10-725/36-725: Convex Optimization **Fall 2019**

Lecture 10: September 30

Lecturer: Ryan Tibshirani *Scribes: Bingqing Chen, Zhaoqi Cheng, Helen Zhou*

10.1 Duality in linear programs

Suppose we want to find lower bound on the optimal value in a convex problem, such that $B \leq \min_x f(x)$.

Consider this linear program (LP) as an example,

$$\begin{aligned} \min_{x,y} \quad & px + qy \\ \text{subject to} \quad & x + y \geq 2 \\ & x, y \geq 0 \end{aligned}$$

Suppose $a, b, c \geq 0$,

$$\begin{aligned} a(x+y) &\geq 2a \\ bx &\geq 0 \\ cy &\geq 0 \end{aligned}$$

Summing the three equations together, gives one

$$(a+b)x + (a+c)y \geq 2a$$

If we constrain $a + b = p$ and $a + c = q$, then $px + qy$ is lower bounded by $2a$.

Thus, the original LP, which we call it the primal problem, is equivalent to the alternative formulation here, which we call the dual problem.

$$\begin{aligned} \max_{a,b,c} \quad & 2a \\ \text{subject to} \quad & a + b = p \\ & a + c = q \\ & a, b, c \geq 0 \end{aligned}$$

One thing to note is that the number of variables in the dual problem is the same as the number of constraints in the primal problem.

10.1.1 Duality for general form linear programs

Now we extend the example to general form LP.

Given $c \in \mathbb{R}^n$, $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, $G \in \mathbb{R}^{r \times n}$, $h \in \mathbb{R}^r$:

Primal LP:

$$\begin{aligned} \min_{x} \quad & c^T x \\ \text{subject to} \quad & Ax = b \\ & Gx \geq h \end{aligned}$$

Similar to the example,

$$\begin{aligned} u^T(Ax - b) &= 0 \\ v^T(Gx - h) &\leq 0, \forall v \geq 0 \\ u^T(Ax - b) + v^T(Gx - h) &\leq 0 \\ (-A^T u - G^T v)^T x &\geq -b^T u - h^T v \end{aligned}$$

Thus, the primal LP is equivalent to the dual LP.

Dual LP:

$$\begin{aligned} \max_{u,v} \quad & -b^T u - h^T v \\ \text{subject to} \quad & -A^T u - G^T v = c \\ & v \geq 0 \end{aligned}$$

10.2 Max flow and min cut problem

One example of LP duality is the max flow and min cut problem. It has an interesting history, with much of the development taking place in the context of rail systems [1].

Flow One could think of flow as some fluid or mass flowing through pipes. The fluid/ mass is nonnegative, pipes have limited capacity, and no mass is lost.

More formally, suppose we have a directed graph $G = (V, E)$. One node is labeled as a source node s, and another as a sink node t. We define a flow as a vector f_{ij}, where $(i, j) \in E$ that satisfies:

- Nonnegativity of flow being pushed in the direction of the edge: $f_{ij} \geq 0, (i, j) \in E$
- Flow capacity per edge: $f_{ij} \leq c_{ij}, (i, j) \in E$
- Conservation of flow (except source/ sink nodes): $\sum_{(i,k)\in E} f_{ik} = \sum_{(k,j)\in E} f_{kj}, k \in V\backslash\{s, t\}$

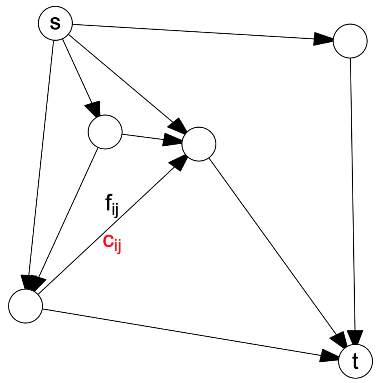

Figure 10.1: Directed graph demonstrating the max flow/ min cut setup.

Max flow problem Given a graph, we want to find the flow that maximizes the the sum of the flows emanating from the source s. As a linear program, this can be written as follows:

$$\max_{f \in \mathbb{R}^{|E|}} \sum_{(s,j) \in E} f_{sj} \tag{10.1}$$

$$\text{subject to } 0 \leq f_{ij} \leq c_{ij} \text{ for all } (i,j) \in E \tag{10.2}$$

$$\sum_{(i,k) \in E} f_{ik} = \sum_{(k,j) \in E} f_{kj} \text{ for all } k \in V \backslash \{s,t\} \tag{10.3}$$

Deriving the dual We can derive the dual in a couple of steps. Since the primal is a maximization problem, the dual will be a minimization problem. Let us assign nonnegative dual variables a_{ij}, b_{ij}, and x_k to the constraints in 10.2 (nonnegativity), 10.2 (capacity), and 10.3 (conservation) respectively.

This gives us:

$$\begin{aligned} &\max_{f \in \mathbb{R}^{|E|}} \sum_{(s,j) \in E} f_{sj} \\ \text{subject to } & -a_{ij} f_{ij} \leq 0 \text{ for all } (i,j) \in E \\ & b_{ij} f_{ij} \leq b_{ij} c_{ij} \text{ for all } (i,j) \in E \\ & x_k \left(\sum_{(i,k) \in E} f_{ik} - \sum_{(k,j) \in E} f_{kj} \right) = 0 \text{ for all } k \in V \backslash \{s,t\} \end{aligned}$$

Adding everything together and rearranging, we have:

$$\sum_{(i,j) \in E} (-a_{ij} f_{ij} + b_{ij}(f_{ij} - c_{ij})) + \sum_{k \in V \backslash \{s,t\}} x_k \left(\sum_{(i,k) \in E} f_{ik} - \sum_{(k,j) \in E} f_{kj} \right) \leq 0$$

for any $a_{ij}, b_{ij} \geq 0, (i,j) \in E$, and $x_k, k \in V \backslash \{s,t\}$.

Pulling out the terms multiplying f_{ij} into $M_{ij}(a,b,x)$, we can rewrite this as:

$$\sum_{(i,j)\in E} M_{ij}(a,b,x) f_{ij} \leq \sum_{(i,j)\in E} b_{ij}c_{ij}$$

Looking at the primal objective, we were only interested in maximizing the sum of f_{sj}'s, not all f_{ij}'s. Thus, we want a, b, and x that satisfy:

$$\begin{cases} M_{sj} = b_{sj} - a_{sj} + x_j & \text{want this } = 1 \\ M_{it} = b_{it} - a_{it} + x_i & \text{want this } = 0 \\ M_{ij} = b_{ij} - a_{ij} + x_j - x_i & \text{want this } = 0 \end{cases}$$

Given that a, b, and x satisfy the above constraints, we have shown that the primal optimal value is upper bounded by $\sum_{(i,j)\in E} b_{ij}c_{ij}$. Therefore, the dual problem is:

$$\begin{aligned} \min_{b\in\mathbb{R}^{|E|}, x\in\mathbb{R}^{|V|}} & \sum_{(i,j)\in E} b_{ij}c_{ij} \\ \text{subject to } & b_{ij} + x_j - x_i \geq 0 \text{ for all } (i,j) \in E \\ & b \geq 0, x_s = 1, x_t = 0 \end{aligned}$$

where we are able to eliminate a_{ij} because we know that $a_{ij} \geq 0$, and since it only shows up in the equality constraints we can change them to inequality constraints and remove the a_{ij} terms.

LP relaxation of the min cut problem At first pass, the above formulation is not immediately interpretable. However, suppose that at the solution,

$$x_i \in \{0,1\} \text{ for all } i \in V$$

Let $A = \{i : x_i = 1\}, B = \{i : x_i = 0\}$. Note that by our constraints, $s \in A$ and $t \in B$. Then, our constraint

$$b_{ij} \geq x_i - x_j \text{ for } (i,j) \in E, b \geq 0$$

implies that $b_{ij} = 1$ if $i \in A$ and $j \in B$, and 0 otherwise. That is, $b_{ij} = 1$ for the edges that cross between sets A and B. Thus, the objective $\sum_{(i,j)\in E} b_{ij}c_{ij}$ is the capacity of the cut defined by A, B.

Therefore, we have argued that the dual is the LP relaxation of the min cut problem:

$$\begin{aligned} \min_{b\in\mathbb{R}^{|E|}, x\in\mathbb{R}^{|V|}} & \sum_{(i,j)\in E} b_{ij}c_{ij} \\ \text{subject to } & b_{ij} \geq x_j - x_i \text{ for all } (i,j) \in E \\ & b_{ij}, x_i, x_j \in \{0,1\} \text{ for all } i,j \end{aligned}$$

Therefore,

$$\text{value of max flow} \leq \text{optimal value for LP relaxed min cut} \leq \text{capacity of min cut}$$

where the first inequality follows from duality, and the second inequality follows from optimization over a larger constraint set.

In fact, there's a famous result called the **max flow min cut theorem**, which states that the value of the max flow through a network is exactly the capacity of the min cut. Thus, all of the above inequalities are actually equalities, so the dual is tight and we get strong duality.

10.3 Alternative perspective on LP duality

The duality of general form LP was derived in section 10.1.1 with algebra. The same results can be derived from an alternative perspective.

$$c^T x \geq c^T + \underbrace{u^T(Ax-b)}_{=0} + \underbrace{v^T(Gx-h)}_{\leq 0} =: \mathcal{L}(x,u,v)$$

So if C denotes primal feasible set, f^* primal optimal value, then for any u and $v \geq 0$,

$$f^* \geq \min_{x \in C} \mathcal{L}(x,u,v) \geq \min_{x} \mathcal{L}(x,u,v) =: g(u,v)$$

Thus, $g(u,v)$ is a lower bound on f^* for any u and $v \geq 0$. We can maximize $g(u,v)$ over u and $v \geq 0$ to get the tightest lower bound.

Note that,

$$g(u,v) = \begin{cases} -b^T u - h^T v & \text{if } c = -A^T u - G^T v \\ -\infty & \text{else} \end{cases}$$

This gives us exactly the dual LP.
Note, that this formulation is completely general and applies to arbitrary optimization problems.

10.4 Example of duality: mixed strategies for matrix games

Consider a game with two players R and J, and a payoff matrix P, where J will pay R amount P_{ij} if they each choose i and j respectively in a round.

		R			
		1	2	...	n
	1	P_{11}	P_{12}	...	P_{1n}
J	2	P_{21}	P_{22}	...	P_{2n}
	...				
	m	P_{m1}	P_{m2}	...	P_{mn}

Table 10.1: Payoff matrix

Both players use mixed stragies, namely, each will first specify a probability distribution

$$\begin{aligned} x: &\quad \mathbb{P}(\text{J chooses } i) = x_i,\ i = 1, ..., m \\ y: &\quad \mathbb{P}(\text{R chooses } j) = y_j,\ j = 1, ..., m; \end{aligned}$$

such that the expected payout from J to R is

$$\sum_{i=1}^{m} \sum_{j=1}^{n} x_i y_j P_{ij} = x^T P y.$$

Now suppose there are two scenarios for this game, depending on which side knows the strategy of his counterpart beforehand.

Scenario 1: R know J's strategy ahead of time

In this case, J will allow R to know his strategy x ahead of time such that R will choose y to maximize his payoff

$$\max x^T P y : y \geq 0, 1^T y = 1 = \max_{i=1,\ldots,n} (P^T)_i.$$

Accordingly, J's chooses x to minimize the payoff of R

$$\begin{aligned} \min_{x} \quad & \max_{i=1,\ldots,n} (P^T x)_i \\ \text{subject to} \quad & x \geq 0, 1^T x = 1 \end{aligned}$$

We can rewrite the problem in the LP form

$$\begin{aligned} \min_{x,t} \quad & t \\ \text{subject to} \quad & x \geq 0, 1^T x = 1 \\ & P^T x \leq t. \end{aligned} \tag{10.4}$$

Scenario 2: J know R's strategy ahead of time

In the alternative scenario, suppose J knows R's strategy y before hand, then the optimization problem becomes

$$\begin{aligned} \max_{y,v} \quad & v \\ \text{subject to} \quad & y \geq 0, 1^T y = 1 \\ & Py \geq v \end{aligned} \tag{10.5}$$

Denote R's expected payoff in the first scenario f_1^* and that in the second scenario f_2^*. Since it is advanageous to know the other player's strategy, $f_1^* \geq f_2^*$. However, by Von Neumman's minimax theorem, the inequality here is actually equality $f_1^* = f_2^*$.

Why? The Lagrangian of the first problem is

$$L(x, t, u, v, y) = t - u^T x + v(1 - 1^T x) + y^T (P^T x - t1)$$

and hence the Lagrange dual function is

$$\begin{aligned} g(u, v, y) &= \min_{x,t} L(x, t, u, v, y) \\ &= \begin{cases} v & \text{if } 1 - 1^T y = 0, Py - u - v1 = 0 \\ -\infty & \text{otherwise.} \end{cases} \end{aligned} \tag{10.6}$$

Comparing 10.5 and 10.6, we can see they are exactly the same problem. In other words, the strong duality holds.

References

[1] Schrijver, Alexander. "On the History of the Transportation and Maximum Flow Problems." Mathematical Programming, vol. 91, no. 3, 2002, pp. 437–445.