

Numerical Optimization
Prof. Shirish K Shevade
Department of Computer Science and Automation
Indian Institute of Science, Bangalore

Lecture - 27
Weak and Strong Duality

Hello, in the last class we studied the KKT conditions for constraint optimization problems, and we also saw some examples and how to use the KKT first order and second order conditions to get local minimum. Now, in today's class we will look at the duality theory, so this duality theory is very important, because in many constraints problems; some constraints problems are difficult to solve directly and instead one can convert them to an equivalent problem which is easier to solve. And, under certain conditions the original problem and dual problem will have the same optimal objective functional values.

(Refer Slide Time: 01:29)

Two-player zero-sum game

A Game between two players P and D

- Game setting
 - \mathcal{X} : A set of strategies for P
 - \mathcal{Y} : A set of strategies for D
 - Payoff function, $\psi(x, y)$, $x \in \mathcal{X}$, $y \in \mathcal{Y}$
- Example
 - Let $\mathcal{X} = \{1, 2\}$, $\mathcal{Y} = \{1, 2\}$
 - Payoff $\psi(x, y) = a_{x,y}$ where $A = \begin{pmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{pmatrix}$
- Game Rules :
 - P chooses a strategy $x \in \mathcal{X}$ and D chooses a strategy $y \in \mathcal{Y}$ independently
 - The referee reveals both the strategies simultaneously

NPTEL Shirish Shevade Numerical Optimization

So, in today's class we will look at the duality theory. Now, before we get into the details of duality theory, let us see some example of two players zero sum game. So, let me first explain what this game is, so this is a game between two players, so let us call these two players as P and D. So, these are the two players which are involved in the game and the game setting is like this, that very player has a set of strategies and each player chooses one of the strategies from the respective sets. So, let script X denotes the set of strategies for the player P; and script Y denotes the set of strategies for player D.

Now, there is a payoff function associated with this game, and let us call that function as $\psi(x, y)$, where x belongs to the set script X and y belongs to the set script Y .

So, each player chooses a strategy from the respective sets of strategies for that player and the payoff function is defined based on which strategies the two players have chosen. So, for example, let us assume that the set of strategies for the player P is 1 and 2. So, the player has two strategies, and the player D also has two strategies, and we will denote them by 1 and 2. And payoff function then can be written in the form of matrix. So, whose entry x comma y denotes the payoff corresponding to the strategies x and y . So, for example, if a player P chooses a strategy 2 and the player D chooses a strategy 1, then the payoff associated with that game is a 2 comma 1. So, this entry of the matrix will be the payoff corresponding to two one combination. Now, this game has some rules; so the first rule is that every player uses the strategy from the respective sets independently. So, without knowing what the other; without knowing the strategy of the other player one the player uses the strategy.

And, what they do is that they mention these strategies to the referee and referee reveals both the strategies independently, the referee reveals both the strategies simultaneously. So, this is the game that given a payoff function ψ each player chooses one of the strategies, and mentions them to the referee and the referee reveals both the strategies simultaneously and then the payoff will decide which player wins or which player loses. So, since there are two players in this zero sum game everytime one player will be the winner and the other will be the loser. So, the game output come it depends on the payoff function ψ .

(Refer Slide Time: 05:22)

The slide is titled "Two-player zero-sum game" in red text. Below the title, it says "A Game between two players P and D ". There are three bullet points: "Game Outcome:", followed by two mathematical implications: $\psi(x, y) > 0 \Rightarrow P$ pays an amount $\psi(x, y)$ to D , and $\psi(x, y) < 0 \Rightarrow D$ pays an amount $-\psi(x, y)$ to P . The third bullet point states: " P wishes to minimize payoff to D , while D wishes to receive maximum payoff from P ". The final bullet point says: "Assume that minimum and maximum exist". In the bottom right corner, there is a small video inset of a man in a light blue shirt, and a logo for NPTEL in the bottom left corner.

Now, let us assume that if $\psi(x, y)$ is greater than 0, then P pays an amount $\psi(x, y)$ to the player D , and if that quantity is less than 0, then D pays the negative of that quantity to the player P , so this is the game outcome. So, naturally every player will try to maximize his or her game. So, P would like to minimize the payoff to the player D and at the same time D would like to choose the strategy such that their payoff that D receives is a maximum payoff. So, this is an important part of the game that one player wants to minimize the payoff to the other player, while the other player D wishes to receive the maximum payoff from P . So, let us see some example to illustrate this point, so throughout this discussion we assume that the minimum and maximum exist. So, if they do not exist then one has to use the (\inf) and (\sup) appropriately, but throughout the discussion we assume that they exist.

(Refer Slide Time: 06:57)

Example: Game 1

$\mathcal{X} = \{1, 2\}$, $\mathcal{Y} = \{1, 2\}$, $\psi(x, y) = a_{x,y}$, where

$$A = \begin{pmatrix} -2 & 1 \\ 2 & -3 \end{pmatrix}$$

Player P's strategy

$$\min_y \{ \max_x a_{1,y}, \max_x a_{2,y} \}$$
$$= \min\{1, 2\}$$
$$= 1$$

Choose $x = 1$

Player D's strategy

$$\max_x \{ \min_y a_{x,1}, \min_y a_{x,2} \}$$
$$= \max\{-2, -3\}$$
$$= -2$$

Choose $y = 1$

min-max \geq max-min

MPTEL Shreshth Sharma Numerical Optimization

So, let us see an example again so as you see here that first player P has these two strategies, and the second player has the two strategies, and the payoff function is like this or it is denoted in the form of this matrix. What it means is that if the player P chooses the strategy 1 and the player 2 also chooses the strategy 1, then the payoff is minus 2. So, which means that the player 2 or player D has to pay an amount equivalent to 2 units to the player P. On the other hand, if player P chooses 2; and the player D choose 1; then player P has to pay an amount of 2 units to player D. So, let us explore the possibilities of different strategies that player P can use so, that it has to pay minimum amount to the player D.

So, suppose player P uses the strategy 1 then what happens is that the player D can choose strategy 2 or the max of these 2 quantities in this row so that player D game is maximized. On the other hand, if player P chooses the strategy 2, then the player 2 will find out what is the maximum of 2 and minus 3, and that maximum is 2 so, player 2 will choose a strategy 1. So, the player P what it will try to do is that, it will try to find out what is the maximum quantity in the first row and the maximum quantity in the second row. So, max of a (1,y) overall y. So, basically max of a (1,1) and a (1,2) and that is 1 and max of a (2,y) is the maximum of the quantities in the second row. So, max of a (1, y) is 1 and max of a (2, y) over all y is 2.

So, the two quantities are 1 and 2 and since the player P wishes to minimize the payoff to player D the player P will choose the strategy, corresponding to the min of these 2 values and mean of these 2 is 1 and that 1 appears somewhere here. So, the player P would like to choose the strategy 1, corresponding to the first row so that it is payoff to the player D is minimized so the player P chooses x equal to 1 at strategy 1 for this particular payoff function. Now, player D will also use this matrix A to decide a strategy and what player D would like to do is that, it would like to maximize the amount that player D receives from player P.

So, player D again has 2 strategies, 1 and 2 corresponding to the first two columns. Let us see what the player D does so let us look at the player D strategies now obviously player D wants to maximize its game or the payoff from this game. So, suppose player D chooses this strategy the first strategy then the player one would like to choose minus the first strategy, so that player 1 or player P receives the payoff of 2 units. So, what player D will try to do is that in the first column it will find out what is the minimum.

So, the minimum in the first column is minus 2, what is the minimum in the second column? Minimum in the second column is minus 3. So, if player P so if player D chooses the strategy 1, player P will get a payoff of 2 units and on the other hand if player D chooses the strategy 2, which corresponds to the second column, then player P will receive the amount of 3 units. So, among these two, two units is smaller than three units. So, player D would like to choose the strategy 1, so between these 2 columns. So, the minimum in the first column is minus 2, and the minimum in the second column is minus 3, and the maximum among these 2 is minus 2. So, if player D chooses the strategy 1 then the amount that that the player D will have to pay to P is only 2 units. So, it is better for D player D to choose the strategy which is y equal to 1. Now, if you look at this strategies used by players P and D, player P uses the min max strategy and the player D uses the max, min strategy. So, in some sense these two problems are duals of each other.

And in this case, the min max strategy used by the player 1, this quantity is greater than or equal to the max min strategy which is altered in the payoff of minus 2. So, [the amount], the payoff achieved using the min, max strategy is always greater than or equal to the payoff achieved by the max min strategy. Now, this game depends on the pay off matrix A now, if you change the entries here, the payoff received by each player would

be different. Now in this matrix suppose we change this minus 3 to the quantity plus plus 3. And suppose we have a new game, which is the game two so only one quantity in the previous matrix which is changed minus 3 to 3 here, the rest of conditions remain the same.

(Refer Slide Time: 14:52)

Example: Game 2
 $\mathcal{X} = \{1, 2\}$, $\mathcal{Y} = \{1, 2\}$, $\psi(x, y) = a_{x,y}$, where

$$A = \begin{pmatrix} -2 & 1 \\ 2 & 3 \end{pmatrix}$$

Player P's strategy

$$\min_y \{ \max_x a_{1,y}, \max_x a_{2,y} \}$$

$$= \min\{1, 3\}$$

$$= 1$$

Choose $x = 1$

Player D's strategy

$$\max_x \{ \min_y a_{x,1}, \min_y a_{x,2} \}$$

$$= \max\{-2, 1\}$$

$$= 1$$

Choose $y = 2$

min-max = max-min

MPTEL Shriah Shekale Numerical Optimization

Now, let us look at the player P strategies which are min max strategy and maximum among the 2 rows so this is 1 and this is 3 and the minimum among them which is 1. So, player P uses the strategy 1 and if you look at the player 2 strategy, player 2 will try to find out the minimum of minus 2 and 2, which is minus 2 and minimum of 1 and 3 which is 1 and the maximum among them so that will be obviously 1. So, player chooses the strategy 1 and if you look at the player P D strategy it will be max min strategy and as I said earlier so max of minus 2 and 1 which is 1. So, player P chooses strategy 1 and player D chooses strategy 2 and they both correspond to this 1.

So, that means that in this game the player P which choose a strategy x equal to 1 and player which choose a strategy y equal to 2 the payoff is 1, that means the player P will have to pay the amount one unit to the player D. But interestingly, this game is different from the previous one in the sense that the payoff that player P has to play and the amount the player D receives they both are same. So, in other words so this is different from the game that we saw earlier where, the min, max and the max, min values were different. So, this is a game where the min, max and max, min values are same and it is

said that the game is in equilibrium that means, that no player can gain by moving away from this strategies. And, this is the situation where the min, max equal to max, min so we have these 2 dual problems min, max and max, min and in the first game we saw that max, min was less than or equal to min max and in this game we saw that min max equal to max min.

(Refer Slide Time: 17:46)

Primal Problem

$$\min_{x \in \mathcal{X}} \underbrace{\max_{y \in \mathcal{Y}} \psi(x, y)}_{\text{primal function}}$$

Dual Problem

$$\max_{y \in \mathcal{Y}} \underbrace{\min_{x \in \mathcal{X}} \psi(x, y)}_{\text{dual function}}$$

- The two problems are *dual* to each other
- For any $x \in \mathcal{X}$ and $y \in \mathcal{Y}$

$$\min_{x \in \mathcal{X}} \psi(x, y) \leq \psi(x, y) \leq \max_{y \in \mathcal{Y}} \psi(x, y)$$

$$\therefore \min_{x \in \mathcal{X}} \psi(x, y) \leq \max_{y \in \mathcal{Y}} \psi(x, y)$$

$$\therefore \max_{y \in \mathcal{Y}} \min_{x \in \mathcal{X}} \psi(x, y) \leq \min_{x \in \mathcal{X}} \max_{y \in \mathcal{Y}} \psi(x, y)$$

Weak Duality

$$\max_{y \in \mathcal{Y}} \min_{x \in \mathcal{X}} \psi(x, y) \leq \min_{x \in \mathcal{X}} \max_{y \in \mathcal{Y}} \psi(x, y)$$

Shrihari Shekhar - Nonlinear Optimization

So, these two problems are the dual problems to each other, the problems solved by the player P so we are going to call it as a primal problem and the problems solved by the player D we are going to call it as dual problem. So, the primal problem is to minimize some function and that the function primal that function is called the primal function and that primal function is obtained by solving another optimization problem which is max of psi (x, y) where y comes from the set of strategies for the player D. If you look at the dual problem, dual problem is a max min problem so the dual problem is to maximize the dual function which is another optimization problem to minimize psi(x, y) where x belongs to X. And, we also saw in the 0 sum 2 player game, that max min is less than or equal to min max or in other words the value of the dual function is always less than or equal to the value of the primal function.

Now, what is important is that we need to be given this payoff function which is psi(x, y) and that strategies for the respective players which are script X and script Y. Now, we will be interested in finding out when the two problems have the same objective function

value. So, in that case it will be beneficial to solve the problem which is easier among the 2 so that, we do not lose on the objective function values. So, these two problems are dual to each other and for any x in the script X , and y in the script Y , we know that minimum of $\psi(x, y)$ is always less than or equal to $\psi(x, y)$ we assume that this x and y are coming from the respective sets script X and script Y .

Similarly, this $\psi(x, y)$ will be less than or equal to the max of $\psi(x, y)$ over all possible y 's. So, therefore we can write that minimum of $\psi(x, y)$ is less than or equal to maximum of $\psi(x, y)$, because of this previous inequality and since minimum of this quantity is less than or equal to max of this quantity, then we can write that, so this is happening for all possible y so even if we take the max of this quantity that will be less than or equal to the min of this quantity. So, max over all possible y 's is min $\psi(x, y)$ will be less than or equal to min x over max of $\psi(x, y)$. And this is important concept in duality which is called the weak duality that max min value is always less than or equal to min max value for corresponding to this payoff function $\psi(x, y)$.

(Refer Slide Time: 21:46)

Weak Duality

$$\max_{y \in \mathcal{Y}} \min_{x \in \mathcal{X}} \psi(x, y) \leq \min_{x \in \mathcal{X}} \max_{y \in \mathcal{Y}} \psi(x, y)$$

- When does the equality hold?

Definition

Let $x^* \in \mathcal{X}$ and $y^* \in \mathcal{Y}$. A point (x^*, y^*) is a **saddle point** for $\psi(x, y)$ if

$$\psi(x^*, y) \leq \psi(x^*, y^*) \leq \psi(x, y^*) \quad \forall x \in \mathcal{X}, y \in \mathcal{Y}$$

- $x^* = \operatorname{argmin}_{x \in \mathcal{X}} \psi(x, y^*)$
- $y^* = \operatorname{argmax}_{y \in \mathcal{Y}} \psi(x^*, y)$

NPTTEL

So, we have seen that this inequality holds for any game. Now, under what conditions this is equality and that is what we are going to study now and for that purpose we need the definitions of saddle points. So, let x star be from the set script X , and y star be the from the set script Y the point x star, y star is a saddle point, for $\psi(x, y)$ if this conditions holds. So, $\psi(x$ star, y star) is equal to $\psi(x$ star, y star) and greater than or

equal to $\psi(x^*, y)$ for all x in the set \mathcal{X} and set is \mathcal{Y} . So, this is very important condition and this is called the saddle point condition. So, as you would see here that if you consider the second part of the condition $\psi(x^*, y^*)$ is less than or equal to $\psi(x^*, y)$ so if we keep y^* fixed then the function has a minimum at x^* y^* in the direction or along the coordinate directions x .

On the other hand, if we keep x^* fixed and vary y , then the function has a maximum at x^* y^* as far as the y coordinate direction is concerned. So, this is a concept of saddle point where along one direction the function value increases and along the other direction the function value decreases and this relationship holds for all x in the set \mathcal{X} and y is the set \mathcal{Y} . Now, as you can see here that x^* is obtained by a minimizing $\psi(x, y^*)$ or fixing y^* and minimizing over minimizing the payoff function over set \mathcal{X} and similarly, y^* is obtained by maximizing the payoff function keeping x^* fixed and changing y . Now, in terms of that under this saddle point condition the equality holds in this case and that is called the strong duality.

(Refer Slide Time: 24:51)

Theorem
The following equality holds

$$\max_{y \in \mathcal{Y}} \min_{x \in \mathcal{X}} \psi(x, y) = \min_{x \in \mathcal{X}} \max_{y \in \mathcal{Y}} \psi(x, y)$$

if and only if there exists a saddle point, (x^*, y^*) , for $\psi(x, y)$.

Proof.
(a) Let (x^*, y^*) be a saddle point for $\psi(x, y)$.


$$\therefore \psi(x^*, y) \leq \psi(x^*, y^*) \leq \psi(x, y^*) \quad \forall x \in \mathcal{X}, y \in \mathcal{Y}$$

$$\therefore \max_{y \in \mathcal{Y}} \psi(x^*, y) \leq \psi(x^*, y^*) \leq \min_{x \in \mathcal{X}} \psi(x, y^*)$$

Note that

$$\min_{x \in \mathcal{X}} \max_{y \in \mathcal{Y}} \psi(x, y) \leq \max_{y \in \mathcal{Y}} \psi(x^*, y)$$

$$\min_{x \in \mathcal{X}} \psi(x, y^*) \leq \max_{y \in \mathcal{Y}} \min_{x \in \mathcal{X}} \psi(x, y^*)$$

 Shreshth Shekhar - Numerical Optimization

And, we will see that now, that the max min of $\psi(x, y)$ is equal to min, max of $\psi(x, y)$ if and only if there exists a saddle point (x^*, y^*) or $\psi(x, y)$. So, under the saddle point conditions the two problems have the same optimal objective function values. Now, let us look the proof of this theorem so this proof has two parts the first is that, we assume that it is saddle point we assume that there exists a saddle point, (x^*, y^*)

and show that this condition holds and in the second part we show that if this holds then there exists a saddle point, (x^*, y^*) , note that this (x, y) always come from the set the respective strategy sets.

Now, let us assume that (x^*, y^*) is a saddle point for the payoff function $\psi(x, y)$ so by the saddle point condition definition we have $\psi(x^*, y) \leq \psi(x^*, y^*)$ and that is less than or equal to $\psi(x, y^*)$ for all x in \mathcal{X} , and y in \mathcal{Y} . And therefore, we can write that $\max_x \psi(x, y^*) \leq \psi(x^*, y^*)$ since, this condition holds it will hold for $\max_x \psi(x, y^*)$ also, $\max_x \psi(x, y^*)$ is less than or equal to $\psi(x^*, y^*)$ and that will be less than or equal to $\min_y \psi(x^*, y)$ over all possible y . Now let us try to get the bounds on this quantities, now if you look at the first quantity so $\max_x \psi(x, y^*)$ is always lower bounded by $\min_y \max_x \psi(x, y)$, and $\min_y \psi(x^*, y)$ is always upper bounded by $\max_x \min_y \psi(x, y)$.

(Refer Slide Time: 27:15)

Proof.(continued)
Therefore,
$$\min_{x \in \mathcal{X}} \max_{y \in \mathcal{Y}} \psi(x, y) \leq \psi(x^*, y^*) \leq \max_{y \in \mathcal{Y}} \min_{x \in \mathcal{X}} \psi(x, y)$$

NPTEL
Shrinath Shrivastava Numerical Optimization

Now, we make use of these two conditions along with this inequality to write that, $\min_y \max_x \psi(x, y)$ is less than or equal to $\psi(x^*, y^*)$ less than or equal to $\max_x \min_y \psi(x, y)$. Remember that, we got this condition assuming that there exists a saddle point; that is why this condition holds and then based on the other conditions we were able to write this condition.

So, what this condition means is that $\min_y \max_x \psi(x, y)$ is less than or equal to $\max_x \min_y \psi(x, y)$, but we have already seen that for any game and feasible strategies x and y $\max_x \min_y \psi(x, y)$ is less than

or equal to min max. So if you combine these two conditions then for this game we can say that $\min \max \psi(x, y)$ equal to $\max \min \psi(x, y)$ and that is equal to $\psi(x^*, y^*)$ so this proves the first part of the theorem.

(Refer Slide Time: 28:19)

Proof. (continued)

(b) Suppose the following equality holds for some $x^* \in \mathcal{X}, y^* \in \mathcal{Y}$,

$$\max_{y \in \mathcal{Y}} \min_{x \in \mathcal{X}} \psi(x, y) = \psi(x^*, y^*) = \min_{x \in \mathcal{X}} \max_{y \in \mathcal{Y}} \psi(x, y)$$

Now,

$$\max_{y \in \mathcal{Y}} \psi(x^*, y) = \psi(x^*, y^*) = \min_{x \in \mathcal{X}} \psi(x, y^*)$$

$$\therefore \psi(x^*, y) \leq \max_{y \in \mathcal{Y}} \psi(x^*, y) = \psi(x^*, y^*) = \min_{x \in \mathcal{X}} \psi(x, y^*) \leq \psi(x, y^*)$$

Therefore, (x^*, y^*) is a saddle point for $\psi(x, y)$. \square

NPTEL
Shrihari Shekhar - Numerical Optimization

Now, the second part we assume that the equality holds and then we want to prove that x^* y^* is a saddle point so which means that it should satisfy the saddle point conditions. So, if you look at the $\psi(x^*, y^*)$, $\psi(x^*, y^*)$ can be written as $\max \psi(x^*, y)$ over all possible \min of $\psi(x, y^*)$ over x . And therefore, so we can say that $\psi(x^*, y^*)$ which is nothing but \max of $\psi(x^*, y)$ and \max of $\psi(x^*, y)$ will always be greater than or equal to $\psi(x^*, y)$ because we already obtained the maximum of $\psi(x^*, y)$ so naturally for any feasible y $\psi(x^*, y)$ will be less than or equal to $\psi(x^*, y)$ and similarly, since this is equal to \min of $\psi(x, y^*)$ that will be less than or equal to $\psi(x, y^*)$. So, \min of $\psi(x, y^*)$ is less than or equal of $\psi(x, y^*)$.

And if you consider this that $\psi(x^*, y)$ less than or equal $\psi(x^*, y^*)$ which is less than or equal to $\psi(x, y^*)$ then we clearly say that (x^*, y^*) is a saddle point for (x, y) . So, this shows that if only if there exists saddle point, then $\min \max$ is equal to $\max \min$.

(Refer Slide Time: 30:17)


Strong Duality

$$\max_{y \in \mathcal{Y}} \min_{x \in \mathcal{X}} \psi(x, y) = \min_{x \in \mathcal{X}} \max_{y \in \mathcal{Y}} \psi(x, y)$$

Consider the problem (NLP):

$$\begin{aligned} \min \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & h_j(\mathbf{x}) \leq 0, \quad j = 1, \dots, l \\ & e_i(\mathbf{x}) = 0, \quad i = 1, \dots, m \end{aligned}$$

- Can we define a game with a payoff function $\psi(\cdot)$ so that the solution to NLP is a solution to the *primal* problem, $\min_x \max_y \psi(x, y)$?
- What is the saddle point condition in terms of f , h_j 's and e_i 's?

 Shreshth Shrivastava Numerical Optimization

So, this condition is called the strong duality condition, that max min equal to min max. So, which means that in the zero sum game that we saw earlier under this strong duality condition the game is said to be equilibrium so that means, no player has any advantage of changing the strategy and getting a better payoff. Now, to define the notion of duality for optimization problem, we need to define a payoff function. So let us consider a optimization problem which we will call it as a non-linear programming problem minimize effects subjects to the condition that less than or equal to zero and $e_i(x)$ equal to zero, there are m equality constraints, and l inequality constraints. Now, the question is that can we define a game with some payoff function ψ so that the solution to NLP is a solution to the primal problem $\min_x \max_y \psi(x, y)$?

So, in other words we are interested in writing this as a min max problem by defining appropriate payoff function $\psi(x, y)$. Now, once we write this as a function the primal problem $\min_x \max_y \psi(x, y)$ then we can possibly write the corresponding dual problem as $\max_y \min_x \psi(x, y)$ and that dual problem maybe sometimes easier to solve. Then, the original problem and if that is the case and if the solutions to the primal and dual problems are same, then that dual problem and if the dual problem is much easier to solve then it will be beneficial to solve the problem other than the primal problem.

And, at the same time we do not get the objective function value which is inferior than the actual optimal objective function value if both the problems have the same solution.

So, one question is that can we define the payoff function so that we can write this problem as the min max problem and what is meant by the strong duality in terms of the functions f , h and e_i 's? Or in other words how do we get the saddle point conditions in terms of the function f h and e_i 's are in terms of the payoff function, and that is what we are going to study now.

(Refer Slide Time: 33:15)

Consider the problem(P):

$$\begin{aligned} \min \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & h_j(\mathbf{x}) \leq 0, \quad j = 1, \dots, l \\ & \mathbf{x} \in X \end{aligned}$$

Define a payoff function as the Lagrangian,

$$\mathcal{L}(\mathbf{x}, \boldsymbol{\lambda}) = f(\mathbf{x}) + \sum_{j=1}^l \lambda_j h_j(\mathbf{x})$$

where $\mathbf{x} \in X$ and $\lambda_j \geq 0, j = 1, \dots, l$

- \mathbf{x} : Primal Variables, $\boldsymbol{\lambda}$: Dual Variables
- $\mathcal{X} = X, \mathcal{Y} = \{\boldsymbol{\lambda} \in \mathbb{R}^l : \lambda_j \geq 0, j = 1, \dots, l\}$

Duality : Define a **min max** problem *equivalent* to the **primal problem P**. Then, the corresponding dual **max min** problem is the **dual problem D**.

NPTTEL Shrihari, Shrivardhan - Numerical Optimization

Now, let us consider a problem, minimize effects subject to less than or equal to zero and x belongs to X now, for convenience we have not included equality constraints here, but they can be included very easily. Now, let us define a payoff function as a Lagrangian function, so in this case the Lagrangian is $f(x)$ plus sigma lambda j where x is from the set x and lambda j greater than or equal to 0. So, in our discussion this Lagrangian will act as a payoff function for this problem. Now, the variable x associated with this problem is called the primal variable, and the variable lambda which is a Lagrangian multiplier corresponding to the constraints that is called the dual variable.

And, if you look at the set x so that set x can be treated as a set of strategies for the first player or the primal problem, and so this x set of strategies is nothing. But set x and set of strategies for the second player or the dual problem is the set of all the lambda's in 1 dimension space because there are l inequality constraints so set of all lambda's in 1 dimension space such that all lambda's are nonnegative for every z . So, the idea is that we define the kin max function for this problem or min max function problem which is

equivalent to this problem and then get max min problem which is a dual to the min max problem. So, naturally that dual problem will be a dual to this problem also. So, we define a min max problem which is a equivalent to the primal problem P and then the corresponding max min problem will be called the dual problem D. And we will be interested in finding out the conditions under which the optimal objective functions are of the two problems P and D are equal and whether the dual problem is easier to solve then the primal problem.

(Refer Slide Time: 36:18)

Assumption: Minimum and Maximum exist for the problems defined here (Use infimum or supremum appropriately).

$$\begin{aligned} \text{Primal Function} &= \max_{\lambda \geq \mathbf{0}} \mathcal{L}(\mathbf{x}, \lambda) \\ &= \max_{\lambda \geq \mathbf{0}} f(\mathbf{x}) + \sum_{j=1}^l \lambda_j h_j(\mathbf{x}) \\ &= \begin{cases} f(\mathbf{x}) & \text{if } h_j(\mathbf{x}) \leq 0 \forall j \\ +\infty & \text{Otherwise.} \end{cases} \end{aligned}$$

Primal Problem:

$$\min_{\mathbf{x} \in X} \max_{\lambda \geq \mathbf{0}} \mathcal{L}(\mathbf{x}, \lambda)$$

That is, (ignoring the possibility of $h_j(\mathbf{x}) > 0 \forall j$),

$$\begin{aligned} \min \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & h_j(\mathbf{x}) \leq 0, j = 1, \dots, l \\ & \mathbf{x} \in X \end{aligned}$$

Now, one assumption that we make here is that minimum maximum exists for all problems which are going to be defined and if they are not if the minimum and maximum do not exist then appropriately one can replace minimum by incrementum and maximum by supremum if they do not exist. So, to avoid notational clutter, we will keep using min max and use one can use incrementum supremum appropriately. Now, the primal function as we saw earlier that the primal problem is min max of the payoff function $\psi(\mathbf{x}, \mathbf{y})$ so the primal function is max of the payoff function. So, in this case the primal function is max of $L(\mathbf{x}, \lambda)$, where L is the Lagrangian, and λ 's are nonnegative and that is nothing but max of $f(\mathbf{x})$ plus $\sum \lambda_j h_j(\mathbf{x})$ and if $h_j(\mathbf{x})$ happen to be less than or equal to zero then this quantity will always be less than or equal to 0. And therefore, the maximum of this objective function will occur at $f(\mathbf{x})$ if $h_j(\mathbf{x})$ is less than or equal to 0 and will be equal to plus infinity otherwise.

So, truly one should write this supremum because under the conditions that $h_j(x)$ greater than zero this is maximum does not exist, but as I said earlier we will continue to use this notation and one can choose to write supremum or incrementum appropriately. So, if you look at this primal objective function value that is nothing but maximum effects if this condition holds. And therefore, we have primal problem which is min max problem so minimize with respect to x and maximize $L(x, \lambda)$ with respect to λ . So, this quantity is the primal objective function value that we saw earlier and we want to minimize this.

So, we can rewrite this problem as minimize $f(x)$ subject to $h_j(x) \leq 0$ and $x \in X$ by ignoring the possibility that $h_j(x) > 0$ so if you are always interested in those parts of the feasible region where x belongs to the set X and $h_j(x) \leq 0$. Then, we can ignore this second possibility of the primal function and then we can write the problem as minimize $f(x)$ subject to $h_j(x) \leq 0$. So, this our primal problem so the min max problem by defining our Lagrangian by defining the payoff function as a Lagrangian function the min max of Lagrangian is over this feasibility conditions is same as our original constraint optimization problem.

(Refer Slide Time: 39:44)

For $\lambda \geq \mathbf{0}$, define

$$\begin{aligned} \text{Dual Function} &= \theta(\lambda) \\ &= \min_{x \in X} \mathcal{L}(x, \lambda) \\ &= \min_{x \in X} f(x) + \sum_{j=1}^l \lambda_j h_j(x) \end{aligned}$$

Dual Problem:

$$\max_{\lambda \geq \mathbf{0}} \min_{x \in X} f(x) + \sum_{j=1}^l \lambda_j h_j(x)$$

The slide also features the NPTEL logo in the bottom left corner and a small video inset of the lecturer in the bottom right corner.

Now, let us look at the dual problem so for that purpose let us define the dual function so let us take λ 's to be greater than or equal to zero so the dual function is let us

define it as theta lambda and theta lambda the dual function is a max min is a dual problem is a max min problem. So, the dual function is a minimum of the payoff function over the variable x so minimum of $L(x, \lambda)$, L this is the Lagrangian function over the set x where lambda is nonnegative.

And if you expand it further it will be minimize $f(x)$ plus sigma lambda j plus s j x. So, this is going to be our dual function and the dual problem is to maximize lambda maximize over lambda the minimum of this function over x. So, typically what would happen is that when we minimize this with respect to x we can write at the minimum x or the minimum value of x in terms of lambda's and write this as a function of lambda and then we maximize that function with respect to the condition that lambda greater than or equal to zero. So, this is our dual problem which is our max min problem.

(Refer Slide Time: 41:15)

Consider the problem:

$$\begin{aligned} \min \quad & x^2 \\ \text{s.t.} \quad & x \geq 1 \end{aligned}$$

- Primal solution: $x^* = 1, f(x^*) = 1$.
- Dual function: $\theta(\lambda) = \min_x x^2 + \lambda(1-x)$. At the minimum, $x^* = \frac{\lambda}{2}$.
- For $\lambda \geq 0, \theta(\lambda) = -\frac{1}{4}\lambda^2 + \lambda$.
- Therefore, the dual problem is

$$\max_{\lambda \geq 0} -\frac{1}{4}\lambda^2 + \lambda$$

- $\lambda^* = 2, \theta(\lambda^*) = 1$
- $f(x^*) = 1 = \theta(\lambda^*)$

NIPTEL Shrish Shekale Numerical Optimization

Now, let us consider the simple problem where we want to minimize x square subject to the constraint that x greater than or equal to one, now as you can see here that this function goes to infinity and has a minimum value which occurs at x star equal to 1 and the optimal objective functional value is 1 square which is nothing but 1. So, the optimal primal objective function value is one and that occurs at x star equal to 1. Now, let us write down the Lagrangian of this problem so the objective function x square plus lambda into 1 minus x by rewriting the constraint as h x less than or equal to zero. Now let us see what the dual function is so the dual function is to minimize 1 x lambda or x so

minimize $x^2 + \lambda(1 - x)$. So, at the minimum we have x^* equal to $\lambda/2$ and therefore, it is easier to see this because it is an unconstrained problem because our set x is nothing but \mathbb{R} .

So, minimize x over \mathbb{R} , $x^2 + \lambda(1 - x)$ and if we differentiate it with respect to x so that is $2x + \lambda = 0$ gives $x^* = -\lambda/2$. And therefore, for $\lambda \geq 0$, $\theta(\lambda)$ the dual objective function value can be written as $-\lambda^2/4 + \lambda$ and this is obtained by substituting this value of x^* in this objective function. Again I note that, this θ is not there therefore, the dual problem is maximize λ maximize $-\lambda^2/4 + \lambda$ subject to the constraint that $\lambda \geq 0$.

So, this is going to be the dual problem. As you saw, that we first consider the Lagrangian and then eliminated the primal variable from this objective function by taking the minimum. And then we substituted that x^* in the objective function and got value of λ and got the function in terms of λ and that objective function needs to be maximized with respect to the condition that $\lambda \geq 0$. Now, suppose we differentiate this, and equate it to zero and what we get is $\lambda^* = 2$ and that quantity is greater than zero.

So, clearly that is the feasible point and at $\lambda^* = 2$ the objective function will be $-1 + 2$ which is 1 . So, $\lambda^* = 2$ and $\theta(\lambda^*) = 1$ and you will see that, the optimal primal objective function value is same as the optimal duality function value. So, by converting this problem to the dual problem and finding the optimal dual value we saw that we got the same optimal objective function value as we would have got in the original case of solving the primal problem.


(Refer Slide Time: 45:15)

Geometric Interpretation

Consider the problem **(P1)**:

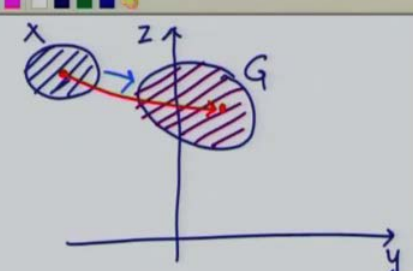
$$\begin{aligned} \min \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & h(\mathbf{x}) \leq 0 \\ & \mathbf{x} \in X \end{aligned}$$

Define $G = \{(y, z) : y = h(\mathbf{x}), z = f(\mathbf{x}), \mathbf{x} \in X\}$.



Now, let us see the geometric interpretation of this duality. Now, for convenience let us assume that there exists only one inequality constraint which is of the type $h(x) \leq 0$ or equal to zero. So, we want to minimize this objective function $f(x)$ subject to the constraint $h(x) \leq 0$ and x belongs to the set X . Now, let us define the set G to be the set of all (y, z) such that $y = h(x)$ and $z = f(x)$ for some x belongs to the set X .

(Refer Slide Time: 46:15)



$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & h(x) \leq 0 \\ & x \in X \end{aligned}$$

So, we have the problem minimize $f(x)$ subject to $h(x) \leq 0$ and x belongs to X . Now, suppose we have some set x and we have defined new space, let us call that space as a $y-z$ space note that we have said that y is equal to $h(x)$ and z equal to $f(x)$ for every x in the set x . So, this set will get transformed to some set G so this set transformed to the set G and so the set G is going to be this set. So, note that what we did was that consider every point here in this set x and transformed it to this $y-z$ space and constructed the image this x in the $y-z$ space.

For example, a point here will get transformed to some point here so, corresponding to this feasible x we take the value of x and then find out what is $h(x)$ and what $z(x)$. So, $h(x)$ and $z(x)$ denote the coordinates of this point in the $y-z$ space and if you collect all such points corresponding to every set in x we get G .

(Refer Slide Time: 48:40)




Geometric Interpretation

Consider the problem **(P1)**:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & h(x) \leq 0 \\ & x \in X \end{aligned}$$

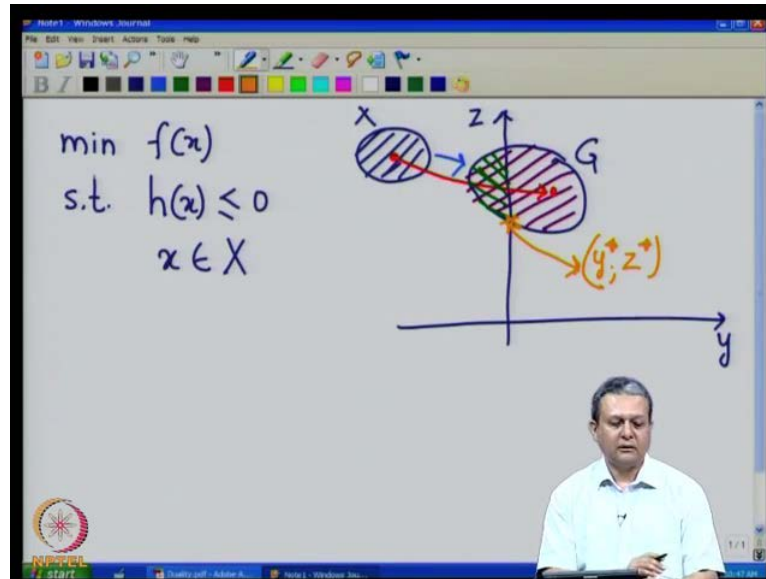
Define $G = \{(y, z) : y = h(x), z = f(x), x \in X\}$.
Let (y^*, z^*) be a minimum of **P1** in $y-z$ space.

⊛

Now, so we have denoted the set G which is obtained from every x in x and by using y equal to $h(x)$ and z equal to $f(x)$. So, let us denote, the minimum of the objective of the problem in the $y-z$ space as $y^* z^*$.

(Refer Slide Time: 49:08)



So, this is our feasible region the set G that is our feasible region. As far as x is concerned, but remember that we also to satisfy that $h(x) \leq 0$ or in other words $y \leq 0$. So, that means the actual feasible region for the original problem is because y has to be less than or equal to zero so this is going to be the actual feasible region for the original problem the set G denotes just the image of the set x and not really the actual feasible region. Because we have ignored $h(x) \leq 0$, but here we consider $h(x) \leq 0$ also and this forms our feasible region.

Now, we want to find out the minimum of f of x in the feasible region and you note that we have defined z to be $f(x)$. So, we want to find out what is the minimum value of z with respect to this feasible region and turns out that this point which is y^* z^* is our minimum point so in this space we are interested in finding out a solution which is y^* z^* .


(Refer Slide Time: 51:06)

Geometric Interpretation

Consider the problem **(P1)**:

$$\begin{aligned} \min \quad & f(\mathbf{x}) \\ \text{s.t.} \quad & h(\mathbf{x}) \leq 0 \\ & \mathbf{x} \in X \end{aligned}$$

Define $G = \{(y, z) : y = h(\mathbf{x}), z = f(\mathbf{x}), \mathbf{x} \in X\}$.
Let (y^*, z^*) be a minimum of **P1** in $y - z$ space.
For $\lambda \geq 0$, define $\theta(\lambda) = \min_{\mathbf{x} \in X} f(\mathbf{x}) + \lambda h(\mathbf{x})$.
Given $\lambda \geq 0$, $\theta(\lambda)$ minimizes $z + \lambda y$ over G in $y - z$ space.

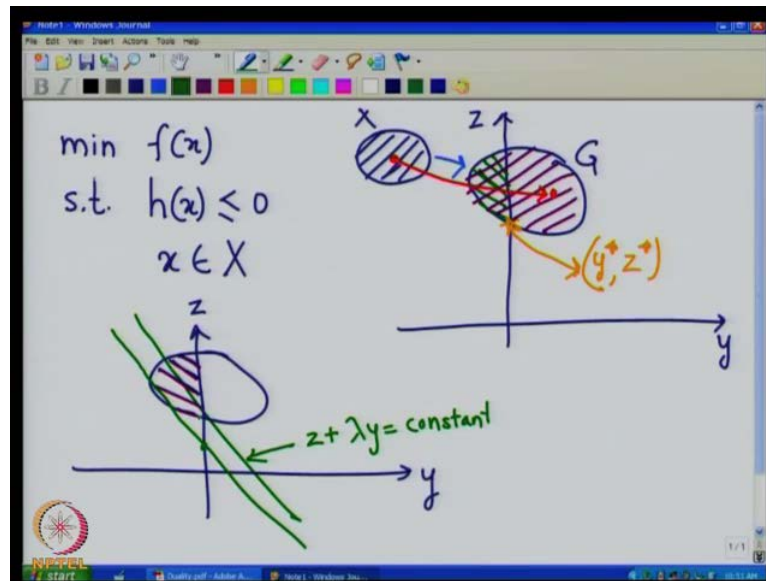


Shreshth Shrivastava Numerical Optimization

Now, let us look at the dual objective function so for lambda which is nonnegative. Let us define, the dual objective function to be theta lambda which is equal to min of the Lagrangian and the Lagrangian function in this case is $f(\mathbf{x}) + \lambda h(\mathbf{x})$ where lambda is a Lagrangian multiplier corresponding to this inequality constraint. So what we want to do is that if we are given lambda which is nonnegative then theta lambda will find out the minimum of $f(\mathbf{x}) + \lambda h(\mathbf{x})$ or in other words theta lambda minimizes in the $y - z$ space it minimizes $z + \lambda y$ over G in the $y - z$ space over that part of G where $h(\mathbf{x}) \leq 0$.

So, if you look at this $z + \lambda y$, $z + \lambda y$ is an equation of a line in the $y - z$ space with $z + \lambda y = \text{constant}$ is an equation of a line in the $y - z$ space where $-\lambda$ is the slope of the line.

(Refer Slide Time: 52:46)



So, if we are given some lambda let us look at this figure, so we are interested in this part. So, $z + \lambda y = \text{constant}$ is an equation of a line with slope minus lambda. And, lambda is greater than or equal to zero, so minus lambda will be less than or equal to zero, so we could have a line. So, this is the equation of a line $z + \lambda y = \text{constant}$ where lambda is greater than or equal to zero. Now, when we want to minimize $z + \lambda y$ we want to find out that $z + \lambda y$ which gives us the least value with respect to the feasible region so with respect to this feasible region will get a quantity which is supporting this set the shaded region here.

And the optimal objective functional value is, the value corresponding to z that we get corresponding to the lines intercept on the z axis. Now, among all those possible lambda's we want to find out that lambda which gives us this objective functional value and how to do that we will see that in the next class.

Thank you.