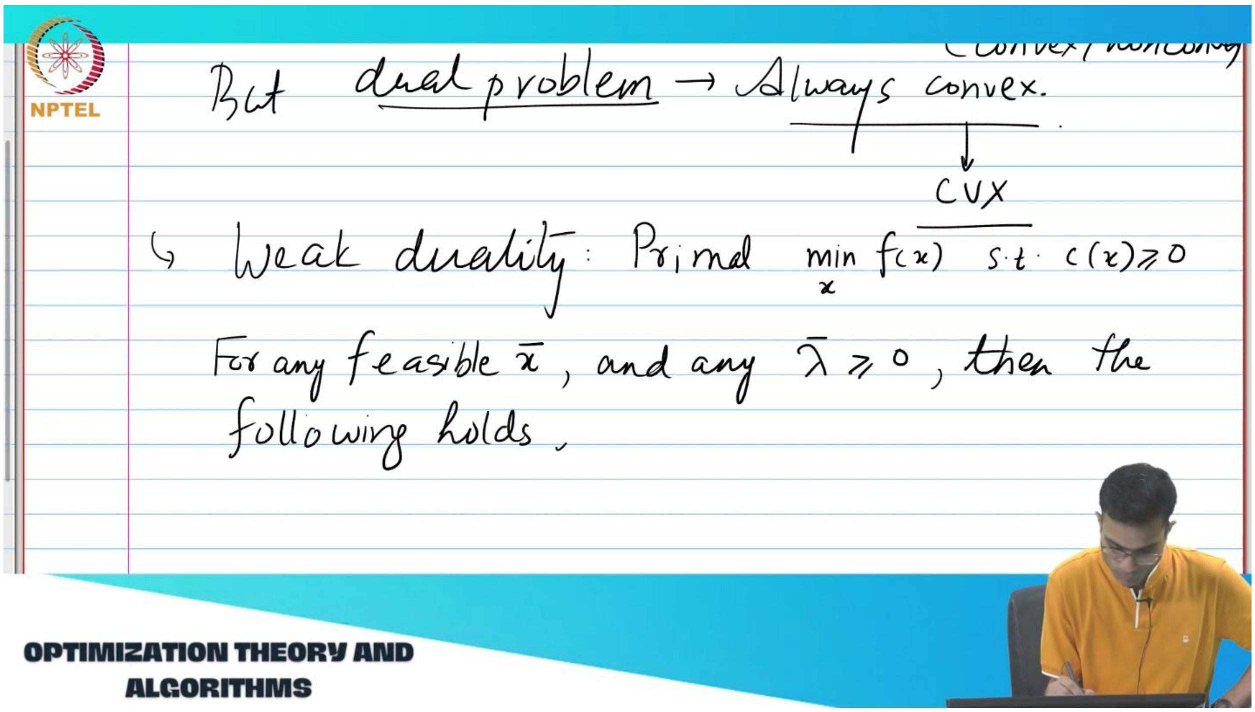
Course Name: Optimization Theory and Algorithms
Professor Name: Dr. Uday K. Khankhoje
Department Name: Electrical Engineering
Institute Name: Indian Institute of Technology Madras
Week - 12
Lecture - 81

Proof of concavity of the dual problem - Part 2

This is what we study under the title of weak or strong duality. So, ok. So, let us just note down what is the primal problem again. So, it is

$$\min f(x) \quad \text{such that} \quad c(x) \geq 0.$$

This is our constrained optimization problem, okay. Now let us see what we say about the dual problem, okay.



So, let us take for any feasible $\bar{x}$, okay. Feasible $\bar{x}$ just means it is any regular guy in the domain of the constraint set. It need not be the best x. So, any feasible x and any $\bar{\lambda} \geq 0$, then the following holds.

So, I am going to note one statement and which we will prove. So, $q(\bar{\lambda})$. So what is this saying? That take any feasible $\bar{x}$ and take any $\lambda \geq 0$, then the value of the dual problem is always less than the value of the primal problem. So this in itself does not look very interesting, but because we have an inequality, we can make this inequality even stronger. For example, I have the left-hand side less than or equal to the right-hand side. If I want to make it really tight, what is the tightest I can do on the right-hand side, a minimization operation or a maximization operation? Minimization operation.

And on the left-hand side, maximization operation. And remember what was the definition of q? q had a minimization operation going on, right? So, let us define f^* as the infimum of $f(x)$, such that $x \in$ constraint set. Instead of infimum, you can also just think of it as a minima. And we will write q^* as the supremum of $q(\lambda)$, $\lambda \geq 0$. Now if I substitute f^* and q^* into this inequality, is it still correct? It is correct because on the right-hand side, I have made it as small as possible, the left-hand side I have made as large as possible.

So this is really something that is very interesting. So when I minimized f, I made f as small as possible while x still belongs to the constraint set. Isn't that the solution to the primal problem, right? That is what I was looking for. Give me the smallest f such that x belongs to the constraint set. So therefore this is the optimum of the primal, $q(\lambda)$, once I got the $q(\lambda)$ in our two-step process, once I defined $q(\lambda)$, to get the dual solution, what did I do to q? I did a maximization problem.

So this q^* is what? It is the solution to the dual problem. So this is the optimum of the dual problem. So there is an inequality here that the optimum of the dual problem can at best reach the optimum of the primal problem. So this is a very very important result in duality. Oh, why is this statement true? Yeah, I will prove it.

We will prove it now. I thought I will just give you the implication and then we will get to the proof. So this is the implication. The proof is actually also very simple. So let us look at the proof of this.

So, we just go to the definition of our, we go to the definitions of Lagrangian rule, Lagrangian and so on. So, the definition of this was minimization over x gives me my Lagrangian rule function. And this is essentially, if I just substitute the definition of this guy, this is the definition of the Lagrangian function, ok. Everyone agrees with this. Now I am going to put a box over here for you to fill in equal to less than greater whatever.

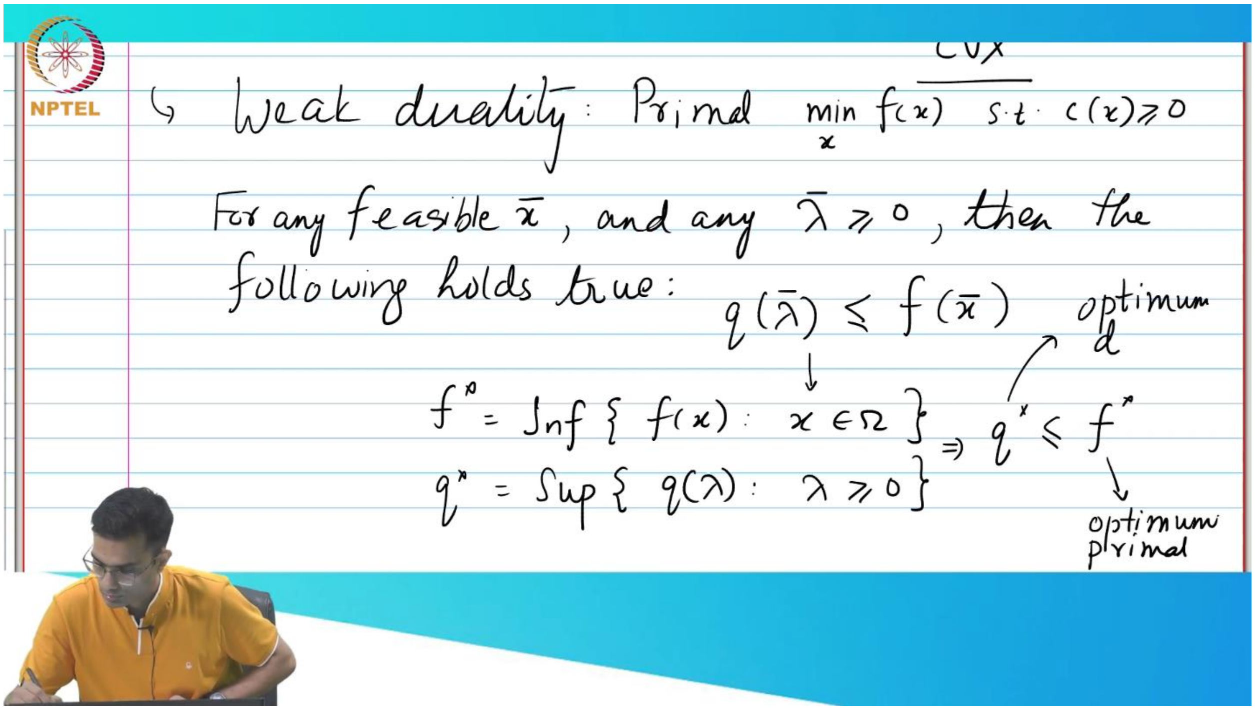
$\leq$, right because the top thing is the minimum. So, if I am not at the minimum, the value of this guy is going to be less than or equal to f of this, right because and x is simply feasible, ok. So, this gave me $q(\bar{\lambda}) \leq \cdots$, right. Now, did I put any constraint on λ? I said $\lambda \geq 0$, right. So, $\lambda \geq 0$.

What about $c(\bar{x})$? Is there an inequality on that? It is also ≥ 0. So, therefore what is this term? Always ≥ 0, right. So, therefore this whole term and I can get rid of a minus term, the inequality would still hold, right. So, surprisingly the proofs in the world of duality are very simple, we just use definitions of concavity, convexity and so on. So it is clear what we did.

We took the definition of the Lagrangian dual function. We substituted any feasible x with inequality hold. Then I noted that the constraint is always ≥ 0. I am saying that this holds true for $\lambda \geq 0$ and then I am done.

So this holds. So this gives us a very important couple of points. If q^* is strictly less than f^*, q^* and f^* are the optimum values, right. Then $f^* - g^*$ is obviously ≥ 0. This is called the duality gap, ok. When there is a duality gap, this is also called weak duality.

In the second case, if $q^* = f^*$, no gap, this is strong duality, ok. And the final statement I am going to make over here without proof, I am going to state it, but this is one of the major things that distinguishes convex and non-convex optimization. For convex problems, the duality gap is always, for convex problems, let us just say always strong duality. So in, so always strong duality, there is one small, you know this is a star, which is the terms and conditions. So there is one terms and condition over here.

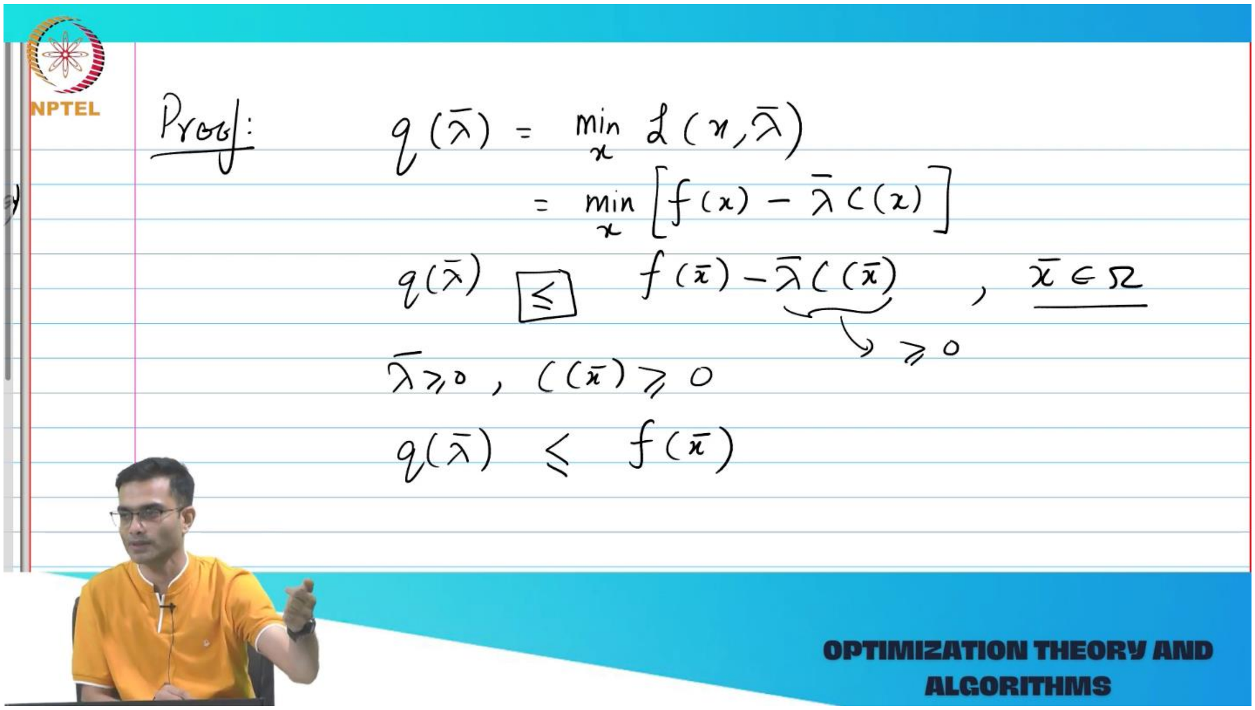


It is the same thing that we have encountered earlier between the discrepancy between algebra and geometry if you remember. That discrepancy needed one small screwdriver to tighten it which was what? Constraint qualification. Same thing holds over here that strong duality holds under constraint qualification, right. So we will just mention this as a star. Under, and an example of constraint qualification we studied in this course is LICQ, ok.

There are other types also, but if this holds, the statement of strong duality also holds. So this is why in the convex world you can solve the primal problem, you can solve the dual problem, you will reach the same optimum value. So you really you have that choice which way to approach it. If you have a non-convex problem, solving the dual problem is still not too bad.

If you have nothing else going on, if the primal problem is very, very difficult to solve, at least the dual problem will give you some kind of a bound on the function value, right? You know that the function, supposing you solve the dual problem, you get q^*. You know that the solution to the primal problem can never be what? Below that value. So, that may be useful information in many problems. So this is essentially what I wanted to cover about duality and the difference between convex and concave problems. And as you have seen the example, I mean the proofs are really simple.

Yeah, question. Yeah, that is the duality gap, f^* and q^*, that is the duality gap. So, yeah, absolutely. So, if I take the optimum value of x^* and λ^* from the dual problem and substitute it back, I am not going to get the optimum f. That is the meaning of duality gap. So, the question is, is there any notion of closeness? Unfortunately not, unless you provide some further information in the problem, we cannot quantify how small is this gap, and so that is also an open area of research to quantify how small is this gap.



Any other questions? So to kind of drill in this idea of the primal problem and dual problem, again I am going to, it is more like a tutorial, we will work through a problem, solve it together so that we get a feel for the dual problem. When I am at the primal problem, we have spent most of the course working on. You need not spend more time on it. But the dual problem can often be a little bit, because it is a two-step process. First is a minimization, second is a maximization.

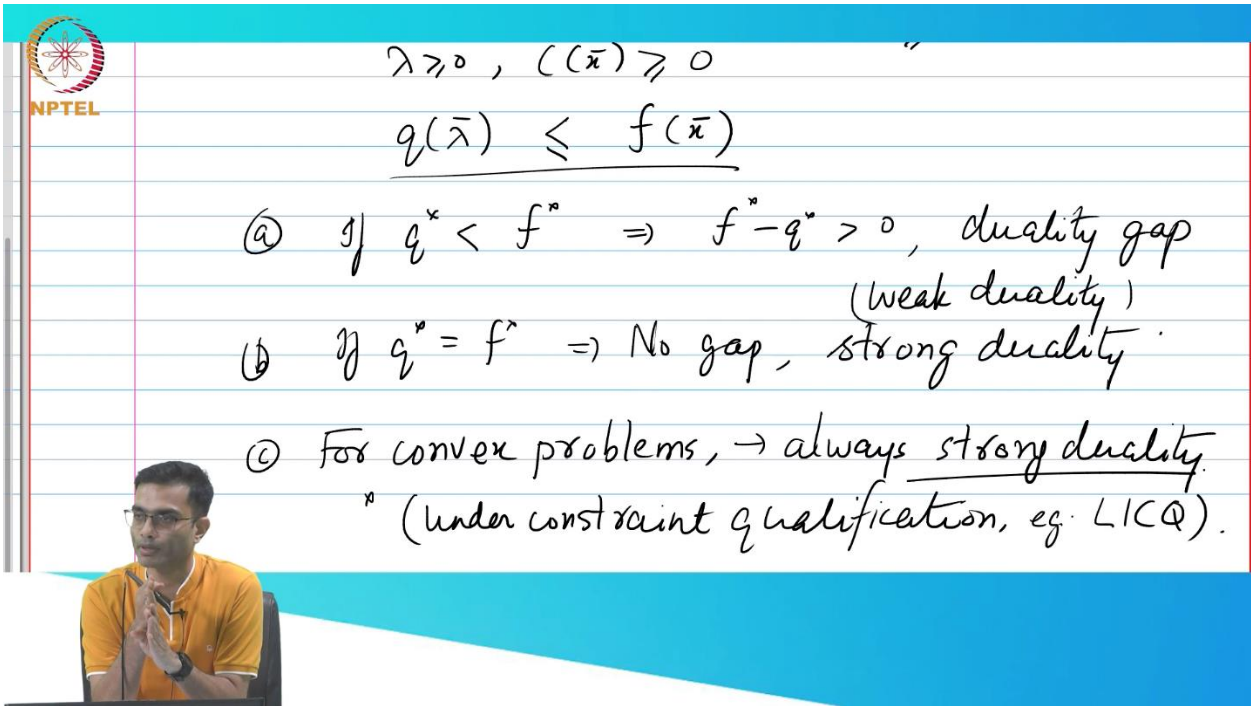
So many times it is like solving two optimizations. It is actually solving two optimization problems. One is a minimization, the second is a maximization. So you need to have a little bit more patience in working with it.

So let us take an example. Okay, so I'm gonna take an example whose primal problem is just graphically I can see it and tell, okay. So minimize, let us say. So you might encounter problems like this. Minimize $-x_1 + x_2$. In fact, in the tutorial, there was one such problem.

There's no problem, right? I mean, minus is there, so what? It may not always be with a plus sign. Subject to, let us say, this is my primal problem. So, constraint 1, what does it look like? Interior region of a circle. So, if I were to sketch this, radius $2\sqrt{2}$.

And so, one region is over here. What is the second constraint? Below, right, $x_2 \leq 6$. So, will this line cut the circle or be above or below? Above. What is $2\sqrt{2}$ roughly? 1.42, 2.8 something, right. So, this is 6 over here, ok. So, this guy is saying that you should be below here, right. So, what is the effective feasible set? It is just the interior of the circle, ok. So, the second constraint is kind of redundant, ok. And now, what are, where is, what is the, what are the contours of the cost function? Straight lines with slope $-45°$, we have seen this before, right, like this.

So in this case where is the solution? Topmost side because there is a minus sign associated with $x_1 + x_2$. So, the solution is sitting over here and this is because at this point $-x_1 + x_2$ is minimized, right. What is that minimum value? -4, right because this point has coordinates $(2,2)$. Therefore, $f^* = -4$ and $(x_1, x_2) = (2,2)$. This is just graphically, we did not solve any Lagrangian to it.



So, now let us solve the dual problem. Two-step process. So, first step, first step I need to first form the Lagrangian, right. So, the Lagrangian of x, λ is:

$$L(x, \lambda) = f(x) - \sum_i \lambda_i c_i$$

So, this becomes a little bit longer to write. So, $-x_1 + x_2$ is my $f(x)$, ok, and then I have my constraints.

So, c_1 and c_2 are there. So, I get:

$$L(x, \lambda) = -x_1 + x_2 - \lambda_1(8 - x_1^2 - x_2^2) - \lambda_2(6 - x_2)$$

This is my Lagrangian. So, if I extract a minus sign common from all of this, do I have a quadratic in x_1, x_2? Right? So, what are the coefficients of... I mean, maybe I do not need to extract a minus anyway. So, let us just write down the terms for x_1^2, x_2^2, and $x_1 x_2$.

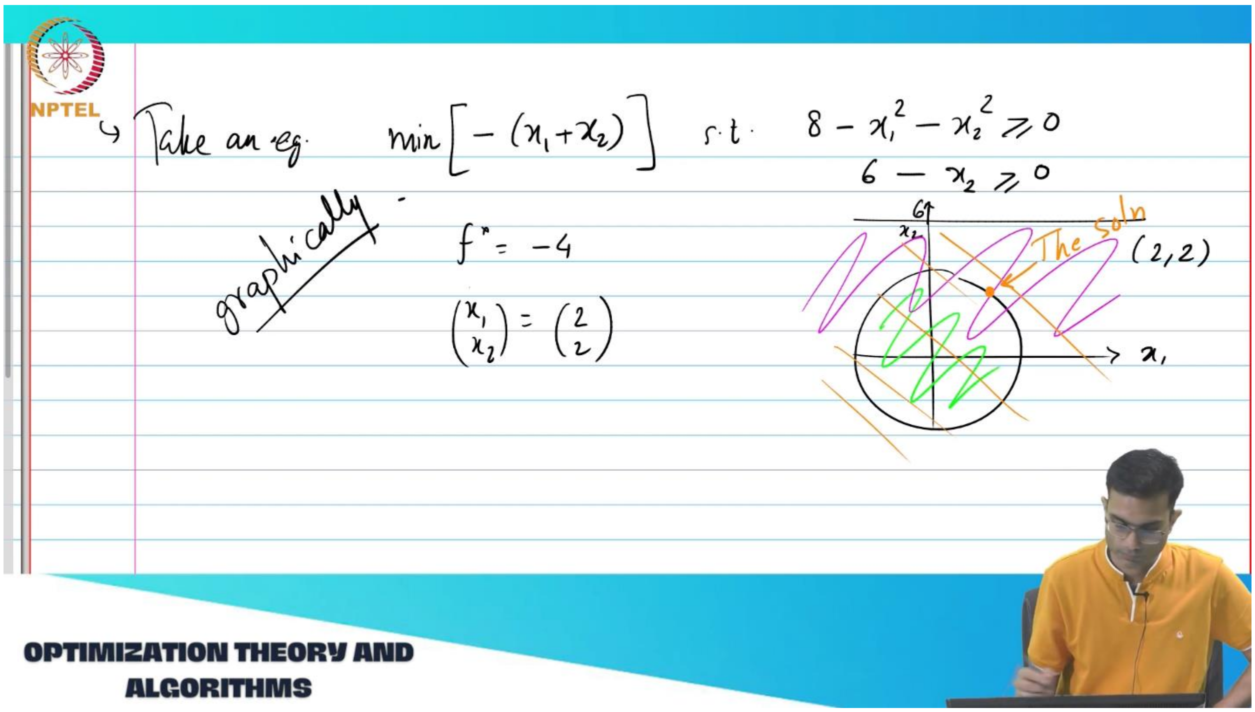
- The x_1^2 term has coefficient λ_1. - The x_2^2 term has coefficient λ_2. - The $x_1 x_2$ term has coefficient 0. - The x_1 term has coefficient -1. - The x_2 term has coefficient $1 - \lambda_2$. - The constant terms are $-8\lambda_1 - 6\lambda_2$.

So, is this a convex function? Yes. Because $\lambda_1, \lambda_2 \geq 0$.

Why are $\lambda_1, \lambda_2 \geq 0$? Because these are inequality constraints, right? So, these are all subtle points. If it were an equality constraint, I cannot say that, right? So, this is convex. Therefore, it makes sense to minimize this with respect to x. Otherwise, it does not make any sense, okay?

So, $q(\lambda)$ is going to be the minimum with respect to x of this Lagrangian. To get this, I would take the gradient with respect to x_1, x_2, right? So, what is $\nabla L(x)$?

- The derivative with respect to x_1 is $2\lambda_1 x_1 - 1$. - The derivative with respect to x_2 is $2\lambda_1 x_2 - 1 - \lambda_2$.



I set these equal to zero, right? So, when I set it equal to zero, what am I looking for in this? I want to find the values of x.

So, x_1^* is going to be:

$$x_1^* = \frac{1}{2\lambda_1}$$

x_2^* is going to be:

$$x_2^* = \frac{1 - \lambda_2}{2\lambda_1}$$

So I have done my part 1. I have basically got my x^* and now I can substitute and get $q(\lambda)$.

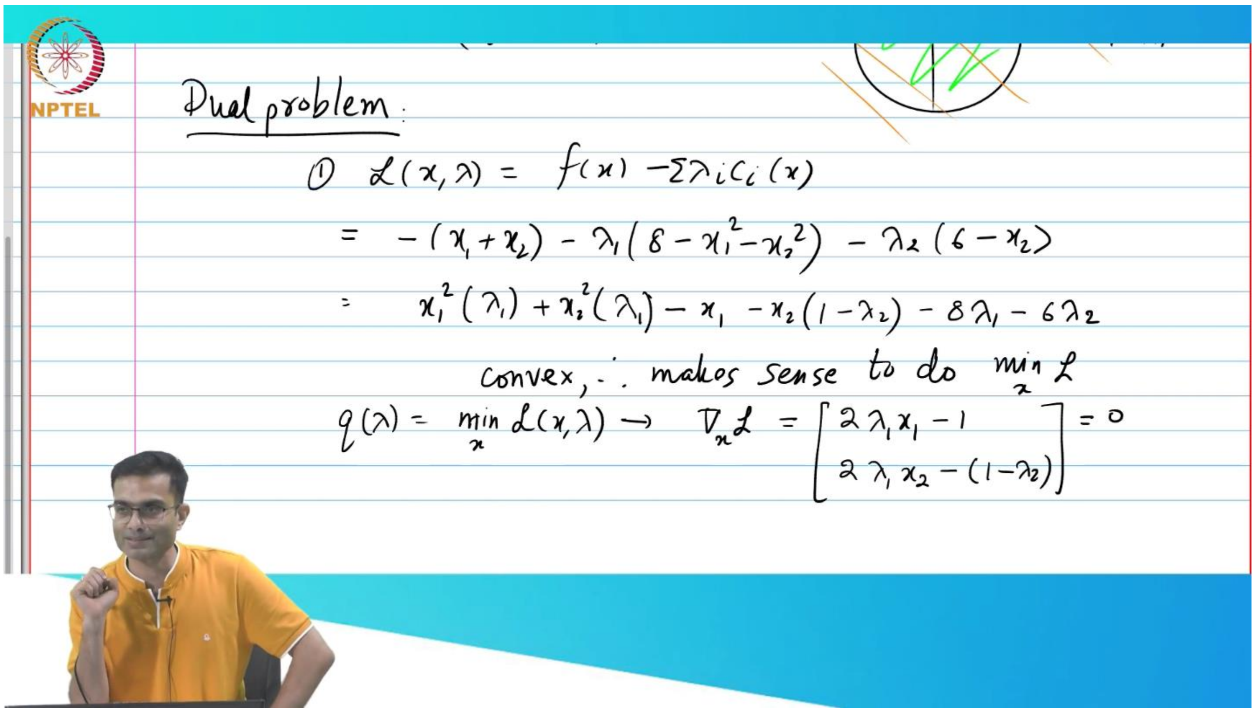
So, $q(\lambda)$ is going to be, so I need to substitute this x_1^* and x_2^* into that q expression in order to get it, right? So $q(\lambda)$, let me write down the expression over here:

$$q(\lambda) = \lambda_1 x_1^2 + x_2^2 - x_1 + \lambda_2(x_2 - 1) - 8\lambda_1 - 6\lambda_2$$

Just rewritten it, and now we need to substitute this into this expression, right? So, this is going to be, when I put $x_1^2 + x_2^2$:

$$x_1^2 + x_2^2 = \left(\frac{1}{2\lambda_1}\right)^2 + \left(\frac{1 - \lambda_2}{2\lambda_1}\right)^2$$

What about the denominator? What is the denominator? It is $4\lambda_1$, that's the first two terms, okay?



And then I have $-x_1$, which will become -2, then okay, then I have x_2. Now, x_2 multiplied by, so there is a, this will become $(1 - \lambda_2)^2$, right? That is this, and then I have $8\lambda_1$ and $6\lambda_2$, sorry. There is going to be a 2 somewhere, right? So, there is going to be 2 here, and there is going to be a 2 here, right? Let me just write it nicely. Oh, actually these guys can just remain as they are: $8\lambda_1 - 6\lambda_2$. Is that right? Yeah? The last two? It is not in the numerator.

The line ends over there. Okay, so this is what it is. Now, step 2. So, it's not a very, you know, convenient-looking form, but it is what it is. So, what is step 2? Once I get q, what do I do with it? Maximize it, right?

Now, step 2: Solve. All right, so:

$$\max_{\lambda \geq 0} q(\lambda)$$

Now, again to verify that this, we should do this step actually, to verify that maximization makes sense, you have to convince yourself that this function that I have is actually concave, then maximization will make sense. It is a function in two variables, so we can verify concavity by looking at its Hessian. That is one straightforward way of doing it, okay. You could also complete the squares and show that it is a concave function. So, all of that.

So, let us assume that we have checked concavity. So, let us do this. I have done a little bit of algebra to speed this process along. There are a whole bunch of minus signs that you can see in the definition of $q(\lambda)$. So I am going to pull those minus signs out. If I pull a minus sign outside, what happens to maximization? It becomes a minimization, right?

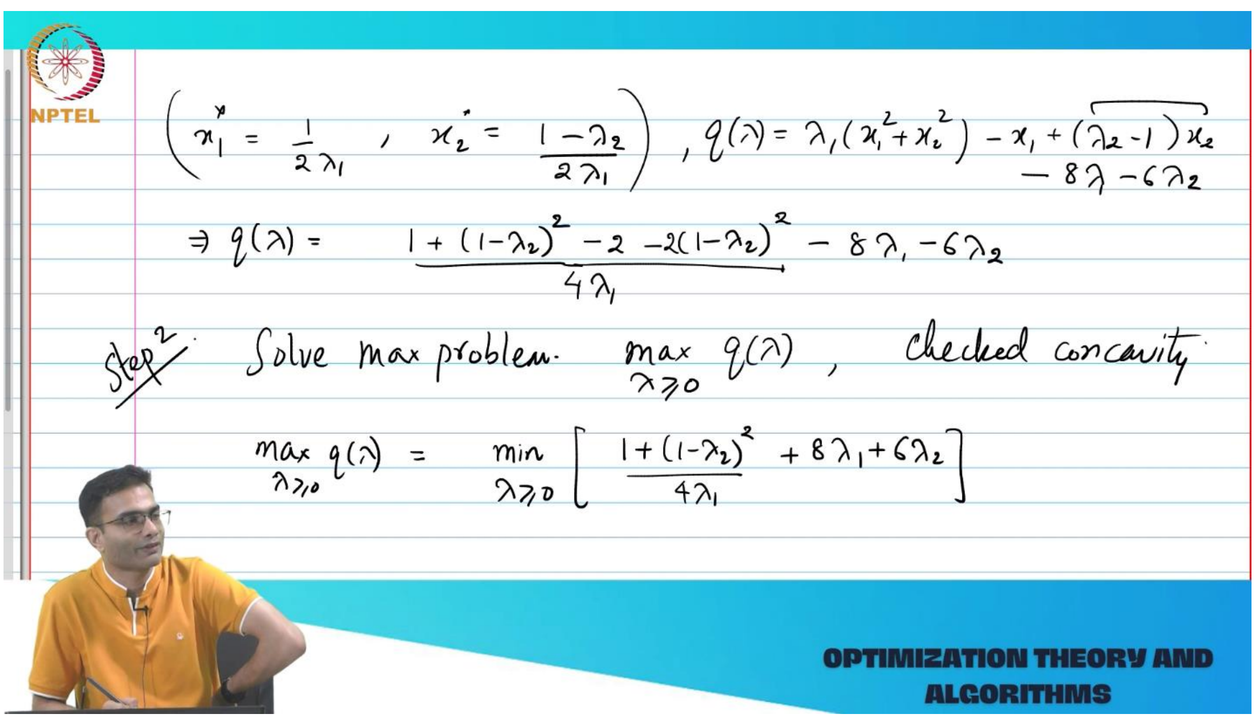
So, I am going to skip a few steps of algebra just in the interest of time and remove that minus sign, which flips the max to min, okay. So that is all I have done. So, I am going to write the final expression that I get.

Correct. In order to get $q(\lambda)$, I cannot verify that independently unless you tell me what λ is. Correct. So, in the way that we have formulated the dual problem, which we did without proof, but just with geometric intuition, we did not impose something extra. It turned out that the second recipe gave me the same answer.

Correct, correct. So, we will have to check once again whether, I mean we will check whether or not this, okay. We have proved, yes. So, why am I doing it again? Just for practice. We know that $q(\lambda)$ is going to be concave, but it's a simple enough example for you to verify.

You don't have to verify it each time. So when does it help to verify? It will help to verify if you have made some simple mistake in algebra. So, some simple steps, these are the things which help to catch your own error that you see, oh, hey, wait, this is not turning out to be concave. Okay. So, this is what this whole problem ends up being, okay.

The line ends over there. Okay, so this is what it is. Now, step 2. So, it's not a very, you know, convenient-looking form, but it is what it is. So, what is step 2? Once I get q, what do I do with it? Maximize it, right?



Now, step 2: Solve. All right, so:

$$\max_{\lambda \geq 0} q(\lambda)$$

Now, again to verify that this, we should do this step actually, to verify that maximization makes sense, you have to convince yourself that this function that I have is actually concave, then maximization will make sense. It is a function in two variables, so we can verify concavity by looking at its Hessian. That is one straightforward way of doing it, okay. You could also complete the squares and show that it is a concave function. So, all of that.

So, let us assume that we have checked concavity. So, let us do this. I have done a little bit of algebra to speed this process along. There are a whole bunch of minus signs that you can see in the definition of $q(\lambda)$. So I am going to pull those minus signs out. If I pull a minus sign outside, what happens to maximization? It becomes a minimization, right?

So, I am going to skip a few steps of algebra just in the interest of time and remove that minus sign, which flips the max to min, okay. So that is all I have done. So, I am going to write the final expression that I get.

Correct. In order to get $q(\lambda)$, I cannot verify that independently unless you tell me what λ is. Correct. So, in the way that we have formulated the dual problem, which we did without proof, but just with geometric intuition, we did not impose something extra. It turned out that the second recipe gave me the same answer.

Correct, correct. So, we will have to check once again whether, I mean we will check whether or not this, okay. We have proved, yes. So, why am I doing it again? Just for practice. We know that $q(\lambda)$ is going to be concave, but it's a simple enough example for you to verify.

You don't have to verify it each time. So when does it help to verify? It will help to verify if you have made some simple mistake in algebra. So, some simple steps, these are the things which help to catch your own error that you see, oh, hey, wait, this is not turning out to be concave. Okay. So, this is what this whole problem ends up being, okay.