

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

Lecture - 18B
Strong duality in convex optimization - II

(Refer Slide Time: 00:23)

Consider
 $\min f(x)$
 $g_i(x) \leq 0 \quad i=1, \dots, m$
 $Ax=b \quad | \quad Ax=b=0$
 $A \in \mathbb{R}^{m \times n}$, full row rank, $\text{rank}(A)=p$.

Suppose $f, g_i(x)$ are convex.
 Suppose $p^* > -\infty$ and attained.

Claim If $f, g_i, i=1, \dots, m$ are convex,
 then g is convex.

proof let $(u^1, v^1, t^1) \in G$.
 let $\alpha \in [0, 1]$. $(\bar{u}, \bar{v}, \bar{t}) = \alpha(u^1, v^1, t^1) + (1-\alpha)(u^2, v^2, t^2)$

We need to show that $(\bar{u}, \bar{v}, \bar{t}) \in G$.
 i.e. $\exists \bar{x}$ s.t. $f(\bar{x}) \leq \bar{t}$, $g_i(\bar{x}) \leq \bar{u}_i$
 $A\bar{x} - b = \bar{v}$

Clearly we have that $\exists x^1, x^2$ s.t.
 $f(x^1) \leq t^1, g_i(x^1) \leq u_i^1, Ax^1 - b = v^1$
 $f(x^2) \leq t^2, g_i(x^2) \leq u_i^2, Ax^2 - b = v^2$

take $\bar{x} = \alpha x^1 + (1-\alpha)x^2$.
 $f(\bar{x}) \leq \alpha f(x^1) + (1-\alpha)f(x^2)$
 $\leq \alpha t^1 + (1-\alpha)t^2$
 $= \bar{t}$

similarly $g_i(\bar{x}) \leq \alpha g_i(x^1) + (1-\alpha)g_i(x^2)$
 $\leq \alpha u_i^1 + (1-\alpha)u_i^2$
 $= \bar{u}_i$

$A\bar{x} - b = A(\alpha x^1 + (1-\alpha)x^2) - b$
 $= \alpha [Ax^1 - b] + (1-\alpha)[Ax^2 - b]$

Ok. So, now, let me actually write out a convex optimization problem. So, consider this convex optimization problem, minimize $f(x)$ subject to $g_i(x) \leq 0$ and $Ax = b$. So, i going from 1 to m $Ax = b$. So, A now is a matrix in $\mathbb{R}^p \times \mathbb{R}^n$, so it has p rows and we will assume that A is full row rank, ok.

So, in short rank of A is equal to p. Now, we will assume. So, suppose f and all these g i's are convex, ok. And suppose p star is greater than minus infinity and attained, ok. So, that means, the optimal value of the primal is finite and it is attained by some point x that is feasible.

Now, first little observation that is quite easy to make is that, and I will leave that as a claim; it is very easy to prove that, if f comma g i's i going from 1 to m. If these are convex, then G, the set G is convex, ok.

(Refer Slide Time: 02:40)

$$\begin{aligned} \min_x & f(x) \\ & g_i(x) \leq 0 \quad \forall i=1 \dots r \\ & h_j(x) = 0 \quad \forall j=1 \dots p \end{aligned} \quad (P) \quad \text{opt. value} = p^*$$

$$G = \left\{ (u, v, t) \mid \exists x \text{ s.t. } f(x) \leq t, g_i(x) \leq u_i, h_j(x) = v_j, \forall i=1 \dots m \right\}$$

$$\text{Optimal value of } (P) = \inf \{ t \mid (0, 0, t) \in G \}$$

$$D(\lambda, \theta) = \inf_x \left\{ \sum \lambda_i g_i(x) + \sum \theta_j h_j(x) + t \mid (u, v, t) \in G \text{ for } \lambda \geq 0 \right\}$$

$$= \inf_x L(x, \lambda, \theta)$$

Optimal value of the dual problem = maximum intercept of any supporting hyperplane to the set G.

Non-vertical supporting hyperplane at $(0, 0, p^*)$

Diagram labels: G , t , p^* , $J(x)$, u , v .

So, to if you want to see this, is not hard to show; but if the if these. So, how does, one show this if you, maybe we can quickly take a look here.

So, if, so what we need to argue is that, if u_1, v_1, t_1 and u_2, v_2, t_2 are in are in G; then every any convex combination of thereof is also in d. So, maybe I will just show you the

proof of this. Let quickly do a proof of this. So, let suppose u_1, v_1, t_1 and u_2, v_2, t_2 ; suppose they both belong, so they all belong to G .

Then ok and let α in $[0, 1]$ and define say $\bar{u}, \bar{v}, \bar{t}$ as $\alpha u_1 + (1 - \alpha) u_2$, $\alpha v_1 + (1 - \alpha) v_2$, $\alpha t_1 + (1 - \alpha) t_2$, ok. We need to show that $\bar{u}, \bar{v}, \bar{t}$ belongs to G , ok. So, how would you show that? Well, you look at \bar{u} .

So, what does it mean to show this? We need to show that there exists; that is we need to show there exists an x , such that $f(x)$ is less than equal to \bar{t} , $g_i(x)$ less than equal to \bar{u}_i and $Ax - b$ is equal to \bar{v} , alright. So, what have I done here? See notice that this is my, this was my optimization problem in the original form; so to bring it in the form as before, I have basically written this as $Ax - b$ equal to 0, ok.

So, I had my, when I took points u_1, v_1, t_1 and u_2, v_2, t_2 ; the 0 got replaced by v_1 and v_2 respectively and now for this particular point I am looking for \bar{v} , ok. So, since u_1, v_1, t_1 and u_2, v_2, t_2 lie in G . So, clearly we have that there exists x_1 .

And we use a superscript, x_1 and x_2 such that $f(x_1)$ is less than equal to t_1 , $g_i(x_1)$ is less than equal to u_i and $Ax_1 - b$ equals v_1 . And similarly $f(x_2)$ is less than equal to t_2 , $g_i(x_2)$ is less than equal to u_i and $Ax_2 - b$ is equal to v_2 , ok.

So, now remember \bar{t} is what? \bar{t} is $\alpha t_1 + (1 - \alpha) t_2$; similarly \bar{u} is $\alpha u_1 + (1 - \alpha) u_2$; \bar{v} is $\alpha v_1 + (1 - \alpha) v_2$. So, what I will do is, I need to show that there exists ok; let us call this \bar{x} , let there exist an \bar{x} such that all this, right.

So, what I will do is, take \bar{x} as $\alpha x_1 + (1 - \alpha) x_2$. In that case in that case, what would I get? Well, $f(\bar{x})$, since f is convex, I would get that $f(\bar{x})$ is less than equal to $\alpha f(x_1) + (1 - \alpha) f(x_2)$, and

that is and that itself is less than equal to $\alpha t_1 + 1 - \alpha t_2$. And that is and that is what? That is simply \bar{t} , ok.

Similarly, g_i of \bar{x} is going to be less than equal to αg_i of $x_1 + 1 - \alpha g_i$ of x_2 and that is equal to; I am sorry that is less than equal to $\alpha u_i + 1 - \alpha u_i$ which is equal, which is nothing, but $\alpha \bar{u}$ sorry, which is nothing but it is not α , which is nothing but, which is nothing but \bar{u} , right.

And $A\bar{x} - b$ is actually equal to $\alpha Ax_1 + 1 - \alpha Ax_2 - b$ and that is equal to. Now, what I will do is, I will write this, I will split this b as $\alpha b + 1 - \alpha b$ and that would give me $\alpha Ax_1 - b + 1 - \alpha Ax_2 - b$. And this is equal to as you can check, $Ax_1 - b = v_1$, $Ax_2 - b = v_2$. So, it will be $\alpha v_1 + 1 - \alpha v_2$ and that is nothing, but \bar{v} , ok.

So, in other words. So, we can, we needed to show, we needed to show that this that this particular thing is true and we; that means that required as to show that there, exists such an \bar{x} that satisfies all of these. And we were able to show that by taking \bar{x} to be $\alpha x_1 + 1 - \alpha x_2$, where x_1 and x_2 are actually the points that, the x_1 and x_2 that give you that because of which u_1, v_1, t_1 and u_2, v_2, t_2 are in G , ok.

So, what is this done? This is given us that, we by doing this what are we shown? We have shown basically this claim here, ok. So, just to make sure everyone is on track make, this is the claim that we have shown that, if f and these g 's are convex; then if these f 's and g 's are convex, then the set G is also convex, ok.

So, now let me move forward to the duality problem, ok. Now, let me define this another set ok, let me define this set T as capital T as follows.

(Refer Slide Time: 12:33)

$T = \{(0,0,s) \mid s < p^*\}$
 Claim $T \cap G = \emptyset$.
 proof. Suppose $(u,v,t) \in T \cap G$
 $\Rightarrow u=0, v=0, t < p^*$
 $\Rightarrow \exists x$ st $f(x) \leq t < p^*$
 $\begin{cases} g_i(x) \leq 0 & i=1,\dots,m \\ Ax=b=0 \end{cases}$
 $\Rightarrow \exists x$ feasible for P st $f(x) < \text{optimal value of (P)}$.
 This is a contradiction.

T is a convex set.
 G we showed is a convex set.
 $\Rightarrow \exists$ a separating hyperplane for T & G
 $\exists (\tilde{\lambda}, \tilde{\theta}, \tilde{\mu})$ st $(\tilde{\lambda}, \tilde{\theta}, \tilde{\mu}) \neq 0$
 $\& \inf_{(u,v,t) \in G} (\tilde{\lambda}, \tilde{\theta}, \tilde{\mu})^T (u,v,t) \geq \sup_{(u,v,t) \in T} (\tilde{\lambda}, \tilde{\theta}, \tilde{\mu})^T (u,v,t)$
 $\inf_{(u,v,t) \in G} \tilde{\lambda}^T u + \tilde{\theta}^T v + \tilde{\mu} t \geq \sup_{t < p^*} \tilde{\mu} t$
 If $\tilde{\lambda} < 0$ or $\tilde{\mu} < 0$, then LHS = $-\infty$, which absurd since RHS is finite.
 $\Rightarrow \tilde{\lambda}, \tilde{\mu} \geq 0$.
 $\inf_{(u,v,t) \in G} \tilde{\lambda}^T u + \tilde{\theta}^T v + \tilde{\mu} t \geq \tilde{\mu} p^*$

It is $0, 0, s$, such that s is less than p^* , ok. So, what is this? Well, this is nothing, but an open segment along the T axis; it is a $0, 0, s$ such that s is less than p^* . So, it is an, it is a segment that goes from minus infinity.

So, if you or rather let me go back to the figure that I drawn here. Well it is a segment that starts from minus infinity and goes all the way up till p^* ; but does not quite touch p^* , so it is open at p^* . So, it is this segment, ok. Now, this segment is obviously a; so this is obviously a convex set, alright.

Now, but now the main the thing that the thing that I that is important for us is that, T intersection G ok; the claim is that T intersection G is empty. Now, this may seem very natural because of the kind of figure I have drawn here, ok. No wonder it has to be, you would think that yes T intersection G is non empty; but it needs a bit of proof, see for

example, the G could very well do this right it could go down here and then intersect here again and so on so.

So, that, so we have to be; we have to be a little careful and rigorous about this particular argument. So, let us do that, ok. So, $T \cap G$ is empty. So, suppose now suppose. So, proof again, suppose u, v, t belongs to G ok; sorry belongs to suppose u, v, t belongs to T , then it means that or rather let us put it this way, let us start with suppose u, v suppose u, v, t belongs to $T \cap G$, ok.

Now, if it belongs to G , if it belongs to T ; then it means that u equals 0, v equals 0, u equals 0, v equals 0 and t is less than p^* , ok. But it since it also belongs to G , it implies that there exists an x , such that $f(x)$ is less than is less than equal to t . So, there exists an x such that $f(x)$ is less than equal to t , $g(x)$ is less than equal to 0 and $Ax - b$ is equal to 0, ok.

But then t itself ok is less than p^* from here. So, if t is less than p^* ; it means that there exists an x which is feasible. So, by this feasible for p such that the value of, such that f of x is less than the optimal value of p . So, what does this mean? It basically means that, if you if $t \cap G$ is not empty; then you should be able to find an x whose optimal value is actually less than the stated value p^* , ok. And that is absurd, because p^* is the least possible value. So, consequently this is not possible.

So, this is a contradiction, ok alright ok. So, now, so what does this mean then? This means that $T \cap G$ is empty; now T itself, because it is at it is this kind of segment that I mentioned from minus infinity, it is an open segment from minus infinity to p^* it that. So, that is that so, the set T is a convex set. We just said that, we actually just showed that G itself is also a convex set, ok. So, T is a convex set, G is also a convex set. So, T is a convex set, G we showed is a convex set, right.

So, now what we have are these two convex sets whose intersection by our claim here is empty. So, there are these so, we got these two convex sets that, that do not intersect. Which means what? Which means, so then what does that mean? That means that, we are in the

regime of now the separating hyperplane theorem. What is this saying? So, which means that, there exists a separating hyper plane, ok.

So, what does this mean? So, there exists a separating hyperplane or a hyper plane separating for T and G. So, there is a hyper plane that separates T and G; so that means there exists, let us call this, I am going to use this notation λ tilde, θ tilde, say let us call this and let us call this μ tilde say such that.

So, there exists this λ tilde, θ tilde, μ tilde such that this hyper plane remember from the separating hyper plane theorem says that, this hyper plane should have a non-zero slope. So, that means such that, this is not equal to 0, ok. And so, is this there is a hyper plane with a non-zero slope, such that now if I look at the set G and look at this at T, they lie on either side of the hyper plane, right.

So, if I look at the least value of over the of this hyper plane over the set G. So, you look at infimum over u, v, t in G of λ tilde, θ tilde, μ tilde transpose u, v, t ; this has to be greater than equal to the largest value over u, v, t of the same function λ tilde, θ tilde, μ tilde transpose u, v, t , as u, v, t ranges now over T capital T, ok. So, this is the separating hyper plane theorem.

So, what is this effectively saying that, entire set G lies on one side of it and the entire set T lies on the other side of it. And there is a weak inequality here; because they could both the infimum and the supremum can end up being equal, alright ok. So, now, let us evaluate these things carefully. So, u, v, t belongs to capital T, if means what? That it has to be that u and v are 0.

So, this particular quantity on the right hand side is actually just supremum over a point, supremum over t less than p star, ok. So, u, v, t belongs to capital T means that, u is equal to 0, v is equal to 0 and t is less than p star, right. So, it is this t less than p star of μ tilde into t , ok. Now, I have not yet, we are not yet in a position to evaluate this; so but, let us leave it at this particular stage, ok.

So, now let us look at the one on the left. And what is the, what is the one on the left? So, why are we, ok just ok yeah; why are not we in a position to evaluate this? Let us let us understand this for the moment. So, why are not we in a position to evaluate μ into t , the sorry the supremum of μ into t over for t less than p star? The reason is because we do not know the sign of μ .

See if the sign of μ is, if μ is, if μ was were positive ok; if μ were positive, you would evaluate this, quantity you would you would maximize it by putting t equal to the supremum would be at t equal to p star, alright. If μ is, if μ is negative alright, if μ is negative; then the problem becomes very different. So, let us just hold our horses here, before we analyze before we go ahead and analyze this.

So, let us first look at what is on the left, ok. What is there on the left? Well, this whole thing I can spread out and write this a little bit; let us write this out in a little more explicit form. So, let us write this as $\lambda^T u + \theta^T v + \mu t$ as u, v, t belongs to G , ok. And so, I should put this as a greater than equal to, ok.

Now, if you look at look, let us observe this particular equation here. So, on this quantity here, this u, v, t belongs to G ; which means that u can be. So, go back to the definition. If you go back to the definition, see what did I what did we say is how does G look. G is unbounded in the u direction. So, whenever there is a u in G , a larger u can also be fit in; because of the way the set has been defined right, these inequalities will continue to hold.

Similarly, if you have a t in G , then you can always fit in a larger T . So, what this means is G is always unbounded in the u and t , along the u and t directions going towards plus infinity, right. So, what this means is that, if u, v, t belongs to G , if u, v, t belongs to G here, then u and t can be made arbitrarily large ok. So, these their values can be made arbitrarily large.

So, now for if their values can be made arbitrarily large; what does this mean? If their values can be made arbitrarily large, what is the; what is the only conclusion that can be drawn from here? Well, the only conclusion that we can have is. So, if the if these. So, let us think of it

this way. So, if λ or if either λ or μ were negative, λ tilde or μ tilde were negative, ok.

So, if λ tilde or μ tilde were negative and the fact that you can drive u and t all the way to plus infinity would mean that, this the infimum here would be would actually be, the infimum here would actually be minus infinity, right. So, if λ tilde and μ tilde were negative, then the infimum here would be minus infinity; but then the infimum is lower bounded by some particular finite value here remember. So, the infimum is lower bounded by some particular finite value, which is the supremum of this linear function over T , right.

So, since this is a supremum, the supremum cannot be minus infinity right; because it is you can always take any for any real value and evaluate it to something finite. So, the fact that this is, this is lower bounded by some finite value; means that you cannot possibly have λ tilde and μ tilde as negative, either of λ tilde, μ tilde is negative.

So, if λ tilde less than 0 or μ tilde less than 0; then LHS equals minus infinity, which is absurd since RHS is finite, since RHS is finite, right. So, if this so, this can be minus so, this cannot be minus infinity, right. So, which means that, both λ tilde and μ tilde are greater than equal to 0; write this separately, λ tilde greater than equal to 0 and μ tilde also greater than equal to 0, ok.

So, now if μ tilde is greater than equal to 0, then that tells us also something what we, about what we were talking about here. Well if μ tilde is greater than equal to 0, then the supremum here is to be equal to μ tilde times p star, ok. So, that has simplified our expression a bit, let me change my color and.

So, you can keep track infimum of. So, the left hand side has now become λ trans, λ tilde transpose u plus θ delta transpose v plus μ tilde times t greater than equal to μ tilde into p star and the infimum here is over u, v and t belonging to G .