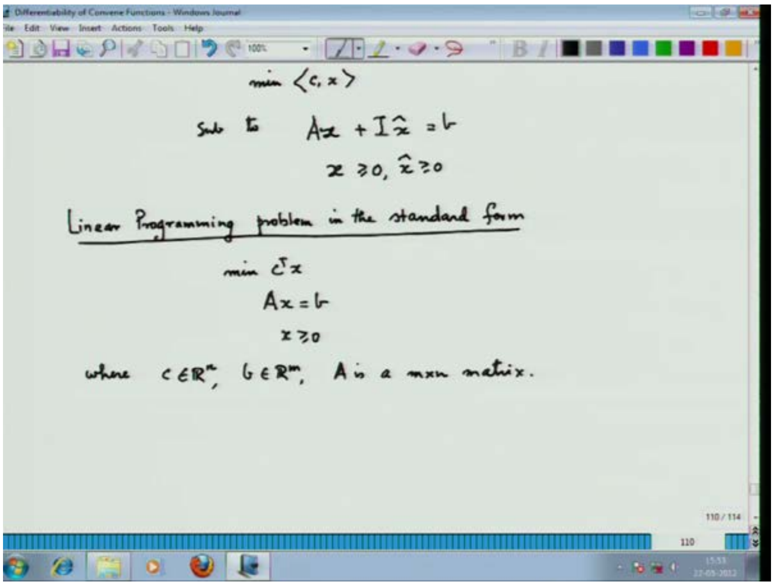
Sorry, this you might not be able to read very clearly, it is a 1 n; same thing what I return earlier, this is a m n. Sorry this would be equal. These variables are slate variables. So, in earlier description, this was less than equal to b 1. So, if I order non negative quantity, I can make it equal to b 1. So, I have x 1 bigger than 0, dot, dot, dot; x n bigger than 0, x 1 hat bigger than 0; x m hat bigger than equal to 0.

So, these variables are call slake variables. More precisely what does happened here; is that, I have a diagonal matrix, which has; I added a diagonal matrix, which diagonal consists of some non-negative numbers. So, I can write this as minimize c, x; subject to A x plus diagonal... Diagonal matrix; when you write a diagonal matrix, you do not write the whole matrix, and write the diagonal and put the 0's in one side. But you just write what is A in the diagonal, because it is understood that everything else is 0, every other element is the 0 entry.

(No audio from 21:56 to 22:16)

And x hat; where, x hat consists of the vector. Now, this greater than equal to 0; is comparison in R m, and this greater than equal to 0; is comparison in R n. So, we can write it much more fully as follows, we can write this as…

(Refer Slide Time: 22:45)



(No audio from 22:40 to 23:02)

So, you see a converted in a problem into an inequality form. So, this leave from a inequality form to any equality form, at least in the main part of the constant. See this constraint usually called in the literature soft constraints, for is which would become clear later on in the course. And these are called hot constraint, which are very which we cannot avoid in some sense. It is very difficult to replace this constant. Here some sort of penalization can be done, some sort of way we can take of this constant, and put it in the objective. But, these constants are not easy to take of as, we as seen in while we constant at the Lagrangian functions.

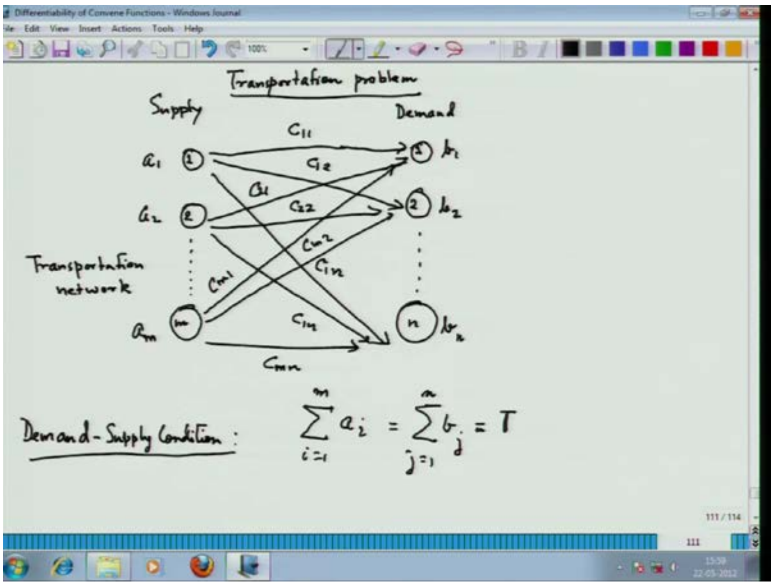
We have never touched this constant, we have touched these constant, which are call the soft constant. We will are never touch the hot constants. You can touch the hot constant, but the whole thing becomes much more complex. So, this much more easier to just take thus hat constant separately, and work with them.

Now, this leads to, what is called linear programming problem in the standard form.

(No audio from 24:03 to 24:26)

These to minimize c transpose x. This is a standard thing; where c is in R n, and A is m cross n matrix. So, let us give an example of such a problem, where a naturally linear programming problem arises.

(Refer Slide Time: 25:08)



So, this would as an example, we will show the transportation problem. This is query interesting, because it is a real life thing. So, you must be quite happy that; too much of math is now gone, but and you are doing with real stuff, but you will soon seen math coming back in as fast as possible, because; you cannot really do the stuff without math you have. To have differed understanding of things in science, you need to have a good understanding of mathematics.

Now here, I have series of m supply points, supply of certain goods. So, these are ware houses storing books. So, these are flip cart ware houses. Everybody flip cart will knows now a days. And these are the demand points. This is store houses, and this is the point; where there is demand for flip card books, books on flip cart or any other thing.

So, the n such points. So, there is the supply point, these are the ware houses of flip cart. And this is the demand point; this is the c t point, say this is supplying to… So, this is the thing and this is the flip cart self-delivery points; is in the n points and m points ware houses. And now, for from each one of them; I can send to any one of this designation.

So, if I ask if I am say, this is Kanpur. Then, when I order for flip cart thing; the demand when I make a demand, any it will check wherever the supply is there, and it will see from where it can in the least cost send me the material.

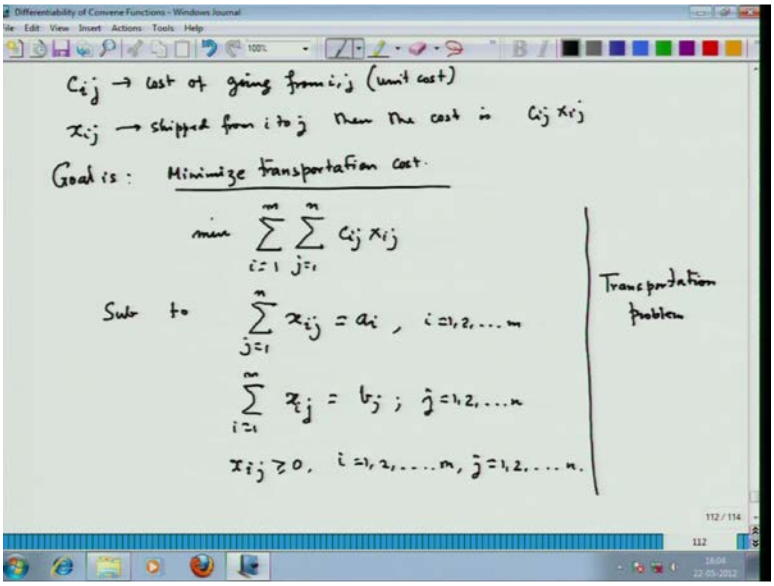
So, from any supply point, I can go to any point for that, there is the cost involves. So, you go from C 1 to 1 as C 11. So, 1 to 2 as C 1 2;1 to n as C 1 n. Similarly, for that second point C to 1, these are the cost involved transforming from this flip cart ware house to us supply point. C 1 n; so, forth you can do for the other also. The m also you can have this C m 1, C m 2 and C m n. So, these is some transportation network these are transportation network. (No audio from 28:27 to 28: 40)

Now, if this is the case; then I do something. Let us do proceed; the every ware house as some amount of say books or any other materials. These where houses a 1, this as a 2, this as a 3. And there is certain amount of demand from each of these faces. So, the demand is b 1, b 2, b m. Sorry, this as a m. A very, very important assumption that will take in modeling, because when we model in we need to assumption.

Any mathematical model cannot just capture real life; just like that. Because realities quite complex. So, you do some have some simplified assumption, and by getting some results from that simplifying assumption. You can have probably a quite a good understanding, what is actually going on.

So, I would know go ahead, and take this demand-supply condition, which is says that the demand must equal the supply. (No audio from 30:03 to 30:13) Say, that the total supply that you have in all ware houses must the match demand, and that is then we can actually sort this suiting, mach the demand that you have. So, this is t; total this is equal, this call the demand supply condition. This is inherent. So, if we if you do not have this balancing equation, there some other you can still get back to this form a work by putting some additional variables. But this is the standard thing that we assume.

(Refer Slide Time: 31:04)



Now, what is our intension? Our intension is a very simple. We want the c i j cost from going from i to j. So, how much I should material I should carry over the whole network. So, that my total shipping cost is minimum. So, in flip cart is doing a self-delivery. It is not going through some other customers. Because now-a-days flip cart is self-delivery is increase. Because, why? Because there observe that if they do, they can during the self-delivery process, deliver you the material at a much, much, much cheaper cost. And so, the total cost, over transporting so many goods over then network is minimum, so that is our goal.

Now, the c i j could be greater than equal to 0, could be even negative. So, cost of shipping a element the quantity element x i j. So, if I ship amount of martial x i j is shipped from i to j. Then the cost is… So this is cost cost of going from i to j, it is a unit cost, in the cost of say; courier cost of one material. Then which we says standard. So, then the cost is… Our goal is to minimize my transportation cost. So, the goal is…

(No audio from 32:39 to 32:59)

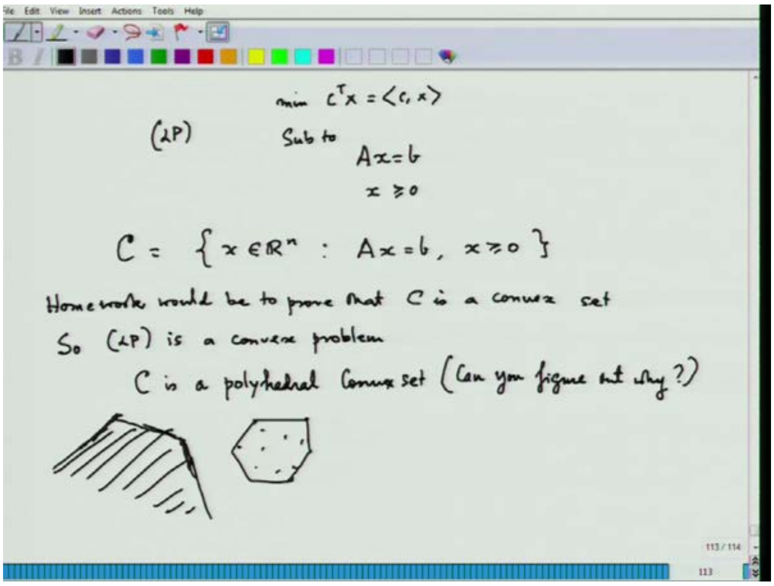
So, problem now is to minimize… subject to… Now, if I fix my i th supply point, and very my j th supply point. The total good z; I can send from i th supply point is a I, which is equal to the amount of total amount of supply that I have, I cannot send anything more.

And total goods… So, if I keep my j th reception point fixed, then whatever the amount of material that I need from total material, that is need from the various supply points as a vary it our supply points. Must equal my demand that is exactly what is required.

And all of this quantities x i j has to be greater than equal to 0, which is quite naturally, because are some books or something like that. What you get? It is a linear programming problem; this is the call the transportation problem. So, you see how very important mathematical real life problem; is been mathematically model as a linear programming problem is not beautiful. It shows the power that is inherent in this whole subject. Now, we have already learned a bit about duality, when in the last lecture very try to find the dual away to linear programming problem have a last two last lecture, in the standard form.

So, we are not get into dual aspect at this moment, but we will not show that you can actually it take a very simple problem; in the natural form. And then, if you have just two variables and two constants or three constants, then using geometry you can very well figure out, how to find the minimum. For that we will do come in to three steps, but first thing we have to determine that the problem of linear programming problem is actually a convex problem.

(Refer Slide Time: 36:02)

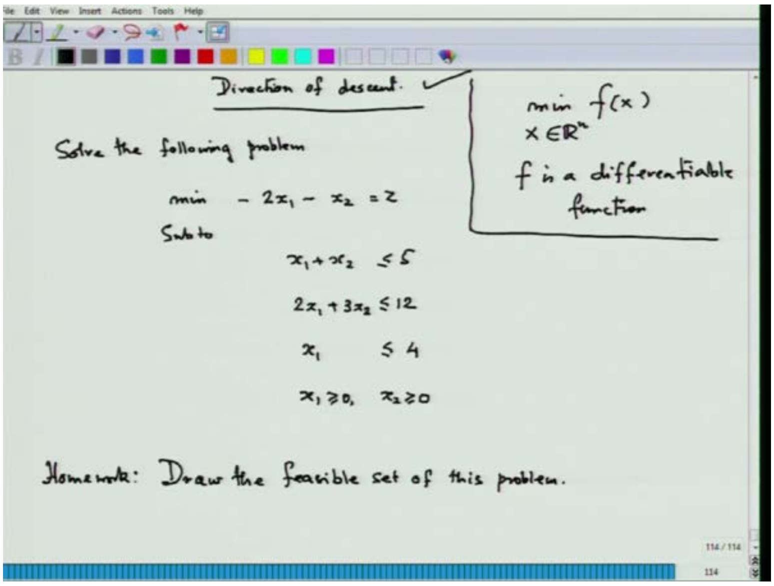


So, I have minimized, so I go back to my standard form. LP; in the standard form which would be always marked as LP not LP c. I am not writing what is A, what is b, and those entire things this is not required.

Now, the set C… the feasible set consist of all x in R n, which satisfy, your homework would be to prove (No audio from 36:39 to 36:49) that c is a convex set. So, what do you conclude is that LP is a convex problem, but it is very important to note that this class of set C is a very special kind of convex set, which we already mention once call polyhedral convex sets. So, the set C is a polyhedral convex set, it means, it is intersection of finite number of hypo planes and half spaces. It means look likes something like this, something this part for (( )) this.

So, something like this, say; set like this, or set like this are is some examples of polyhedral convex sets. So, this is a polyhedral convex set. Can you figure this, it can figure out. So, definition of polyhedral convex set is, I call that is in the intersection of finite number of half spaces. Now in order for you to understand where how does, how do we figure out of the minimum of linear programming problem? You must understand the notion of the direction of decent. That is mean which direction I will move. So, that my function value is decreases.

(Refer Slide Time: 38:53)



So, this is something would be important. So, we do not a much time to discuss this direction of decent today. So, what we will do is will postponed to for the next class, but what will do is that over idea would be to solve out of this particular problem, in a geometrical fashion.

(No audio from 39:19 to 39:39)

Many of this would say, we have doted this in are under graduate days, those who are already under graduate, this watching this program, graduates students, but lot of things which you have not really understood when you did the undergraduate classes.

So, we would write to makes certain things clear, rather than mechanically show you what happens. We would like to tell you what the hell is actually going on, beyond the set of rules what is the math. So, deduction of decent is very important thing, that is in which deduction; if I c x this problem is C transpose x, in which direction I should move in R n. So, that my functional value will continuity decrease.

Because I am want to minimize a function. So, I will end here with giving homework; draw the feasible set of this problem. (No audio from 40:35 to 40:48) So, will come back and start tomorrow, considering a problem of this form minimizes f x, x to minimize over the whole R n, and f is a differentiable function. You know by the by this time what is a minimum, differentiable function we have already discussed. So far this in which suppose I have a point where the function is not minimized, there is the gradient of f at x at that point not 0. Then I need to move towards another point, where the function value actually decreases.

So, I am actually going towards a minimum, so minimum or local minimum. So, that is. So, how do which I how do I know how direction to choose. The directional, and I get this degrees of function value is a deduction of decent, and like is a coming down in mountain basically. So, we will start over discussion with the deduction of decent tomorrow. Thank you very much.