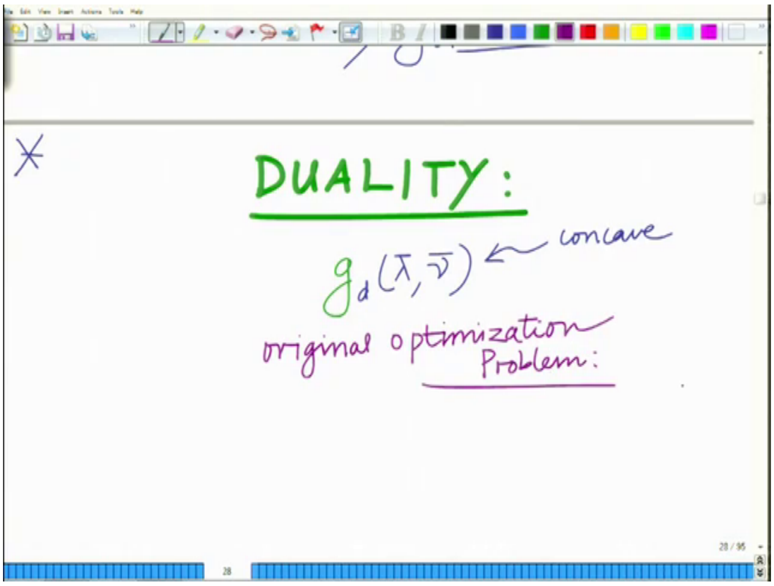
Applied Optimization for Wireless, Machine Learning, Big Data
Prof. Aditya K. Jagannatham
Department of Electrical Engineering
Indian Institute of Technology, Kanpur

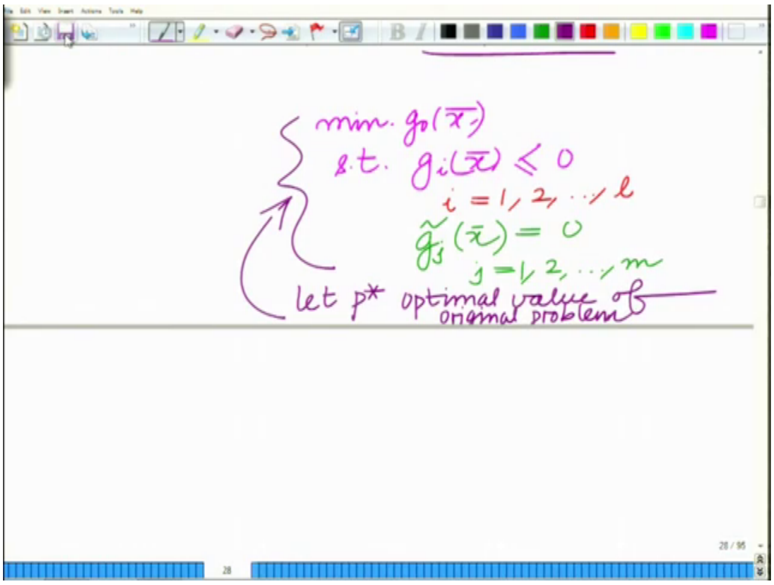
Lecture-64
Relation between optimal value of Primal and Dual problems, concepts of Duality gap and Strong Duality

(Refer Slide Time: 00:23)



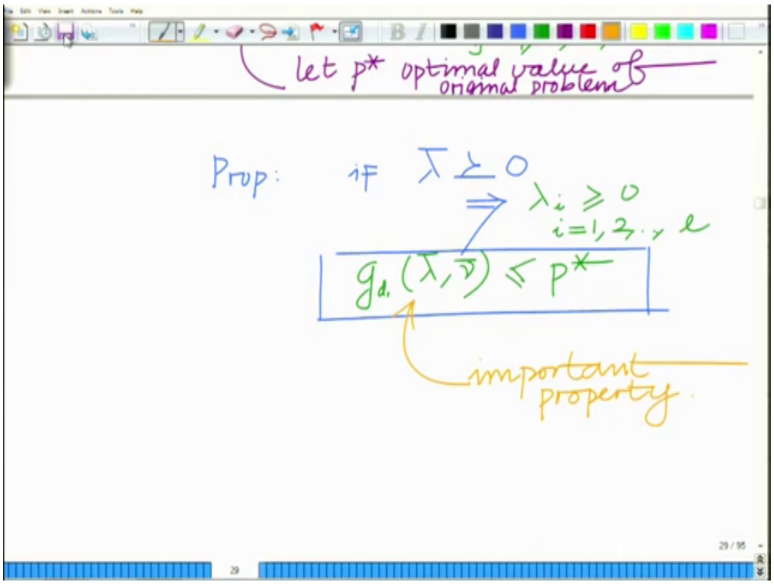
Hello. Welcome to another model in this massive open online course. And we are looking at the concept of duality for optimization alright. So, let us continue our discussion. And what we have shown is that well this dual function that is the Lagrange dual function g d lambda bar comma nu bar, we have shown that this is concave correct. And now let for a moment let us go to the original optimization problem. Let us go back the original possibly not known not necessarily convex optimization problem.

(Refer Slide Time: 01:19)



And that is as follows what we are doing is we are minimizing g naught of x bar, subject to a constraint g i x bar less than equal to 0, i equals 1, 2 up to l and the equality constraint g j tilde x bar equals 0, j equals to 1, 2 up to m. This is a original optimization problem. Now, let P star denote the optimal value of this original optimization problem that is if you perform this optimization, we are obtain P star as the optimal value of the objective function correct subject to the constraint of this original ok. So, P star a denoting by P star the optimal value of the original optimization problem that is the what we have obtained as optimal view if we solve this original, possibly non-convex optimization problem ok.

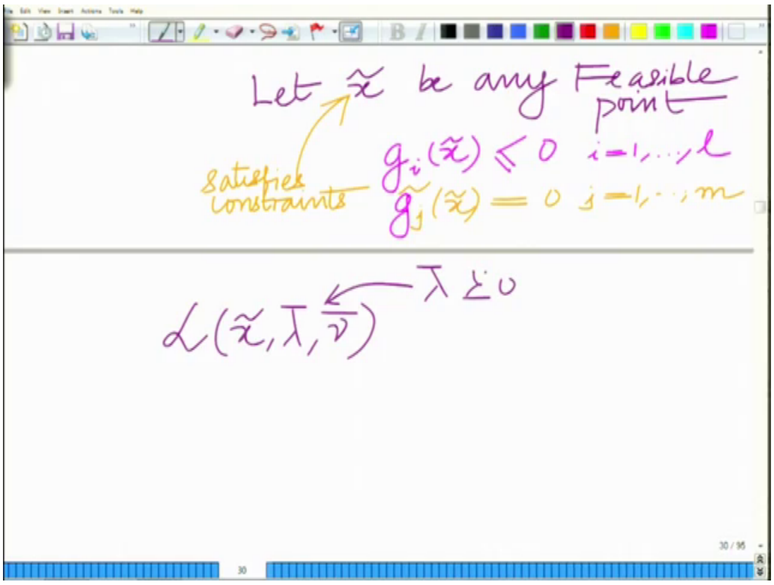
(Refer Slide Time: 02:43)



Now, the interesting result that you want to show is that and this is the very interesting, and very important property of the dual. We want to show that if this vector of Lagrange multiplier associated with the inequality constraints that is lambda bar is component wise greater than 0 that means, that each lambda i right so each lambda i is greater than equal to 0 or i equals to 1, 2 up to l. Remember lambdas are the Lagrange multiplier associates with the inequality constraints.

Now, each lambda is greater than equal to 0 all right, then the dual that is g d of lambda bar, nu bar that is less than always less than or equal to P star, which is basically the optimal value of the original optimization problem. And this is a very important property. This is a very important property of the dual function, which states at the dual function that is if the Lagrange multiplier lambda i are greater than equal to 0. Then the dual function g d lambda bar, nu bar is always less than equal to P star, which is the optimal value of the original optimization problem and this is the very interesting property.

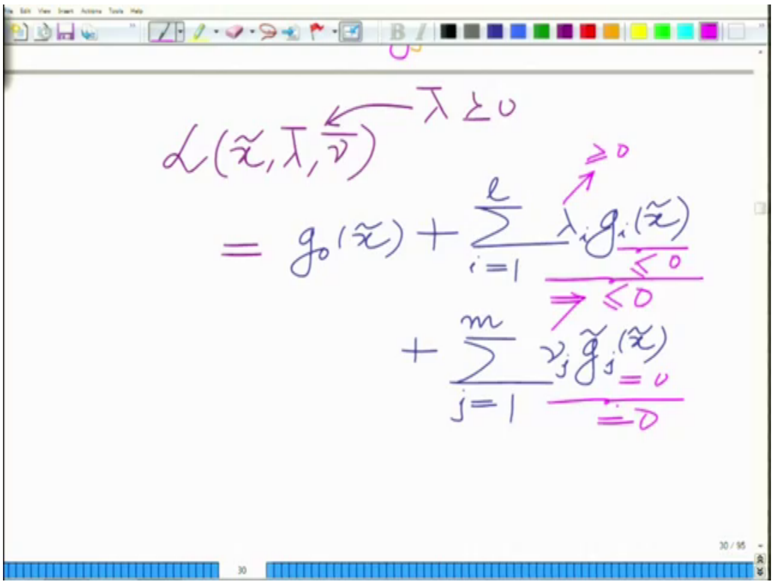
(Refer Slide Time: 04:37)



Let us try to understand, why this arise ok. Let us try to briefly demonstrate this, so this is the very important properties. So, let us try to demonstrate this, so observe now let or let us starts with the following. Let x tilde be any feasible point. Feasible point mean, it satisfies the constraint. Satisfies the constraint is sense that you have g i of x tilde is less than or equal to 0, i equals 1 up to l. And g j tilde of x tilde equals 0; j equals 1 up to m ok.

So, this is a feasibles of point in the sense that this satisfies the constraints and therefore, now if you look at the Lagrange dual if you look at the Lagrangian correct L of x tilde lambda bar, nu bar, now remember you have to keep in mind that lambda bar is component wise greater than equal to 0, which means each element lambda is greater than or equal to 0.

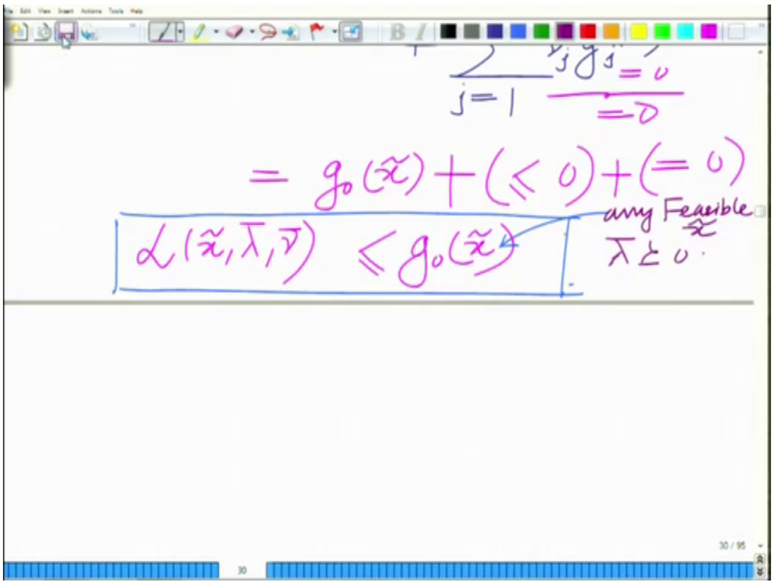
(Refer Slide Time: 06:07)



So, this is basically now let me just write it a little bigger, so that you can observe this. So, this is g naught x tilde plus summation i equals to 1 to l lambda i g i x tilde plus summation j equals 1 to m nu j g j tilde x tilde. Now, what you have observe is now g naught x tilde plus now look at these each lambda i is greater than equal to 0. X tilde is a feasible point, we have seen that each g i of x tilde is less than or equal to 0, which implies that lambda i that is we look at these net quantity lambda is greater than equal to 0, g i of x tilde is less than or equal to 0. Implies this lambda i into g i of x tilde is less than or equal to 0.

On the other hand g j of x tilde equal to 0, we do not care about nu j. g j of x tilde this is equal to 0. This implies that this whole quantity nu j into g j of x tilde this is equal to 0.

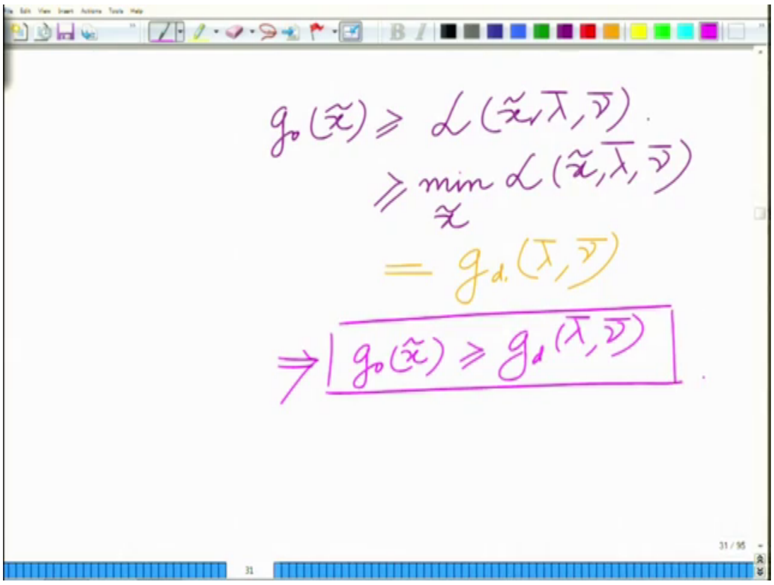
(Refer Slide Time: 07:27)



So, we have this is equal to g naught of x tilde plus some quantity that is less than or equal to 0 plus some quantity that is equal to 0. And so your g naught of x tilde plus some quantity is less than or equal to 0, which is the negative quantity, which implies that this is therefore less than or equal to g naught of x tilde that is basically your quantity on the left this is the Lagrangian L of x tilde lambda bar, nu bar this is basically. So, this original this is basically your the Lagrangian for the optimization problem, this is less than or equal to g naught of x tilde for any feasible point.

So, this holds for now remember this holds for any feasible point. And as long as all the lambda i is are greater than that is lambda bar is component wise greater than equal to 0 very good all right. So, we have shown that the Lagrangian dual function at any feasible point x tilde, when the vector each lambda is greater than equal to 0 is always less than or equal to the or is always is always less than or equal to g naught of x tilde that is the value of objective function at that point x tilde at that feasible point x tilde.

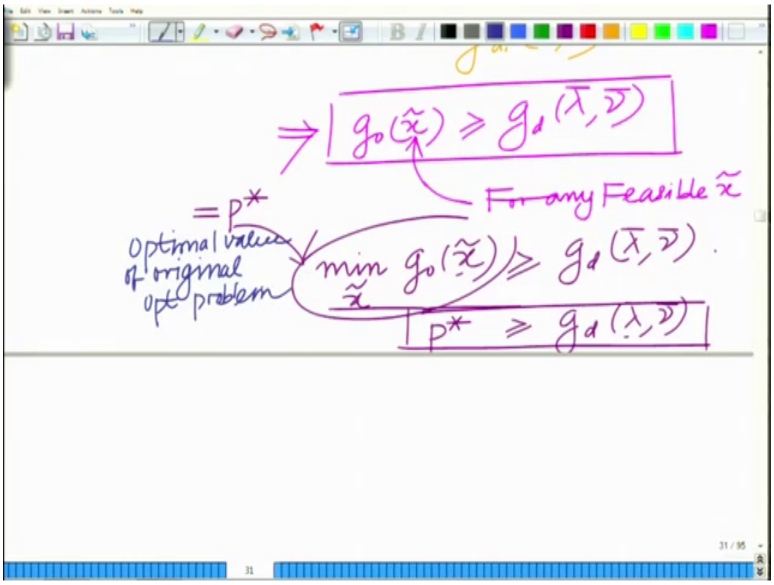
(Refer Slide Time: 09:07)



Now, now therefore what we have is g naught of just rewriting this g naught of x tilde greater than or equal to the Lagrangian ok. Now, g naught of x tilde is greater than or equal to l of I am sorry the Lagrangian at x tilde, which implies that this is greater than or equal to the minimum over all such feasible points ok. So, what you have now shown that g naught of x tilde is greater than or equal to Lagrangian at x tilde. And therefore, you take the minimum of the Lagrangian or the set of all feasible points x tilde.

Naturally g naught of x tilde is going to be greater than equal to the minimum value on the right hand side. And the minimum value of the right hand side is nothing but the Lagrangian dual function. So, therefore this is basically nothing but this is if we take the minimum that is your g d lambda bar, nu bar, so that is basically equal to your g d lambda bar, nu bar. And therefore that implies that, your g naught of x tilde is greater than or equal to g d ok. So, this is greater than or equal to correct this is greater than equal to g d lambda bar, nu bar for any feasible point x tilde for any feasible x tilde.

(Refer Slide Time: 10:45)

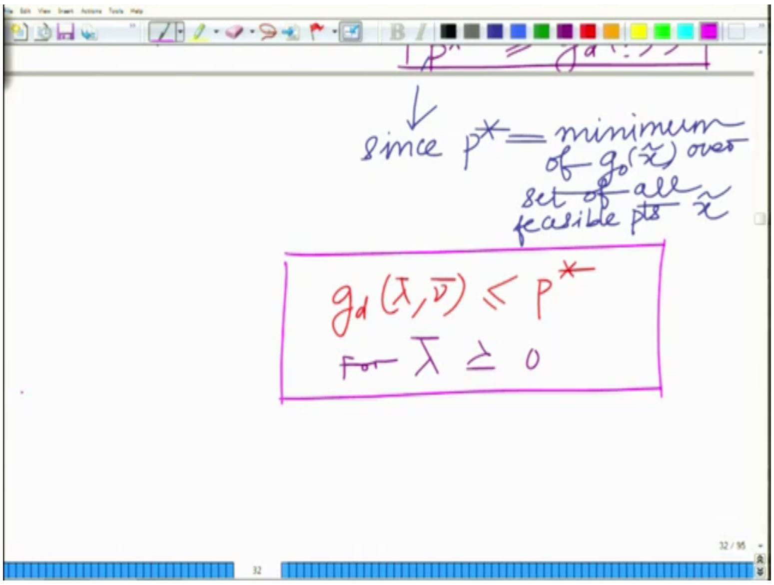


Now, naturally if you take the minimum of this, so for any feasible x tilde, this is greater than or equal to g d lambda. So, if we take the minimum of g naught x tilde over the set of all feasible points, this is going to be greater than or equal to. Now, this holds for any x tilde correct, this is an equality holds for any x tilde g naught of x tilde is greater than or equal to the dual function lambda bar g d of lambda bar, nu bar. Naturally at the optimal that is if we take the minimum over the set of all, that is this holds for any feasible point x tilde.

So, naturally it holds for that feasible point x tilde at which the minimum is achieved, correct that is the minimum overall said of all feasible point x tilde g d of x tilde g naught of x tilde that is nothing but P star, which is the optimal value of the original optimization problem. So, this is basically your P star, which is therefore greater than equal to less than.

This is basically nothing but this minimum this is equal to P star, this is the optimal value of the original optimization problem. This is the optimal value of the optimization problem ok. So, P star is greater than or equal to g d lambda bar, nu bar ok.

(Refer Slide Time: 12:43)

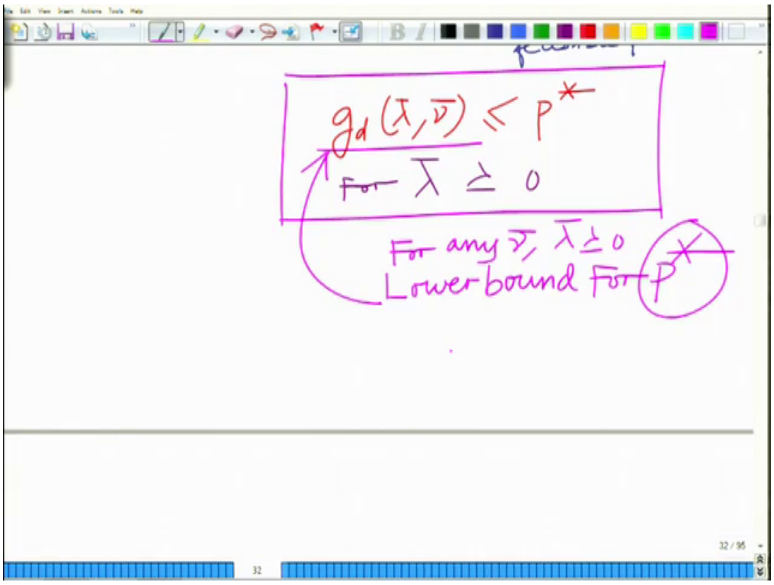


Since, P star is minimum of since basically P star is the since P star equals the minimum of the objective of g naught x tilde over set of all feasible points x tilde or the set that contains all the feasible points x tilde that is what is meant by this. And therefore, you have g d of lambda bar, nu bar is less than or equal to P star ok.

Now, mind you have to still keep in mind that this holds to only if lambda bar that is you also have that let me just try these things together, for lambda bar greater than or equal to that is component wise that is each component of the vector lambda bar greater than equal to 0, which is means that each Lagrangian multiplier lambda associated with the inequality constraints is greater than or equal to 0 ok.

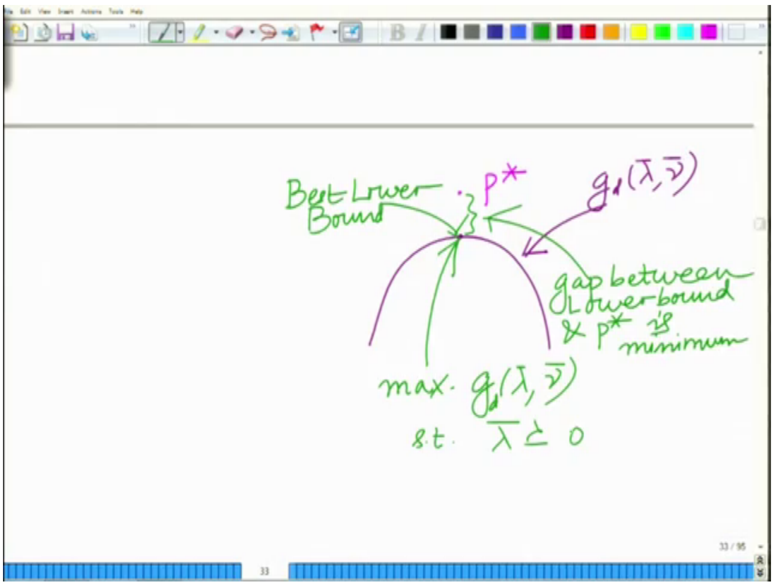
So, this basically that and therefore what this shows is g d of this Lagrange dual function for any lambda bar, nu bar. So, this holds for any lambda bar, nu bar. I mean this holds for any nu bar and any lambda bar such that each lambda bar is greater than or equal to 0, which means that this Lagrange dual function forms is a lower bound for this P star. For any lambda bar greater than or equal to 0 that is component wise greater than equal to 0 and nu bar. This Lagrange dual function is a lower bound for P star that is you can say that the optimal value of the original optimization problem is always going to be greater than equal to the Lagrange dual function.

(Refer Slide Time: 14:49)



So, this is basically a lower bound, which is basically optimal value of the original optimal; lower bound for P star, which is the optimal value of the original optimization problem ok. So, for any lambda bar greater than for and this is for any nu bar comma lambda bar component wise greater than equal to component wise now naturally if the Lagrange dual function is the lower bound, right for any lambda bar component wise greater than equal to 0 and nu bar.

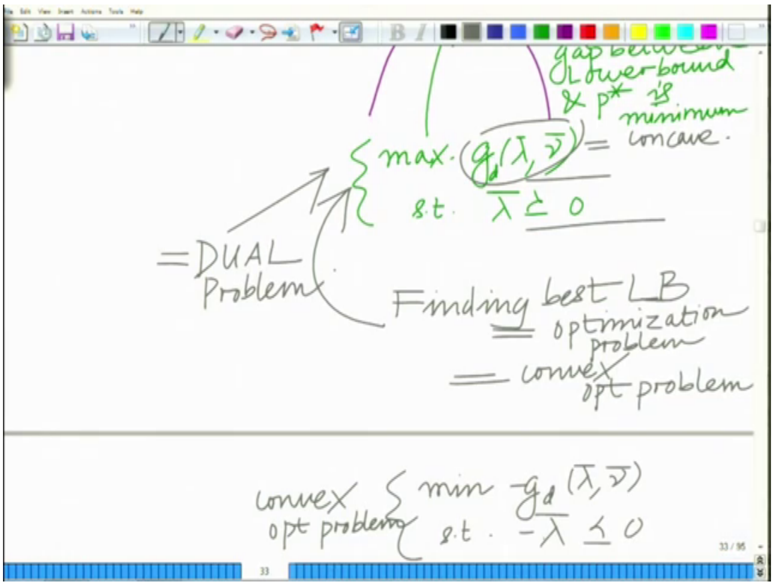
(Refer Slide Time: 15:43)



One can ask the question, what is the best lower bound, right because look at this we have this P star, we have this concave function, which is g d of lambda bar, nu bar all right. This is always lower than P star. One can ask the question, w what is the best lower bound? Best lower bound is the maximum value of this lower bound, which is you want to be as close as the optimal value P star.

Because, if you look at it technically, if you look at the lower bound of minus infinity that is always going to be lower bound right any P star is always going to be, but we want the best possible lower bound. What is the lower bound, which is as close as possible to the P star, so that this gap between the lower bound and P star is minimized so gap between lower bound and P star is minimum ok. And that is basically given as the maximum value of the Lagrange dual function g d. Of course, subject to the constraint now that lambda bar is component wise greater than equal to 0.

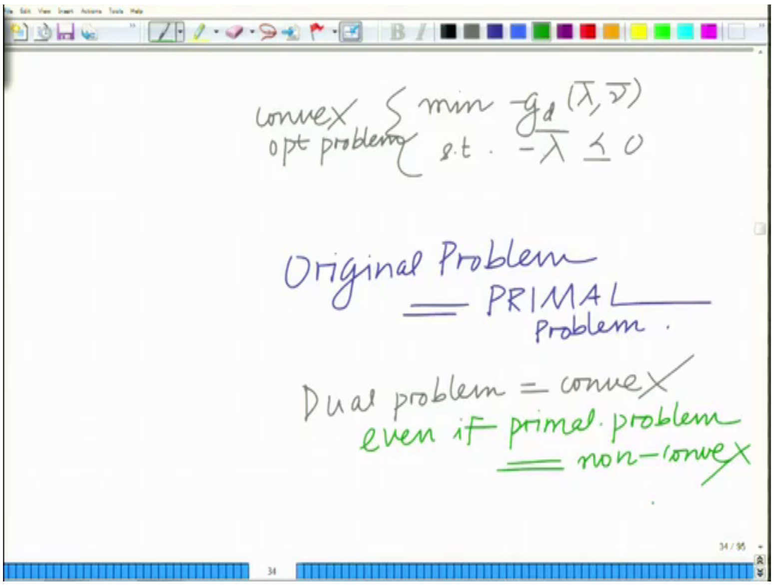
(Refer Slide Time: 17:21)



And now we can see the there is an optimization problem, this is also (Refer Time: 17:20) best lower bound is also an optimization problem. Finding best lower bound, and in fact the interesting thing about this, this is the convex optimization problem although the original problem need not be convex, because this is concave.

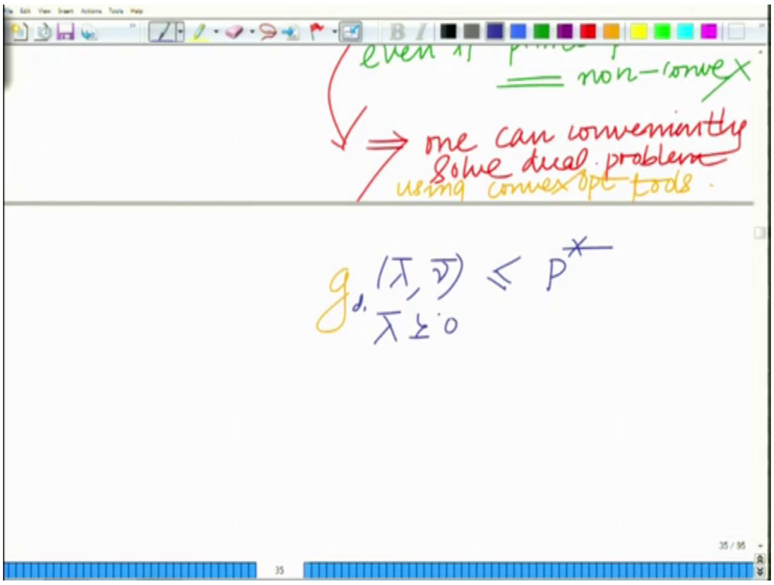
So, I can equivalently write this as minimum of g d minus g d lambda bar, so g d lambda bar, nu bar is concave minus g d lambda bar, nu bar is convex. I can so maximizing g d lambda bar, nu bar, I can minimize minus g d lambda bar nu bar. Subject to the same constraint lambda bar is greater than equal to 0. In fact, I can put a negative sign here, and I can make this reverse minus lambda bar is less than equal to 0. And this is your standard form convex. So, you can use all the techniques of convex optimization. To conveniently solve the dual problem, this is known as the dual problem, where you are optimizing the Lagrange. So, this is basically the interesting dual problem.

(Refer Slide Time: 19:03)



The original problem is termed as the primal problem, primal and dual. So, even if the primal problem is the non-convex, the dual equivalent dual problem that is derived from the primal problem is convex. And therefore, one can conveniently use all the techniques of convex optimization to solve the Lagrange dual problem. So, primal problem is the dual problem is convex.

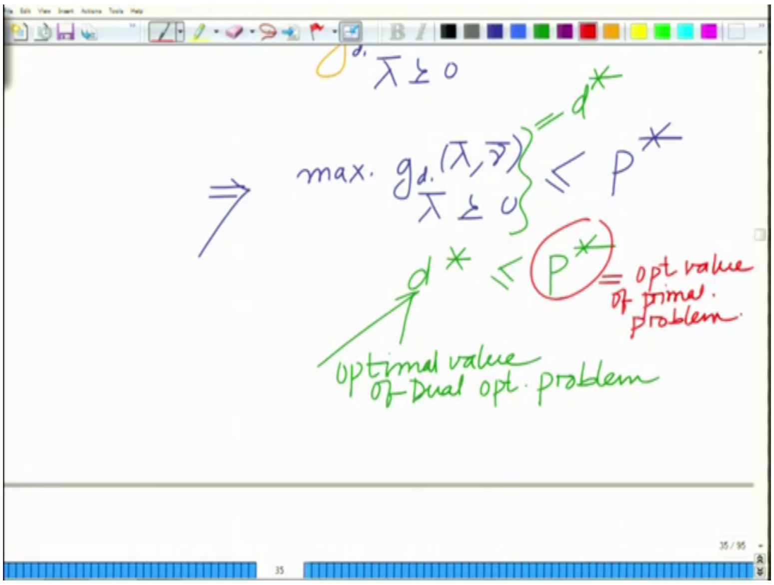
(Refer Slide Time: 20:07)



Even if primal problem is non- convex ok which implies one can conveniently solve the dual problem using tools of convex optimization, convex optimization tools or convex optimization solver, and so on. So, one can use all the tools and techniques for convex optimization available for convex optimization into conveniently solve the dual problem. And remember, because you are taking the best lower bound that is going to give you something that is as close as possible the best lower bound.

There is lower bound that is as close as possible to the optimal value P star, but still it is going to be lower than P star. So, what you get by solving the dual optimization problem is always going to be a lower bound, you are going to get the best lower bound, but still it is a lower bound, it is lower than P star ok. Now, when is it equal to P star ok, now it is always going to be because remember so this is g d lambda bar, nu bar is less than or equal to P star for all lambda bar greater than equal.

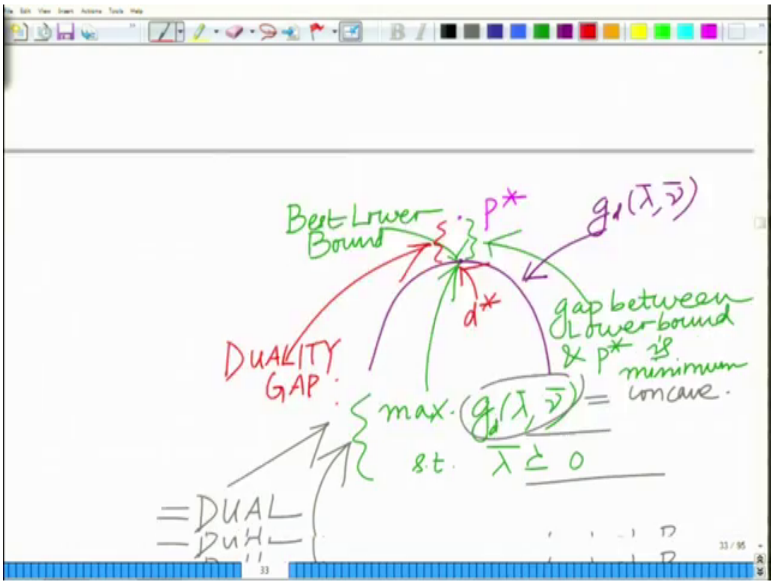
(Refer Slide Time: 21:35)



So, if we take the maximum of this, remember this holds to for any lambda bar, nu bar ok. So, at the maximum also which holds for some particular lambda bar, nu bar, this is going to be less than or equal to P star. So, implies if we take the maximum right, so for any lambda bar, nu bar less than equal to P star. So, if we take the maximum of this for some optimal value of lambda bar, nu bar, this is still going to be less than or equal to P star ok.

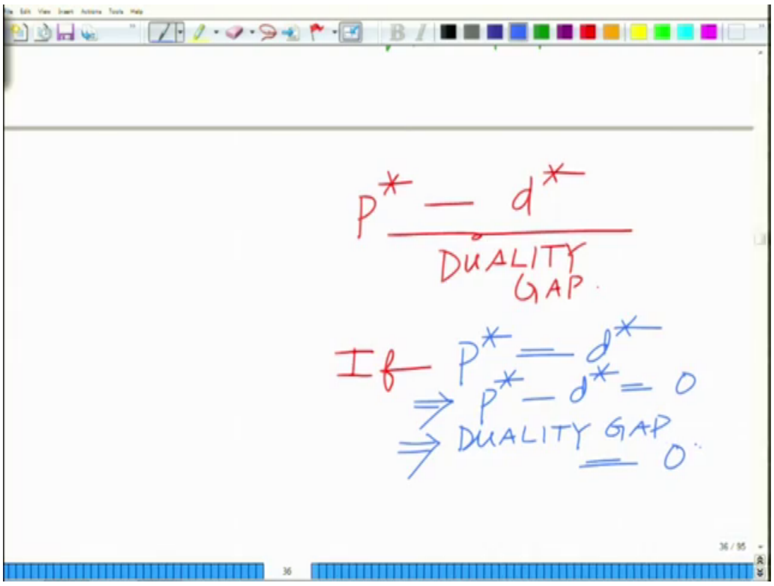
And therefore, this lower bound if you called this dual optimization problem. If the optimal value of this is d star d star, so d star is always going to be less than or equal to P star ok. So, d star is the optimal value of the d star is the optimal value of dual optimization (Refer Time: 22:43) P star equals optimal value of the primal problem. This is the optimal value of the primal problem.

(Refer Slide Time: 22:59)



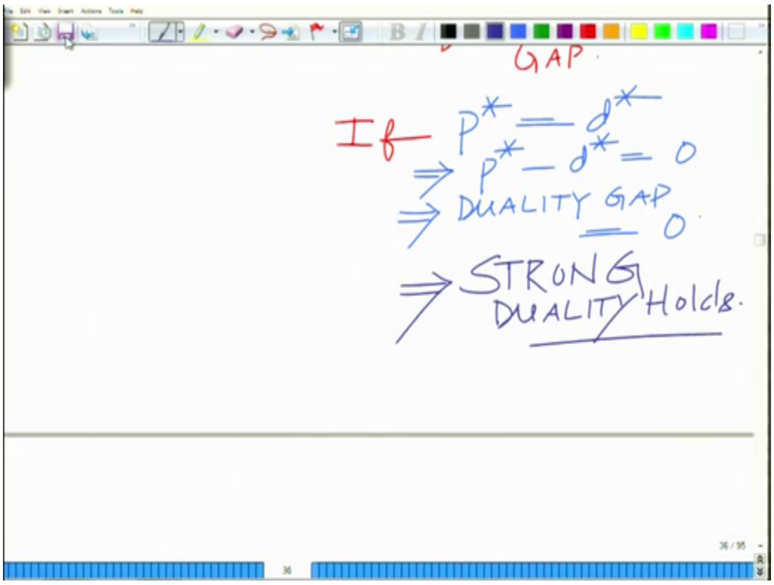
And this is basically your d star this is basically where the d star, which is the maximum value of the Lagrange dual function ok. And this gap now that you see between d star and P star that is your duality gap.

(Refer Slide Time: 23:23)



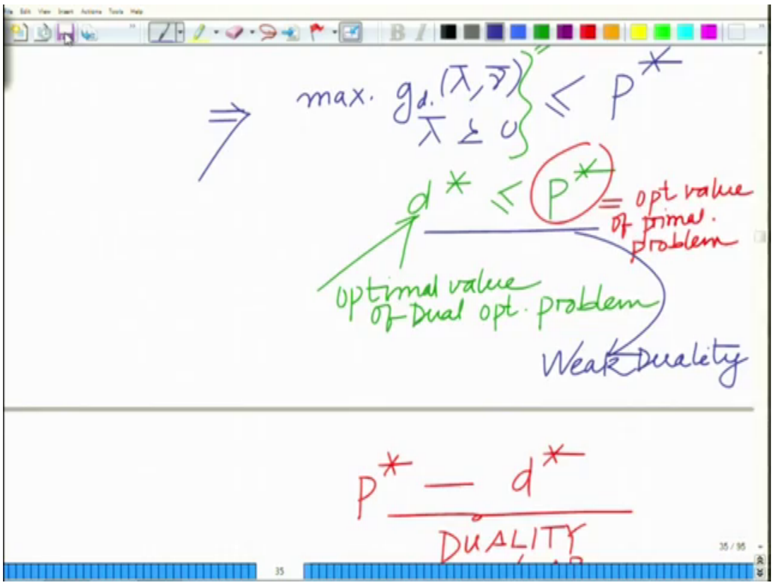
So, d star is less than equal to P star. And this gap P star minus d star this is your duality, this is termed as the duality gap. Now, if P star equals d star, that implies P star minus d star equal to 0, that implies duality gap equals 0.

(Refer Slide Time: 24:05)



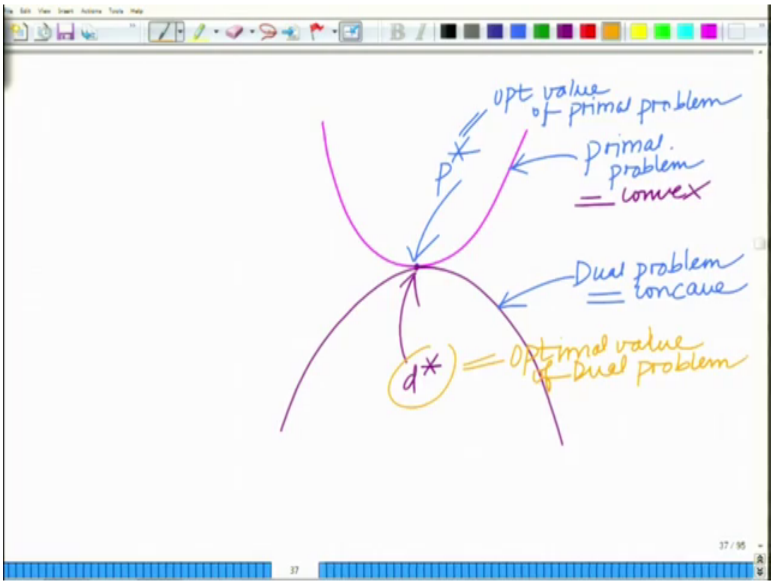
When this happens, it is said that strong duality holds, so this implies that strong duality holds for the problem.

(Refer Slide Time: 24:25)



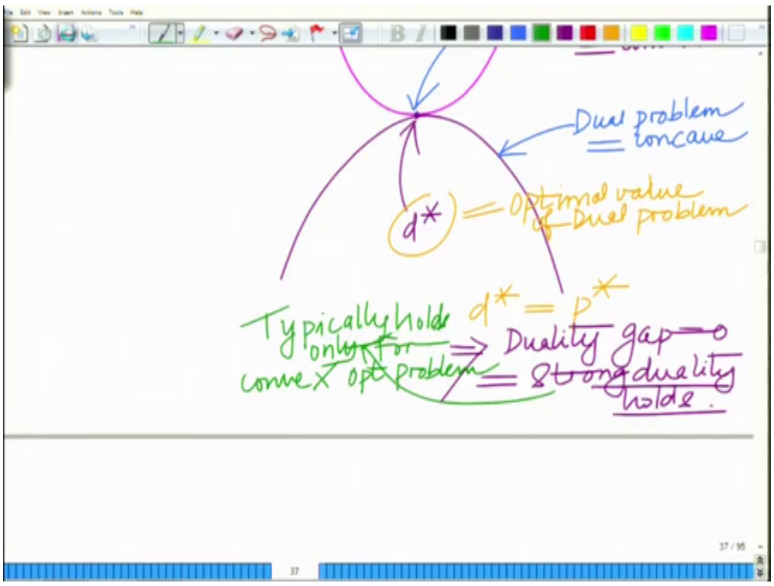
Otherwise, if d star is less than equal to P star, this is simply weak duality is simply weak duality, it is always holds.

(Refer Slide Time: 24:41)



And the special case that P star equal to d star that is we look at this the primal problem this is convex ok. So, this is your primal problem, which is convex. And this is the optimal point, which is P star, which is the optimal value of the primal problem. And then you have the dual, which is concave. And you have the optimal value of the dual problem, which is so the primal problem is convex. And you have the dual problem or you have the, which is concave ok. This is equals optimal value of primal problem; this is optimal value of dual problem ok.

(Refer Slide Time: 26:09)



And when P star equal to d star, when this d star equal to P star, this implies that basically you are duality gap is 0, implies strong duality holds, this implies that strong this implies that strong duality holds. When duality gap is 0, strong duality holds for the problem. And this is typically to loosely speaking the strong duality holds for any convex (Refer Time: 26:54) although this does not hold. Although one can form the dual optimization problem, solve the dual optimization problem for any possibly non-convex problem, but strong duality typically holds only for a convex optimization problem.

So, one can solve the primal problem, so this typically holds only for a convex optimization problem. Let us say that is the theory of duality. So, you have the primal problem, optimization problem from the Lagranagian dual function all right. View from the Lagrangian take the (Refer Time: 27:43) or minimum with respect to x, you get the dual the Lagrange dual function all right, you can take the which always a lower bound per P star, which is optimal value of the original problem. What is the best lower bound, you maximize the Lagrange dual function subject to the constraint that lambda bar is component wise greater than equal to 0 that we will give you d star, which is the optimal value of the dual optimization problem; d star is always less than equal to P star.

If d star equals P star, then we say that the duality gap is 0, strong duality holds. And this is typically true for a convex optimization problem. And often solving the dual is either easier or solving the dual e problem yields valuable insides into the original optimization problem. It is always a good exercise for the primal problem.

And even though if you are even if you are able to solve the primal problem, it is always a useful exercise to construct the dual optimization problem solve that see if this solution to these two match. And then see what insides the dual optimization problem (Refer Time: 28:45). So, the primal or dual they always go hand in hand for any optimization problem in particular for a convex optimization problem, because the duality gap is 0. So, we will stop here, and continue in the subsequent models.

Thank you very much.