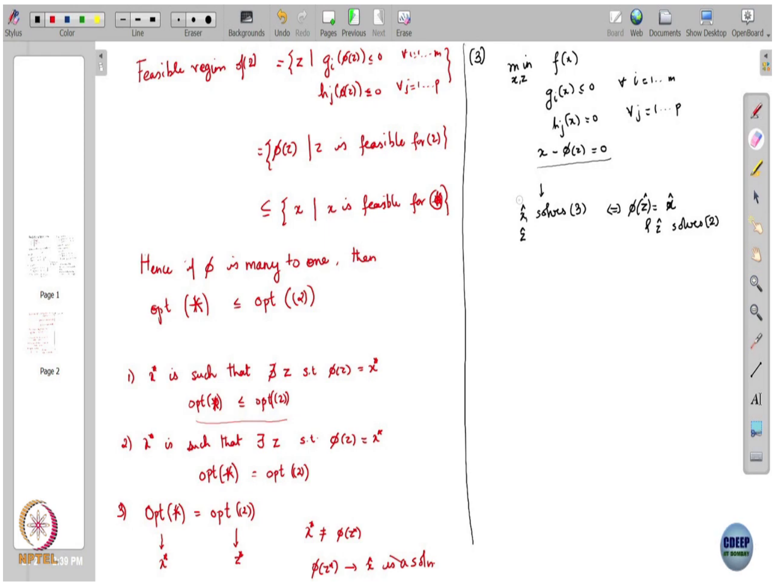
Optimization from Fundamentals
Prof. Ankur Kulkarni
Department of Systems and Control Engineering
Indian Institute of Technology, Bombay

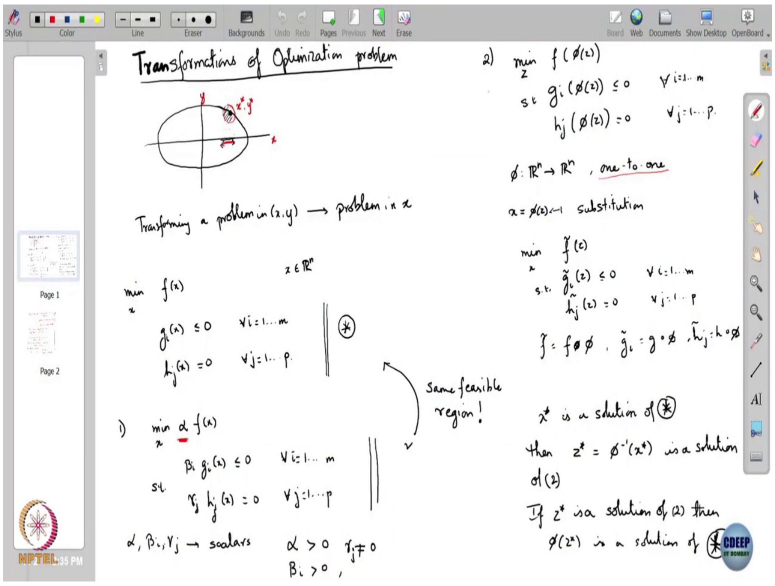
Lecture – 7B
Transformation of optimization problems - II

(Refer Slide Time: 00:26)



Now, if you did not have. So, let me let us bring go from here into a problem 3. Now, if I have phi which is not 1 to 1, suppose it is many to 1. Phi is many to 1 then in that case, how can I still maintain an equivalence between 2 and star?

(Refer Slide Time: 00:48)



So, that is the question here is without when it is not 1 to 1. How can I maintain the correspondence between 2 and star?

Student: Phi (Refer Time: 01:00).

So, if there are multiple z's, that correspond to the same that give you the same x or for ok. If you are on the other or the basically the issue is that not a for every x there need not be a z, right. For every x there need not be a z in that case, what can one, how can we still ensure a correspondence between 2 and star? So, the trick to do that is the following, see the what why did we fail here?

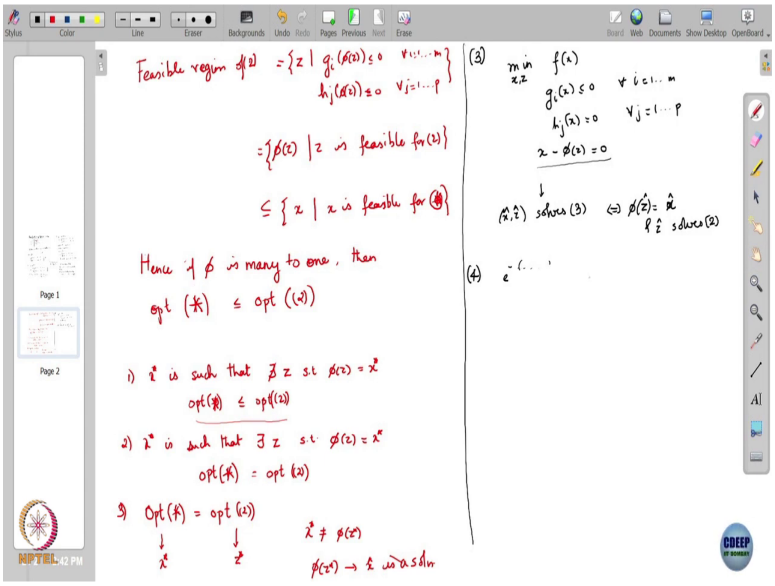
When we substitute, when there are multiple there are x's for which there is no z in that case what can what kind of problem happens? We substituted phi of z as x and we got to problem star, right. We substitute we got we substituted phi of z as x and got problem star and in that case what the, what has happened as a result? We are searching over too many x's, right far more many than we were supposed to in problem 2 right.

So, what we can do is bring that back in some through the constraints ok. So, the. So, suppose I look at this problem. So, this has all the constraints that star has and in addition to that, it is requiring that x must be equal to phi of z. So, let me write it like this x minus phi of z is equal to 0. So, x should be such that there is a phi there is a z for which x is equal to phi of z ok.

And now, I am minimizing over both variables x and z. So, as x varies z will also vary, but the, but my additional constraint here, x equal to phi of z this ensures that I am searching only over those x's for which there is a corresponding z. This automatically constraints my problems star to search only over those x's for which there is a corresponding z. And this therefore, becomes equivalent to problem 2. Is this clear?

So, if I get from here, if I get a get if I have a solution x hat here suppose x hat solves 3, if and only if z which is equal to phi of sorry phi inverse of not phi inverse sorry, if and only if so, actually let us take x hat z hat this thing solves 3 ok.

(Refer Slide Time: 04:56)



So, if x hat z hat this they that solves 3, let me put it try to clear this x hat comma z hat solves 3. Then you must have that x hat is equal to phi of z hat that is because of the constraint and z hat solves 2, right and this is equivalent. Is this clear? So, you can bring it do the substitution or change of variables, but then if you are losing the , if you are losing information if you are losing information by about the z by transforming it to star, you can bring it back through the backdoor by putting it as a constraint ok.

Now, it may or may not simplify the problem that depends on the problem at hand. Sometimes for example, one of the one of the issues that occurs is, you can have algorithms might not be able to process these kind of functions. You know function passed to another function it may make errors in calculating gradients and so on. Those kind of issues can occur.

So, it is much cleaner to then write out give them one function one clean function in each constraint, and then put this as an additional constraint. The cost you pay for this is now you have additional variables also. So, your problem size has increased in terms of the number of variables you are optimizing over. So, this this can sometimes be a cleaner way of formulating the problem than formulating in in this sort of in this kind of complicated manner ok.

Student: (Refer Time: 06:28).

Yes.

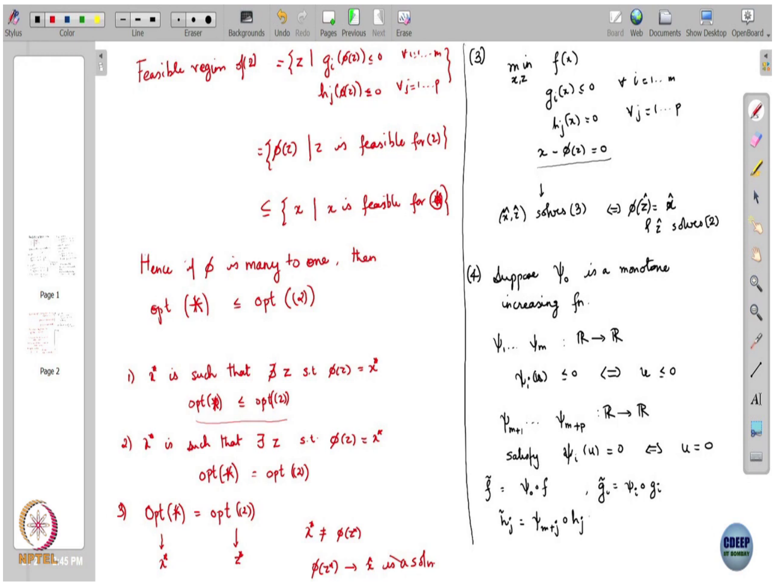
Student: (Refer Time: 06:29).

Yeah. So, that could be another. So, a many a case where you cannot for which case where you have for every x there need not be a z that sort of case would be when you are changing dimension also. So, if you are mapping down for example, if you are going from Rn to something smaller some Rk, which where k is smaller than n.

In that case, there will be a vectors in the range space which are not for which there is no corresponding z, right. Yeah ok. So, let another commonly used trick which is probably you already know is to do transformations that are of a monotone kind, but I will explain this in a much more in a better way now in a more general way.

So, you. So, we you may have seen for example, when you want instead of trying to maximize instead of trying to maximize say the exponential of a certain function. One of the things, that people do is instead they simply take logarithms on this on the same function and then you maximize the log of it. And the reason for that is that log is a monotone function right. So, that is the thing that I will now tell you in a much more general way.

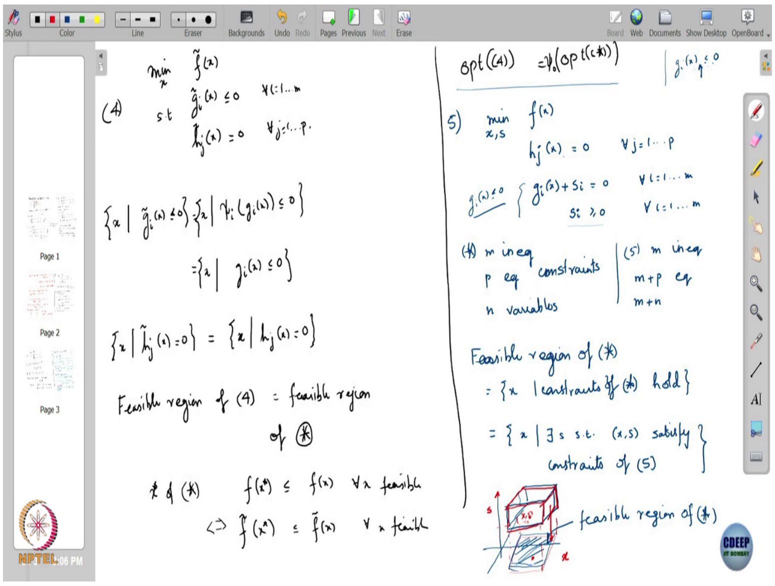
(Refer Slide Time: 08:13)



So, look at this. So, suppose here psi 0 is a monotone increasing function. Now, suppose also psi 1 till psi m these are all functions from R to R and they are such that they satisfy psi i of x less than equal to 0, if and only if not x let me denote this by u psi i of u less than equal to 0, if and only if u is less than equal to 0, ok. And suppose psi m plus 1 till psi m plus p, these are also functions from R to R.

They satisfy psi i of u equal to 0, if and only if u equals 0. Now, you define say f tilde as psi 0 composed with f and g tilde i as psi i composed with g i and h tilde j as psi m plus j composed with h j. So, I am composing this time from the left alright so, in the end, now you look at this optimization problem, which is my problem 4.

(Refer Slide Time: 10:33)



I am now minimizing f tilde of x this does not this transformation does not change my space still x subject to g tilde i of x less than equal to 0 for all i 1 to m, h tilde j of x equal to 0 for all j equals 1 to p ok. So, this is my this is my optimization problem. Now, this is actually equivalent to 2 star. So, how do you see? Again, you want to say that these are equivalent, equivalent in what sense? Equivalent means solving one gets you to the solution of the other, right.

Now, in this case actually it is much it is even better. So, first let us compare the feasible regions of the two, can you say something about the feasible regions of the two? Feasible region of this problem 4, how does that compare with the feasible region of star? So, suppose I have an x that is feasible for 4, let us start with 4 I have a x that is feasible for 4, which means it satisfies all these constraints, satisfies all the inequality constraints all the equality constraints.

Now, gi tilde of x is less than equal to 0. Now, let us take this as just as an example, let us take a gi tilde of x less than equal to 0 ah, gi tilde is what was gi tilde? gi tilde was psi composed with gi, right. So, it says. So, this is saying psi i of gi of x less than equal to 0 and if you see my condition on this psi i for i from 1 to m, the condition is simply that condition is that psi i of u is less than equal to 0 if and only if u is less than equal to 0, right.

So, this. So, if I look at all the x's, if I look at all the x's such that this is less than equal to 0, which is equal to the x's such that this is less than equal to 0. That is actually the same as the x's for which g i of x is less than equal to 0. Because psi i of g i of x is less than equal to 0, if and only if gi of x is less than equal that follows from the underlined thing property here. And similarly you can so, you can do this for all the inequality constraints together.

Likewise for the equality constraints, you have that psi i of u of is equal to 0 if and only if u is equal to 0, ok. So, which means that u is the only 0 of this function psi. I can have whatever shape you want, but it has it has only one 0 which is at 0 right. So, in that case again the what will happen is that x such that h tilde j of x equals 0, if this set the x's for which h tilde j of x is equal to 0 is the set for which h j of x is equal to 0, right.

So, their feasible regions are so, the feasible region of the two problems are the same. Feasible region of 4 is equal to the feasible region of feasible region of star. What about the objective? The objective has been transformed through a monotone increasing function, right. So, if you have a solution x star of star so, you have that f of x star is less than equal to f of x for all x feasible , then it will also be that f tilde of x star is less than equal to f tilde of x for all x feasible.

This is because I can I on an inequality I can always apply a monotone increasing function and that does not change the direction of the inequality, right. So, this the right. So, this is if an only if this is an increasing function so, it is also invertible so, I can go back and forth. So, what this would mean is that you all so, the optimal value of 1, optimal value of 4 is going to be psi 0 applied to the optimal psi 0 of the optimal value of star, right.

So, the optimal value of 4 is optimal value of star psi 0 outside. Let us look at another example of transforming problems, this time which concerns introducing additional variables and changing the nature of constraints, ok. So, look at this problem. So, minimizing now f x subject to I have my equality constraints h j of x equal to 0 the original equality constraints, but I will also do the following thing, I will I write g i of x which was an inequality.

Let me write it like this, gi of x plus si equal to 0, ok. This is for all I from1 to m, well, but this si is a new variable that I introduced. Now, remember gi of x, remember was gi of x was in the original problem was less than equal to 0. So, there is a gap between so, gi of x is and 0 there is a gap here. So, I can fit in a positive number here between gi of x and 0 and make the two make such that this becomes equal, right.

So, that is that number s ok. So, si is now something that is greater than equal to 0 that has been fitted into that. Now, as my x changes, it is for to in order to maintain this; in order to maintain this equation here, ok. In order to maintain this equation, my s will change value ok. So, it is I need to bring that in also as a variable that cannot it is not a constant anymore, ok. It has to be if I make it a constant, then that will change the meaning of my constraint, ok.

It has to also float as my x changes. So, my variables now are x as well as S. So, this. So, what has happened as a result of this transformation? What is happened is my I have earlier I had inequalities on g, ok. Now, I have an equality constraint in g, but I have introduced new variables, which are these S's and I have now an inequality on the S's right.

So, my earlier problem, if you look at star it had m inequality constraints and p equality constraints and of course, n variables. On the other hand 5 has now m inequality constraints, how many equality constraints? m plus p equality constraints and how many variables? m plus n variables. Now, what is the advantage of doing something like this?

There is a there is one there are there are sort of it has its one key advantage is that it standardizes things quite a bit for us. So, all equalities will now be of a very simple form, which is all sorry all inequality constraints can be without loss of generality considered to be of this very simple form; that means, there are some variable greater than equal to 0, right.

So, you; so, inequality constraints will simply take the form of defining some quadrant or some orthent in your space and equality constraints will define surfaces in the in the space, right; whereas, earlier you had surfaces defined by equality constraints and you had regions defined by inequality constraints and that interaction can be a little complicated this tends to make things a lot simpler, ok. So,

Student: (Refer Time: 21:10).

Yes.

Student: Why (Refer Time: 21:12).

Yeah so, that is a so, si greater than equal to 0 is the inequality constraint. So, the question is what are the inequality constraints in problem 5? This is these are the inequality constraints

Student: (Refer Time: 21:27).

It is a variable its so, it is of the function si itself is being asked to be greater than or equal to 0, right.

Student: Sir.

Yes.

Student: (Refer Time: 21:39).

It is should be treated as an independent variable. See eventually all these variables will interact, you have to pick an optimal choice for all of them together. Jointly need to be they need to be chosen ok, but when we write variables in an optimization problem, if one is known to be a function of the other, right. Then you should substitute and get rid of that, right. So, that is the so, eventually left you are left with these bunch of independent decisions, that are bound by constraints that is how imposes the problem.

So, si is si will is another variable, but then it is the values it can take is dictated by this you know these two constraints that. So, once you tell me the x, it fixes the range for si you know for every i. Is this clear? And of course, the in the optimal value of optimal value they have to all be chosen together they cannot be chosen you know independently of each other.

Because the you know their regions are such that you know one you need to know one for to know what the others. So, as I was saying the so, can we can you observe why the two are equivalent first, we are not I have not discussed that. So, why is 5 equivalent to star? So, see remember so, the way to one the way to observe this is look at if you look at the feasible region of star.

You look at the feasible region of star this is the x is such that all the constraints of star hold. Now, how do I compare this with the feasible region of 5 which is not in the same space anymore? Because now it has more variables, right. 5 is both in the x and the s space. So, you have it is in a larger dimensional space. So, how do I compare the feasible region of star and the feasible region of 5?

No, no, no, see si is not present in my original problem in present in star. So, how do I compare the feasible region of star with that of i? Well, that is the one way to do this is to observe the following, see the feasible region of star is the x such that the constraints of star hold, ok. Now, the feasible region of 5 can be brought in the following way, you look at the x such that together with s ok.

So, what I have written here is this is. So, look at the set of x's for which you can find some xs, ok. Such that x and s together satisfy the constraints of 5 so, x and s if they together satisfy the constraints of 5, right. And you look at such x's only the x component of that x comma s there, ok. Now, for those x's would not the constraint of constraints of star hold automatically, right.

Because if x comma s satisfy 5 satisfies the constraints of 5 what of course, g h j of x h j of x would be equal to 0, all of these would be satisfied h j of x would be equal to 0. And here, you would have that gi of x plus si equals 0 and si would be greater than equal to 0 ok, but gi of x plus si equal to 0 and si greater than equal to 0 would mean that, gi of x is less than equal to 0, right. So, these two together implicitly automatically imply that g i of x is less than equal to 0 right.

So, consequently what is happened is that the x's for which the, for which star the constraints of star are satisfied. Is the same as the x's for which you can find some s, such that the constraints of 5 are satisfied with x and s together. So, that basically ensures that these feasible regions, this is the sort of way you can establish a link between the two feasible region. Geometrically what is happening is, if you want to, if you want a picture here.

Suppose here, you know here is the space here is a region of on which you have defined the feasible region of star. This is the feasible region of star. What you have? What we have done here is sort of by going by introducing an additional variable, the feasible region of 5 has become something like this.

So, feasible region of star here is this box is this a flat square here whereas, the feasible region of star is this red 3-dimensional box and its such that if I take the shadow of this box on this on to one of the variables that we look at its projection onto one that gives me back the feasible region of star. So, this is region of x this is this is the space of x's that the additional x's that we have that is the space of s.

So, any point x here has a corresponding s such that x comma s together lie in this floating box on top, that is your x comma s. And likewise if I take an x comma s here, I can project that back down to get to a point in the feasible region of star, right.

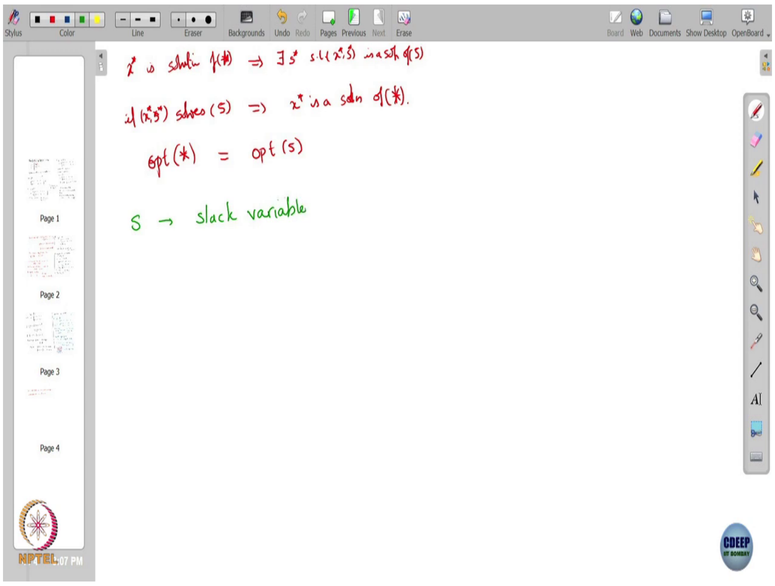
So, the feasible region of star is in some sense a shadow of the feasible region of 5, right shadow or projection or; however, you want to think of it, ok. So, now, but this this simply relates the two feasible regions, why are they equivalent optimization problems?

Student: (Refer Time: 29:20).

So, the reason they are equivalent is because see in star you are optimizing only over this plane here, this, this, this, this, this sort of square that region over the feasible region on down here , but and in 5; in 5 the objective function has not changed going from star to 5.

Object you have introduced this new variable s, but that variable does not appear in the objective. What matters for the objective is the value of x alone. It is still f of x which is the objective, which was the objective earlier also in star also. So, the f of x is that you can get the values of f of x that you can get in star, is the same as the values of f of x that you can get in 5, right. So, as a result the this the objective values will actually be the same ok.

(Refer Slide Time: 30:24)



So, the bottom line is that you are you if you have a solution x star is a solution of star this implies that there exists a s star such that x star s star is a solution of 5. And likewise, if x star s star solves 5, then x star is a solution of the solution of star, ok.

So, this is how you can go back and forth and in fact, moreover the optimal value of star would be equal to the optimal value of 5. So, the this new variable s that we just introduced has a name it is what is called a slack variable, ok. So, the s is what is called a slack variable that is the.