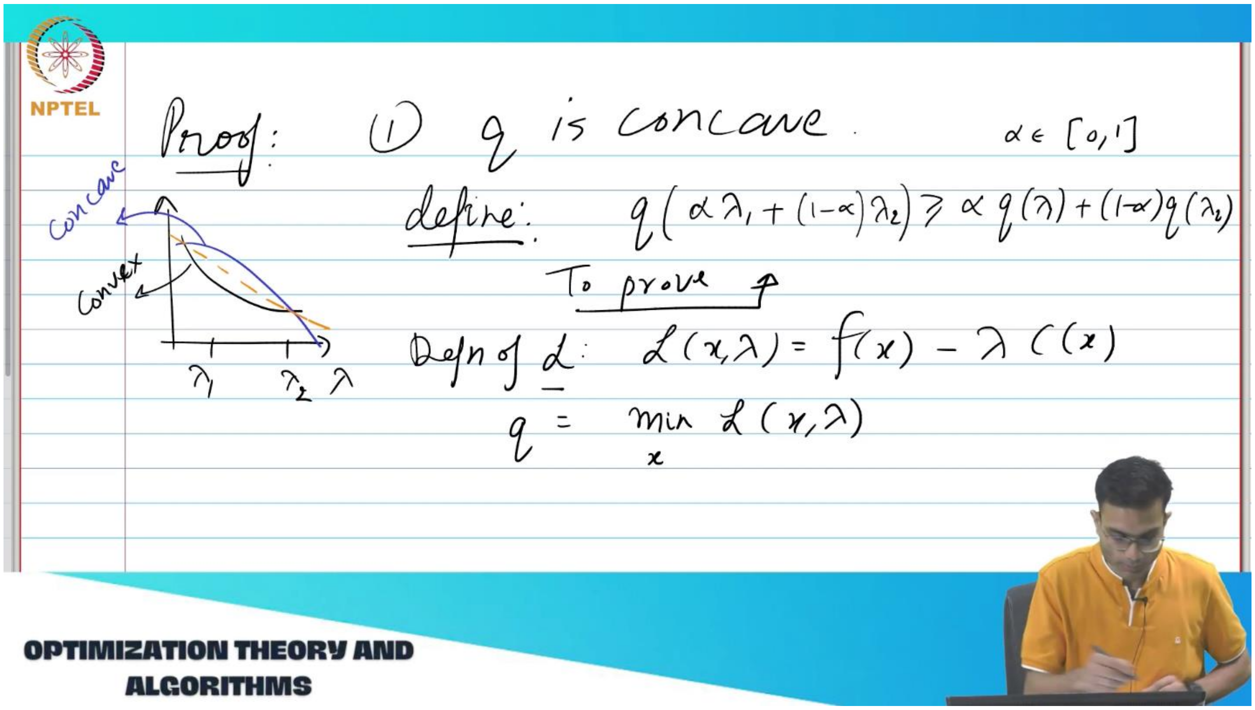
Course Name: Optimization Theory and Algorithms
Professor Name: Dr. Uday K. Khankhoje
Department Name: Electrical Engineering
Institute Name: Indian Institute of Technology Madras
Week - 12
Lecture - 80

Proof of concavity of the dual problem

So, we were talking about the Lagrangian dual function q and we discussed two very interesting properties of it. One is the fact that this function is concave and the other that this function domain is convex. And what is unique is that no matter what function you take, whether the function is convex, non-convex, linear, non-linear, these two properties hold. And this is why duality is a very useful property. So, what is the implication? Why do we care? Implication, the take-home message from here is that solving the dual problem, can you say that is always a convex optimization problem? Why? Because $-q$ is convex. So remember in constraint optimization how do I define a convex optimization problem? What were the two criteria? Function is convex and the constraint set is a convex set, both of these are needed.

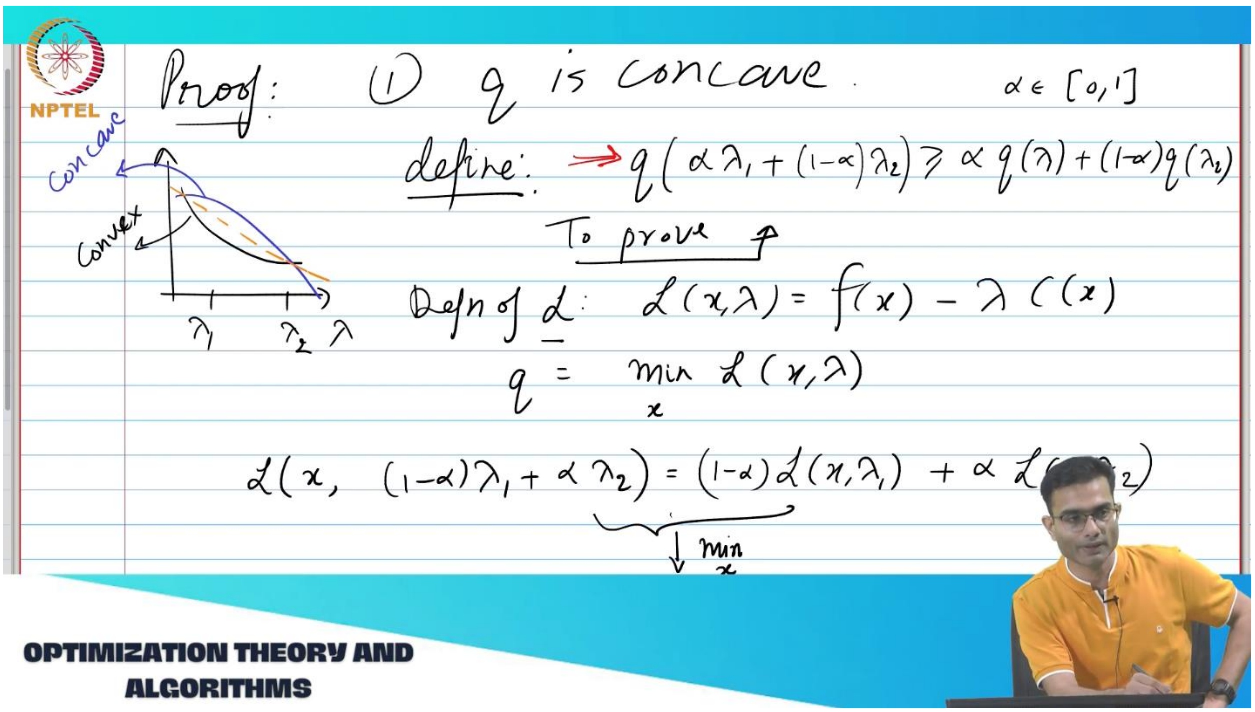
So this is some place where you have to be careful. If your function is convex but your constraints give a feasible region which is not convex, then this problem is not a convex optimization problem. So, be careful of this one small catch, ok. So, now let us prove these statements.



They are actually very, very simple to prove, ok. So, no big tricks involved over here. The first thing that we have to prove is q is concave, right. So we can basically go back to our very simple definitions of convex functions, right? So supposing let us take λ as the parameter here. There are two points over here.

And I have drawn three possible functions over here. So, which is a convex function among these three? Black, blue, orange. So, strictly speaking orange also because it is the equality. So, this is our convex and then our, this is our concave. So, how would I define a concave function? So, let us say this is λ_1, λ_2.

So, how would, what would be the definition of a concave function? This analogous to convex, but what is the difference? Correct. So, in words the concave, if I, the function of, function value lies above the concave the line joining the two points right. So, it would be something like, so if I take $q(\alpha_1\lambda_1 + (1-\alpha)\lambda_2)$, this should be greater than or equal to $\alpha q(\lambda_1) + (1-\alpha)q(\lambda_2)$, right? And is there any slight something that we are missing? Is there a range on α? Right, $\alpha \in [0,1]$.



This is the definition of a concave function, ok and this is what we have. So, we have to prove this ok. So, remember the definition of the Lagrangian is what?

$$L(x, \lambda) = f(x) - \lambda c(x)$$

This is how I define the Lagrangian, right? And how was q defined? Minimize this with respect to x, right? So, with these two definitions in mind I have to work on proving this concavity idea. So, let us take one big Lagrangian. So, let us take Lagrangian.

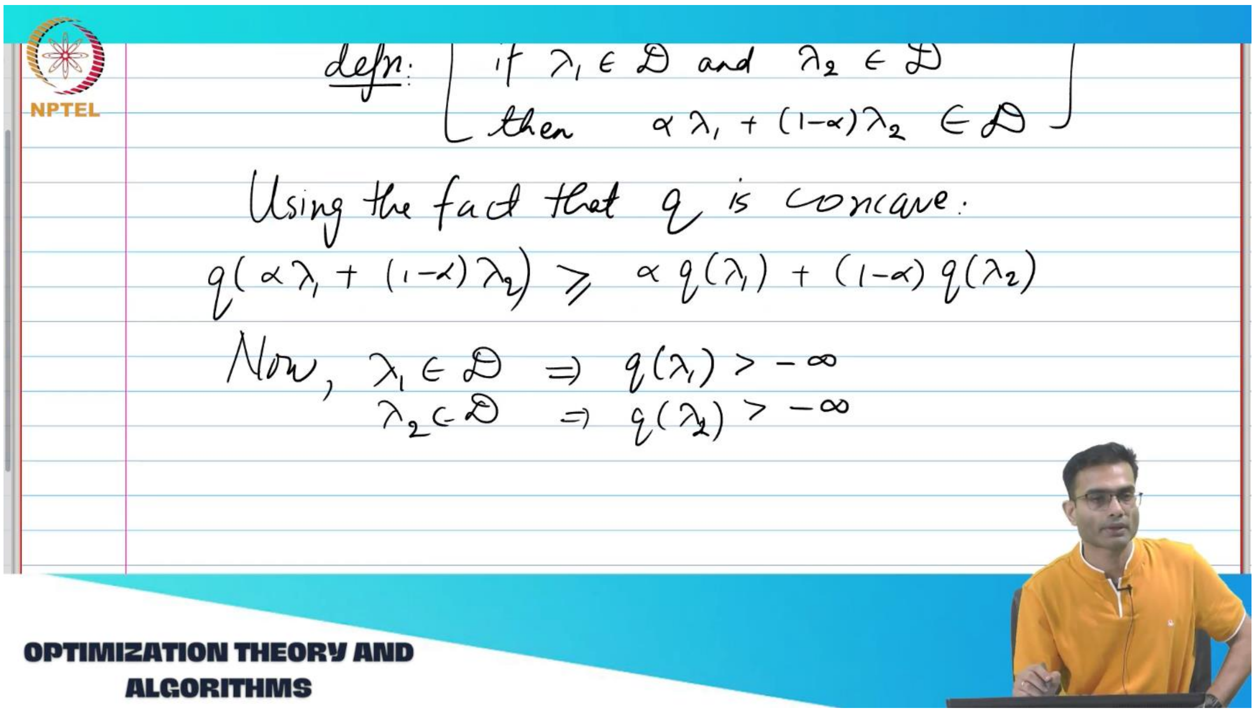
What is the parameter that I should be focusing on x or λ? λ, right because when we are talking about concavity of q in terms of $q(\lambda)$. So I am interested in what happens to λ. Anyway, x gets factored out over here because I am minimizing with respect to x. I will be left with a function in λ. I have to prove something in the λ-domain right.

So I should take a linear or a convex combination in the λ space. So that would be something like

$$(1-\alpha)\lambda_1 + \alpha\lambda_2$$

This is one big Lagrangian that I can take, right? So if I substitute this into the definition of the Lagrangian, will it split into two Lagrangians? Yes, because it is linear in λ, it will just split up. So what will the first Lagrangian be? $(1-\alpha) \times L(x, \lambda_1)$. It is simply linear and then this becomes $\alpha \times L(x, \lambda_2)$.

So, this is true. Now to reach the Lagrangian dual from here, what operation do I apply on both sides? I have the correct answer, what is it? I have to minimize both sides with respect to x that will give me the Lagrangian. So, if I take this equation and do a minimize with respect to x. So, what will become the left-hand side? q of this whole thing



$$(1-\alpha)\lambda_1 + \alpha\lambda_2$$

just by definition. What happens to the right-hand side? Are you sure? So, one answer is that this becomes

$$(1-\alpha) \times q(\lambda_1) + \alpha \times q(\lambda_2)$$

So, before we go that maybe I just need to mention one thing. What is this? Min of $a + b$ and min of $a + b$? What is the relation between these two? Equal, greater, less than? I hear greater, I hear less. So, supposing I take $a = x$ and $b = -x$, then what will happen? The left-hand side will become 0.

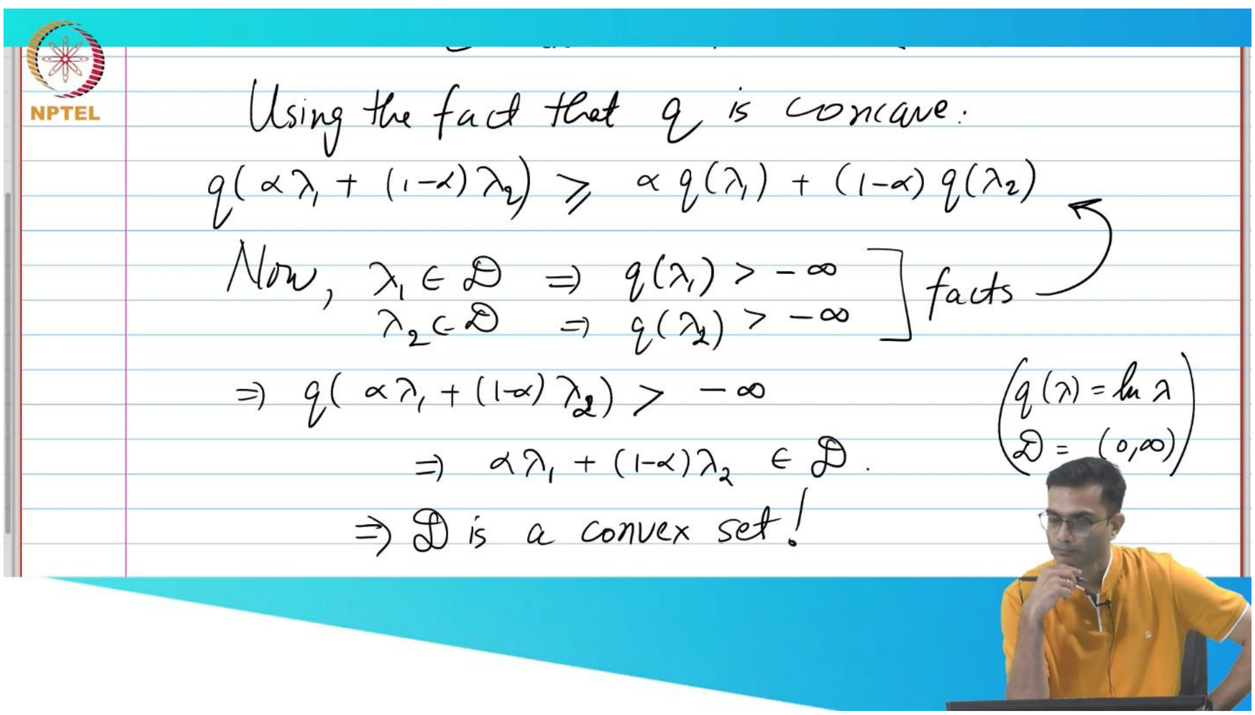
What is? I am minimizing, min means not, I mean min is standing for minimize with respect to, that is a minimization operation. Correct, a and b are functions. Correct, a and b are functions of x. So, then what would you say? Which way should it go? Okay, I hear less than equal to, I hear greater than equal to.

Take some, I mean. So, it can go lower because each of these terms I can minimize separately. So, this is actually going like this. So, this is actually, so on the left-hand side I have q of this expression and on the right-hand side I do not have q of this plus q of that. Instead, what will I have? This will become greater than or equal to, now I can apply the minimization on each of these separately. So, this would become

$$(1-\alpha) \times q(\lambda_1) + \alpha \times q(\lambda_2).$$

Actually, we are done. This is what we wanted to prove. If you look back at the definition of concavity, here is the definition of concavity, right? The function graph lies above the convex combination of the function values, right? So this is $q(\alpha_1\lambda_1 + (1-\alpha)\lambda_2)$, right? So this is, whatever, you can see what the difference between this and this is. So, therefore we get. So, it is actually a very simple proof, does not involve any complicated proof by contradiction or this, that, and the other.

So, did everyone follow this? This is something that you can convince yourself by just taking a few examples that this inequality is going to be this way. So, we have got that it is a concave function. What is the next thing that we have to prove? What is the second thing? The domain of D is convex. Okay, so again let us go back to the definition of, so this is the definition of, I mean this is convexity of a, not of a function, but what? For set, right? This is, so if I give you some region like this, is this a convex set? No, because there is some convex combination of two points which does not lie in the set, right. So, if I take a line like this from here to here, all of these points over here are from here to here are outside.



So, this is not a convex set right. So, we have to use. So, if I were to define it, when do I say that a set is convex? I would say if λ_1 belongs to D and λ_2 also belongs to D, then what should hold true? Convex combination of this, so $\alpha\lambda_1 + (1-\alpha)\lambda_2$ must also belong to D, right? This is the definition of convexity which we want to prove, ok.

So, let us go to prove it. We will use the fact that q is concave, which we just proved. So, if I take

$$q(\alpha\lambda_1 + (1-\alpha)\lambda_2),$$

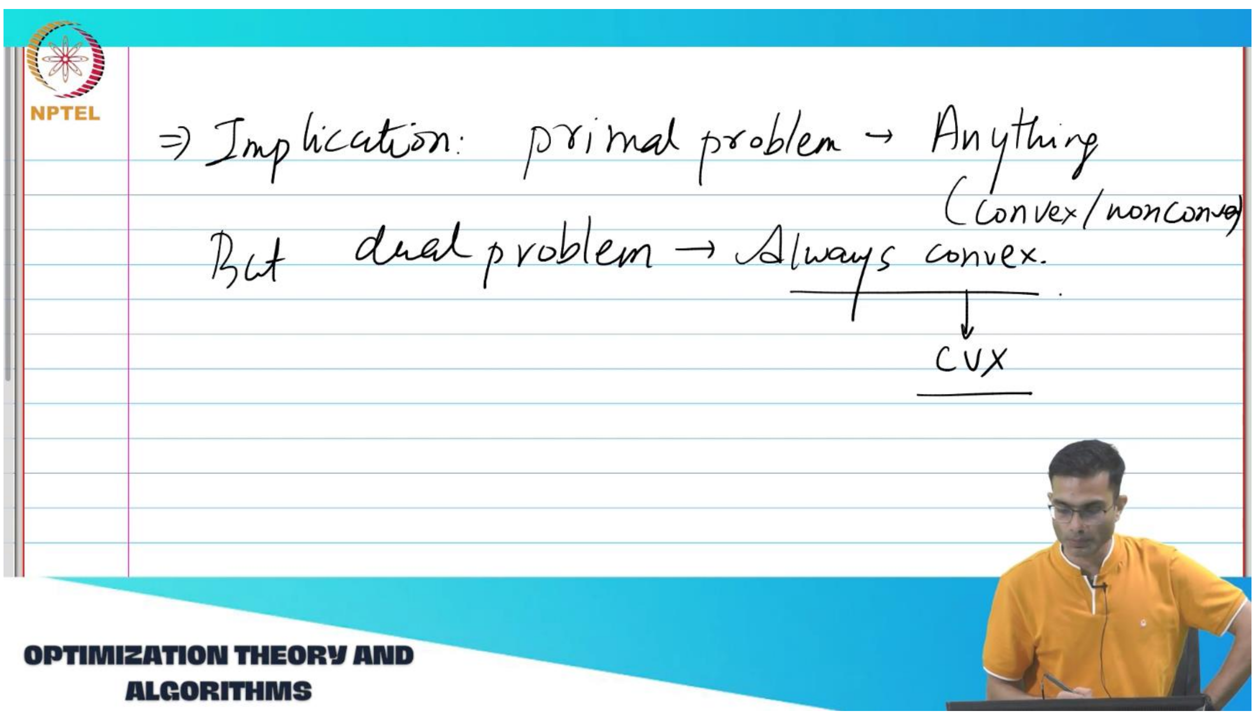
this should be greater than or equal to, I am using the property of concavity,

$$\alpha \times q(\lambda_1) + (1-\alpha) \times q(\lambda_2).$$

Now we are saying that λ_1 is in D. We are saying λ_2 is in D, right? So we are saying that λ_1 belongs to D. What is the meaning of λ_1 belongs to D? Operationally, what does it mean? Just go back to the definition of the domain of D.

When do I say that λ is in D? What should happen? $q(\lambda)$ should not blow up. It should be greater than $-\infty$. So if I say that λ_1 belongs to the set, what does it imply? What does it say for $q(\lambda_1)$? Simply greater than $-\infty$. Inwards, it does not blow up. Similarly, if λ_2 belongs to this, it implies that $q(\lambda_2)$ also does not blow up.

So now if I take these two facts, and plug it into this, what can I say about $q(\alpha\lambda_1 + (1-\alpha)\lambda_2)$? This is going to be greater than $-\infty$, and this is all that I need to say that the convex combination $\alpha\lambda_1 + (1-\alpha)\lambda_2$, therefore this also belongs to D, right. So, therefore D is a convex set. So, did everyone follow this very simple logic? We used the previous fact that the function is concave. Concave means that the function value is above the convex combination of the function values and that is it. If q, if λ_1 is in D, if λ_2 is in D, then put together this implies that their convex combination also is in D.



Means it is a finite number. So, ok. So, what that means? It is a finite number. For example, if I ask you what is $\log(x)$ at 0? So, it is not a number. So, if I had, if $q(\lambda)$, supposing $q(\lambda)$ were equal to $\log(\lambda)$, let us say in some problem. So, in this case what is the domain of q? So, in this case the domain would be the open interval $(0, \infty)$ because 0 is not allowed.

q blows up. So it is excluded from the domain, but any positive quantity is fine, right? So that is the meaning of saying that those λ for which $q(\lambda)$ is greater than $-\infty$. It is a funny way of saying it, but it is a shorter way than saying a finite number or whatever. Fine? So what is the, I mean I have already said this before, but what is the implication? The implication is that the primal problem can be anything, by anything I mean convex, non-convex, but the dual problem is always convex. So that's a very useful fact to know. In fact, previously we had taken an example, right? When we saw that we took some simple example, we saw that the primal approach gave us one solution, we saw the dual approach gave us the same solution.

So, the primal problem is difficult to solve, but I can formulate its dual, and if I can solve it, I have a second way of arriving at the solution. So, that is the power of it. So, there is something I want to mention at this point which we will not cover in the course but is extremely useful, particularly for those of you who will go into numerical applications of optimization. There is a tool called CVX. Who has heard of CVX? Okay, a fair number of people have heard of it.

So, CVX is a tool built by a bunch of people at Stanford. It's a MATLAB toolbox available for convex optimization. So, in that, you can — the way we speak in words is almost how you can specify the problem that minimizes. So, you define f, then you say minimize f subject to $x \geq$ blah or $g(x) \geq$ something. So, the way we speak in words, you can specify your convex problem in it.

And so, it means that given any optimization problem, whether or not you can solve the primal problem, you can always solve the dual problem using tools like this. And like CVX, there are other tools, but CVX is the most popular tool. But keep in mind that CVX is extremely slow. So, once you know that something is working, it's always better to move to your own code implementation of it. But for example, for your course project, if your problem happens to be convex and you want to do some programming, CVX is a very convenient tool to use.

Just install in MATLAB and you are good to go. It is also there for Python, I believe. Okay, so that is a slight detour into CVX. Any questions on this? Okay, we are coming right to that. So, his question is that if the problem is not convex, then what is the meaning of the dual, the solution to the dual problem? Does it hold any meaning for the solution of the primal problem? Okay, so that is exactly what I am going to talk about next.