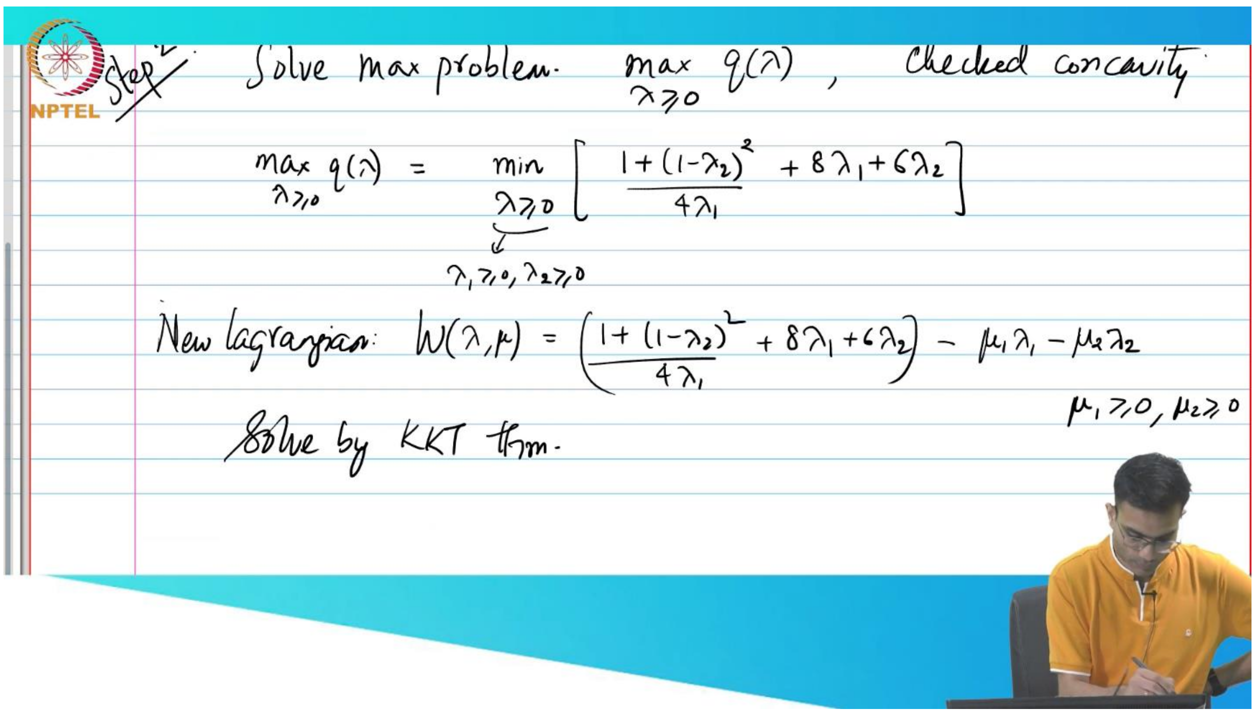
Course Name: Optimization Theory and Algorithms
Professor Name: Dr. Uday K. Khankhoje
Department Name: Electrical Engineering
Institute Name: Indian Institute of Technology Madras
Week - 12
Lecture - 82

Proof of concavity of the dual problem - Part 3

So I need to minimize this. Now this looks like a slightly messy problem, right? Why? Because λ_1 is in the denominator also, it is not a very nice function. So this is where you know having a little bit more patience is required in working with a dual problem. How will I solve this? We solve this as though it's a new primal problem. So I have a function in two variables, λ_1, λ_2. Is there a constraint on this variable? Yes.

So this simply means that $\lambda_1 \geq 0$, $\lambda_2 \geq 0$, right? So if you were faced with this as a fresh primal problem, what would you do? You would formulate its Lagrangian and then solve it using KKT conditions. So let's do that. So the new Lagrangian, Let's call it w. This is going to be a function of lambdas and let's say μ.



So μ are the new Lagrange multipliers of this problem, right? So what will I get? So I'm going to get, what is the, so function minus $\lambda_i c_i$. What are the constraints in this problem? In terms, if I write in terms of $c(\lambda)$, what are the constraints? $c_1(\lambda)$ is what? λ_1, $c_2(\lambda)$ is λ_2, right. So, I am going to write this whole expression this is my $f(x) - \mu_1\lambda_1 - \mu_2\lambda_2$. This is the Lagrangian of the new problem right. So, this is going to be:

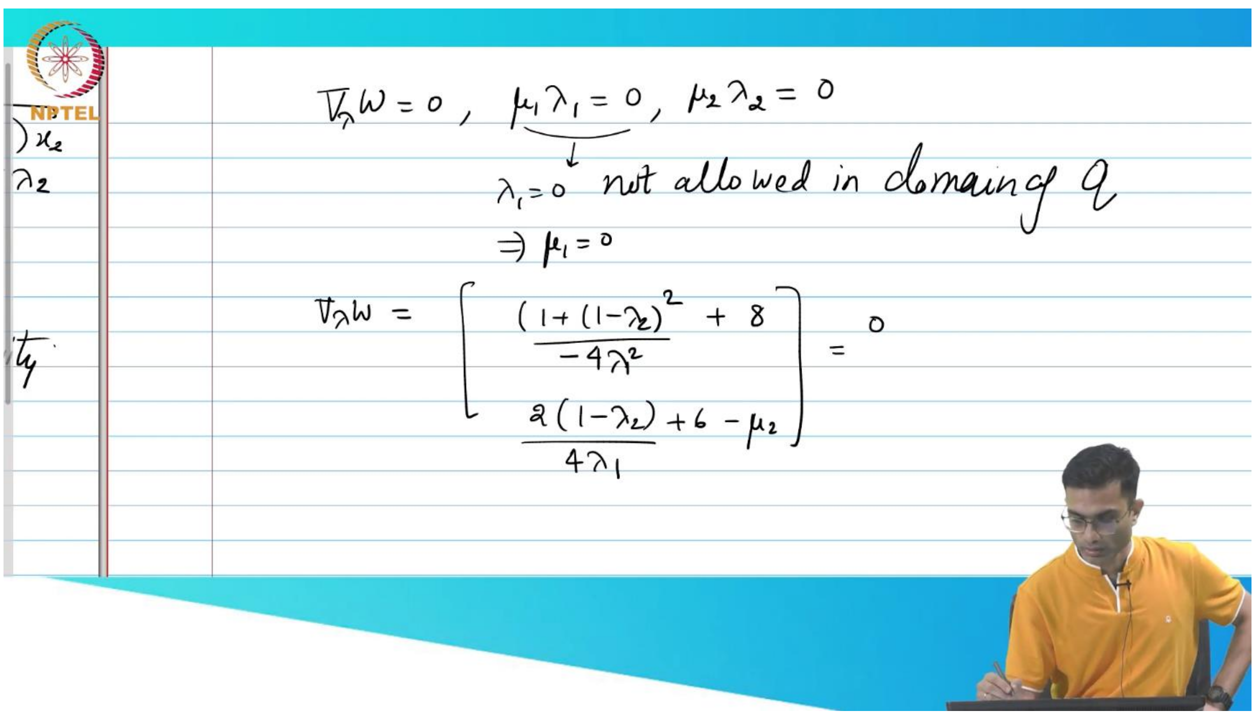
$$1 + 1 - \frac{\lambda_2^2}{4\lambda_1} + 8\lambda_1 + 6\lambda_2$$

This is the first part minus $\mu_1\lambda_1 - \mu_2\lambda_2$. Okay. Is there a constraint on μ_1, μ_2? Should be greater than or equal to 0 because these are inequalities. So, $\mu_1 \geq 0$, $\mu_2 \geq 0$, okay, solved by the KKT theorem. So, let us do that. So everyone is following what we are doing? We have got a second problem which needs to be taken care of.

So we are taking care of it. Okay. So KKT will tell me what this Lagrangian with respect to what variable will I minimize? λ. Right. So this should be equal to 0.

What do the complementarity conditions tell me? μ_1 which is a Lagrange multiplier multiplied by the constraint, what is the constraint? λ_1 should be equal to 0, $\mu_2\lambda_2$ should be equal to 0, ok. So, let us before we find this gradient, let us look at this, $\lambda_1 = 0$, is it allowed? actually it is not going to be allowed because why? Because $\frac{1}{\lambda_1}$ is there right. So $\frac{1}{\lambda}$, so therefore $\lambda_1 = 0$ needs to be excluded from the domain of q. So that is why $\lambda_1 = 0$ is not allowed right. So we will just make a note over here, $\lambda_1 = 0$ not allowed in domain of q.

If $\lambda_1 = 0$ is not allowed therefore what is the implication for μ_1? Therefore μ_1 has to be 0, ok. That is great because that will help me to make my Lagrangian solving a little bit easier. So what is this Lagrangian? Let us look at the expression you have for w and take a derivative with respect to λ_1. What is the first term that you will receive? λ_1 is only in the denominator, correct? Yes, I will get a minus $\frac{1}{\lambda^2}$. So, this is going to be:



$$1 + 1 - \frac{\lambda_2^2}{-4\lambda_1^2} + 8$$

This is the derivative with respect to λ_1; what will I get with respect to λ_2? I am going to get:

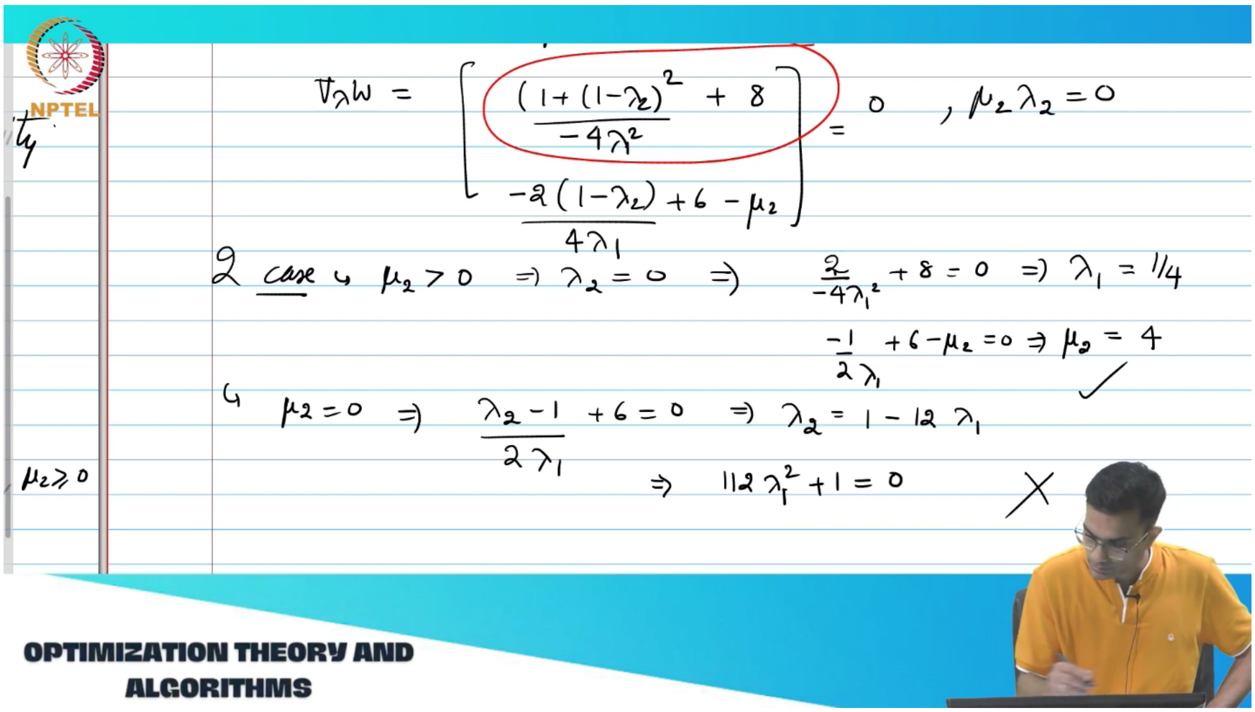
$$2 - \lambda^2\lambda + 6 - \mu_2 = 0$$

Ok, and also, $\mu_2\lambda_2 = 0$, right? So, this $\mu_2\lambda_2$ again gives us two cases, right? So, in two cases, $\mu_2 > 0$, which implies that $\lambda_2 = 0$. Okay? So, in this case, what is my solution? So, $\lambda_2 = 0$, right? So, what happens to the first term?

$$\frac{1}{-4\lambda_1^2} + 8 = 0$$

Okay. Is there anything left on the second term? $\lambda_2 = 0$. So, subtract 2 from 4. So, oh sorry, that is the first term.

The second term is going to give me what? $-\frac{2}{4\lambda_1}$, so that is $-\frac{1}{2\lambda_1} + 6 - \mu_2 = 0$, right? So, from the first expression, do I get a value of λ_1? Right, does $\frac{1}{4}$? And from here, what value of μ_2 do I get? 4. Is $\lambda_2 = 4$ allowed? It is a Lagrange multiplier for the new problem. It was an inequity. So, it had to be greater than or equal to 0. So, it is allowed, right? So, this is legitimate, okay.



So, this was easy to do. The second case is when $\mu_2 = 0$, okay. So if $\mu_2 = 0$, what does the first expression give me? If I take the gradient, what will I get? Or I can just get λ_2 directly from the second part of the gradient, correct? So, what will I get?

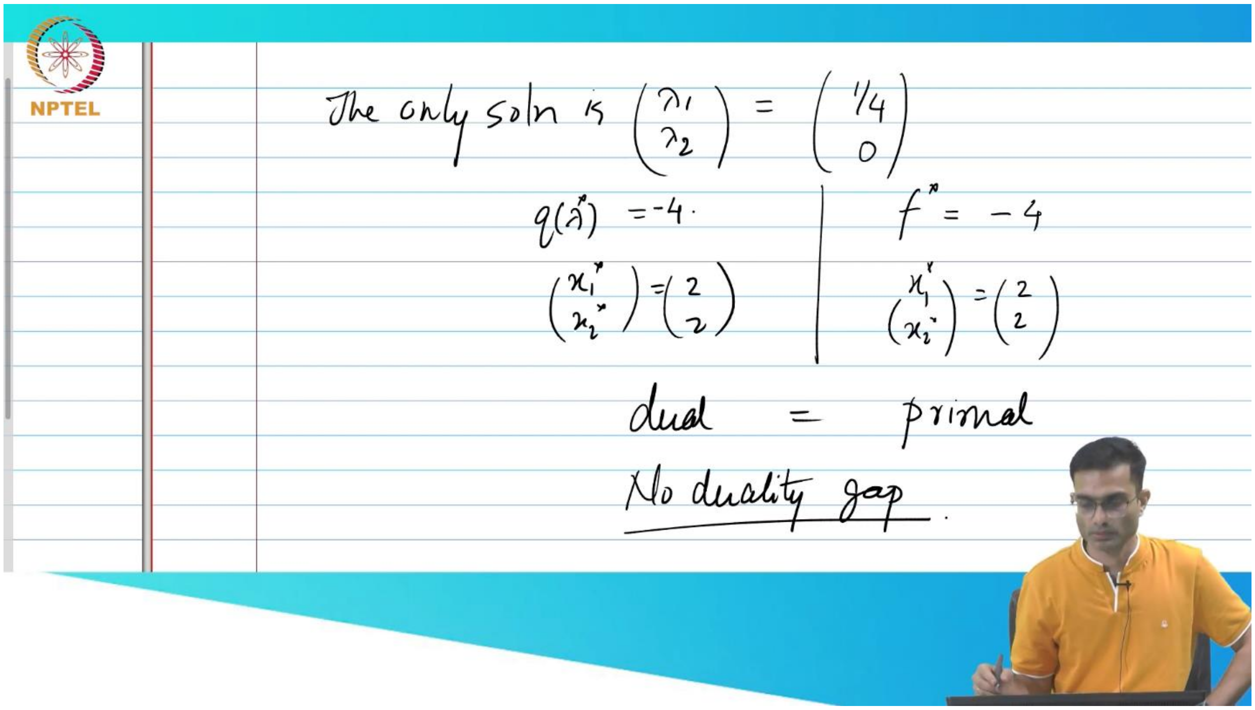
$$\lambda_2 = \frac{1}{2\lambda_1} + 6 = 0$$

Because $\mu_2 = 0$. So that implies that $\lambda_2 = 1 - 12\lambda_1$.

Now I can take this value of λ_2 and substitute it in. If you look at the first part of the gradient over here, it is purely λ_1 and λ_2. I have a relationship for λ_2 in terms of λ_1. I can put it back inside. So, yes, if λ_1 is correct, then that's right.

So, we can stop at this one. Are we required to have λ_1 and λ_2 greater than 0? Yes, that's correct. So this shows us that we need not proceed further because any λ_1 greater than 0 does not apply. What if λ_1 is between 0 and $\frac{1}{12}$? That's a possibility, isn't it? So we actually have to go ahead and solve it. So when you substitute this back, I will spare you a couple of steps of algebra.

What you get is that this is the solution for λ_1. So, what is the λ_1 that satisfies this? It's imaginary, right? So, obviously, that is not feasible. So this case, where $\mu_2 = 0$, does not work. I am only left with $\mu_2 > 0$.



So, let us see. We did not calculate it, so in this case, 1. Okay, we already know λ_2, right? So, let's create a new page. So, the only solution is λ_1 and λ_2. What was the value of λ_1? $\frac{1}{4}$. And what is the value of λ_2? 0, right? This is the only solution that works.

So now, what is $q(\lambda)$? Let us look back at this. Where has our $q(\lambda)$ gone? This expression right here, this big expression right here. So, in this, I am going to substitute λ_1 with $\frac{1}{4}$ and λ_2 with 0, okay? So, you can do it mentally because $\lambda_2 = 0$ and $\lambda_1 = \frac{1}{4}$. What is the optimum value that you get? There should not be a minus sign.

Oh, I'm sorry. Ah, right. And we had a relationship for x_1 and x_2 in terms of λ, right? If you go back, where did that go? Here is $x_1^*, x_2^* = \left(\frac{1}{2}\right)\lambda_1$ and $(1-\lambda_2)(2\lambda_1)$. So, if I substitute $\lambda_1 = \frac{1}{4}$ in this, what do I get for x_1^*? 2, 2. And if you recall our primal problem, what was our f^*? -4 and $x_1^* x_2^*$ was also equal to 2, right? So, this was the only KKT point that worked correctly throughout. So, here you have the dual equal to the primal right, and that is not a surprise because the objective function, I mean, the optimization problem is convex, so there is no duality gap.

So this was just to provide you with an example of problem solving using the duality method. So you can see that there can be many solutions in between that arise due to the various cases. But you have to eliminate them based on the previous, the step knowledge of the previous step. For example, the domain of q should exclude places where it blows up, and so on.

So that helps you rule it out. So, if we just do a quick refresher on what we did, we took a very simple problem like this. Minimizing $-x_1 + x_2$ with these constraints. We saw that the solution graphically is correct, right? The optimum value is -4 at the point $(2,2)$. Then we took the dual problem formulation; the objective function remained as it is, which is here, and I had the constraints over here. To get to the Lagrangian dual, step 1 was to minimize the Lagrangian with respect to x, which I did over here.

So, by taking gradients with respect to x_1 and x_2, I arrived at this relationship between x and λ. I substituted that back into the Lagrangian dual, and I obtained one big expression for the Lagrangian dual. Fine. Then, that allows us to move on to step 2 of duality. Step 2 of duality is to maximize the Lagrangian dual norm. So I take this problem and solve the maximum problem.

I did some algebra to convert the maximization into minimization, and this is the expression that I obtained. This, in itself, is an optimization problem. So, I applied KKT and formed a new Lagrangian for this function. I called it W, and the Lagrange multipliers are μ_1 and μ_2, right? So, this is my $f(x)$, and this is my $\lambda_i c_i$, right? This is how we write the Lagrangian function with respect to W. So, that is what I have done here: $f(x)$ is as it is, and I have my μ_1, λ_1, μ_2, and λ_2.

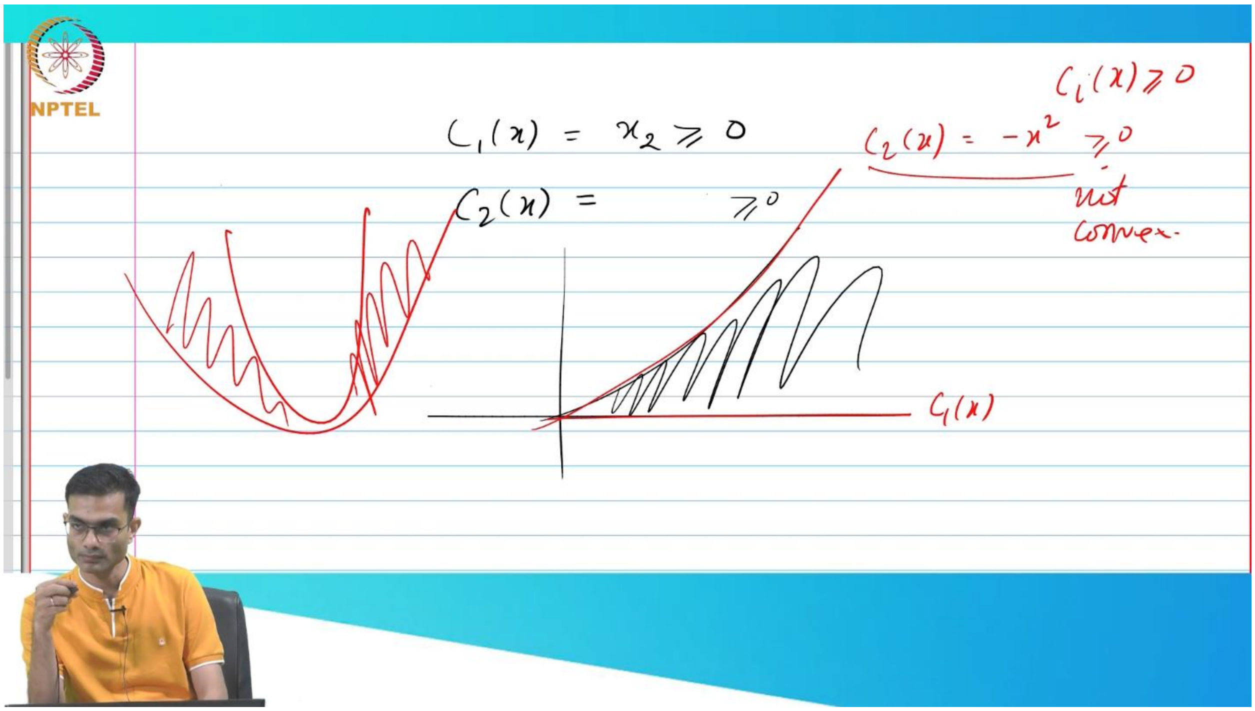
Okay? But with the constraints that μ_1 and μ_2 have to be greater than or equal to 0 because we are dealing with inequalities, okay. So, we solve this using KKT by setting the derivative with respect to the variable, which is λ, equal to 0. So, with respect to λ, I took the gradients and essentially solved for it. We encountered some cases. In case 1, I found the correct solution.

In case 2, I looked at something that led me to an incorrect solution. So, I eliminated, "Oh, we not only talked about lambdas being real, but we also said they should be non-negative, right?" Why should lambdas be non-negative? Correct, that condition is already present in λ, so being complex is obviously ruled out, right? So that’s how you would solve a problem using duality, step by step, right? So, it requires a little more patience because you’re solving two optimization problems, right? So, if the primal problem is easier, you should solve that. If the dual problem is easier, you should solve that, okay? You can always use some technology. CVX is a technology, isn’t it? Because it’s an automated set of rules that solves your problems and provides you with the solution. So, as homework and to learn a new tool, go back and install the CVX toolbox on MATLAB or Python, and try to solve this problem, you know.

It will provide you with a lot of extra information. It will explain what the duality gap is, what the Lagrange multipliers are, and the intermediate steps. So, it is also a nice way to see what is going on behind the scenes in solving these problems, okay? But just keep in mind that if you start applying this to real-world problems, it is very slow. So keep that in mind. That is the definition of a convex optimization problem.

The objective function is a convex function, and the constraint set is a convex set. Okay, that is a good question. So, if my constraints are convex functions, does that mean that the feasible set is a convex set? Please say that again. Correct, no; however, the feasible set is now being defined by multiple convex functions. The correct sentence is: "$c_1 x \geq 0$, and $c_2 x \geq 0$; these are convex functions."

Does it imply that the feasible x is? So, let us take an example to see if that always holds true. So, let us say that $c_1 x = x_2$, which is greater than or equal to 0. Okay? Let us say that $c_2 x = 2x$ and $e^x - 1 \geq 0$, right? So, no, that does not make sense, okay? So, what if this is my convex function, and let us say my feasible set becomes this? Then, let us say this is my $c_1 x$, and let us say that this is my $c_2 x$. Is it possible? So, c_2, the intersection that I have here, is not convex, but how will I define $c_2 x$? So, the constraints have to be defined such that $c_i x \geq 0$, right? So, c_2 is not, right? So, this is more like $-x^2 \geq 0$, right? So, you are correct. So, if the functions are convex, then the intersection seems to be convex.



So, this is not a good example because the constraint function is not convex. So, you have a counterexample where the c's are convex, but the feasible set is not convex. Yeah, but considering a parabola is not enough because the constraint has to be that the function is greater than or equal to 0. So, if I take two parabolas—this one and this one—to define, and you are thinking of this as the feasible set, right? So, one of the constraints will be concave, while the other constraint will be convex. So, that means it is not a counterexample, right? So, you're saying this was one of the tutorial problems, right? Intersection of convex sets.

Yeah, but here we are talking about functions. No, I can define a function $f(x) = x_1^2 + x_2^2$. But I am telling you that this is my domain. It is up to me to define the domain. I said that this is the domain of the function. In this domain, the function is $x_1^2 + x_2^2$.

So, the domain is concave; I defined it to be so, but the object, the function, is convex. So, I have a convex function on a concave domain, right? So, it is not necessary that just because the function is concave, the domain also has to be convex. What is concave? The c's. Correct,

correct, right, right, right, and right. So if it is a con, then the intersection of concave functions is also a con; it seems like it, right? Correct, correct, correct.

Let us look for a formal statement of this. I mean, by arguing amongst ourselves, we can only come up with some counterexamples. So, I do not want to make a definitive statement. We will, or maybe I will, post on Google Classroom what the conclusive information is about it. It is convex. It is convex, right? convex function He is talking about the intersection of the domains; you cannot intersect functions, but you can intersect the domains.

I'll post a definitive answer for this, okay? The domain of what? The feasible region is concave. Then it is not a convex optimization problem. Right away. The domain is concave; I mean, the domain is non-convex.

Like this. So, you have to take care of the projection operation. What we have defined in class is projection onto convex sets. That is a well-defined operation. A projection onto a non-convex set is, first of all, lacking uniqueness.

So, that is the tricky part. So, people may perform tricks such as expanding their domain to make it convex in a minimal way, and then solving it, and so on. Yes, I will look it up, and there may be other ways to define a convex set besides this one. If that point happens to be in the interior, that is fine. If you are in the interior of the domain, then the constraints are inactive; it is like an unconstrained optimization problem. In fact, the solution to this problem is 0, which is the interior of the domain.

So, it does not matter what the domain set is, right? But if it is on the boundary, that may be the tricky part, right? Because of that boundary point, how did it come to be? Did it get projected from somewhere else? Okay, let's say you solve it numerically. Fine. Which is strictly inside the set. Then you are done. Yeah, but I doubt these things happen very often in real life.