10-725/36-725: Convex Optimization **Fall 2018**

Lecture 10: October 1

Lecturer: Lecturer: Ryan Tibshirani *Scribes: Minshi Peng, Chenlei Fang, Qicheng Huang*

Note: *LaTeX template courtesy of UC Berkeley EECS dept.*

10.1 Lower bound in linear program

Suppose we want to find lower bound on the optimal value in our convex problem, $B \leq \min_x f(x)$. For example, consider the following simple linear program

$$\min_{x,y}, \text{ subject to } x + y \geq 2, x, y \geq 0$$

It is easy to see that the lower bound is $B = 2$, because one of the constraints is exactly the same as the objective function.

Let us try other problems. Suppose the linear program is

$$\min_{x,y} x + 3y, \text{ subject to } x + y \geq 2, x, y \geq 0$$

We can have a lower bound for the objective function by $x + 3y = (x + y) + 2y \geq 2$

More generally, for linear program

$$\min_{x,y} px + qy, \text{ subject to } x + y \geq 2, x, y \geq 0 \tag{10.1}$$

the constraint can be equivalently represented as

$$ax + by \geq 2a, ax \geq 0, cy \geq 0, a, b, c \geq 0$$

Adding them together, we will get

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

Let $a + b = p$ and $a + c = q$, then we obtain the lower bound of linear program 10.1 as $B = 2a$, for any a, b and c such that

$$a + b = p, a + c = q, a, b, c \geq 0$$

This gives us one linear bound. But what is the best lower bound. Simply we can maximize our lower bound over all possible a, b, c:

$$\max_{a,b,c} 2a, \text{ subject to } a + b = p, a + c = q, a, b, c \geq 0$$

This is also called the dual linear program of primal problem 10.1. Note that the number of dual variables is the number of primal constraints.

Now let us see another problem

$$\min_{x,y} px + qy, \text{ subject to } x \geq 0, y \leq 1, 3x + y = 2$$

The constraint of the linear program can be equivalently represented as

$$ax \geq 0, -by \geq -b, 3cx + cy = 2c, a \geq 0, b \geq 0$$

Adding them together, we get

$$(a + 3c)x + (-b + c)y \geq -b + 2c$$

Let $p = a + 3c$ and $q = b + c$, we obtain the dual problem

$$\max_{a,b,c} 2c - b, \text{ subject to } a + 3c = p, -b + c = q, a, b \geq 0$$

Here we formulate the Duality for general form LP as following. Given $c \in \mathbb{R}^n, A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m, G \in \mathbb{R}^{r\times n}, h \in \mathbb{R}^r$, the primal LP problem is

$$\min_{x} c^T x, \text{ subject to } Ax = b, Gx \leq h \tag{10.2}$$

Then the dual LP is

$$\max_{u,v} -b^T u - h^T v, \text{ subject to } - A^T u - G^T v = c, v \geq 0 \tag{10.3}$$

Observe that for any u and $v \geq 0$ and x is primal feasible, we get

$$u^T(Ax - b) + v^T(Gx - h) \leq 0$$

and by rearranging it becomes

$$(-A^T u - G^T v)^T v \geq -b^T u - h^T v$$

so that if $c = -A^T u - G^T v$, we get a bound on primal optimal value.

10.2 Example: Max Flow and Min Cut Problems

The max flow and min cut problems are essentially equivalent because of duality. First let's describe the problem setting: given a directed graph $G = (V, E)$, define $f_{ij}, (i, j) \in E$ as the flow from node i to j . Also given c_{ij} as the capacity of the edge, which is the maximum amount of flow that one can push through that edge. In addition, the flow going into the node has to be equal to the flow coming out of the node. That is true for all nodes except for the source (s) and the sink (t) nodes. These constraints can be formulated as:

$$\begin{aligned} f_{ij} &\geq 0, (i, j) \in E \\ f_{ij} &\leq c_{ij}, (i, j) \in E \\ \sum_{(i,k)\in E} f_{ik} &= \sum_{(k,j)\in E} f_{kj}, k \in V \setminus \{s, t\} \end{aligned} \tag{10.4}$$

The **max flow problem** wants to maximize the total flow leaving the source, subject to satisfying the constraints. It can be expressed as an LP:

$$
\begin{aligned}
&\max_{f\in\mathbb{R}^{|E|}} \sum_{(s,j)\in E} f_{sj} \\
&\text{subject to } 0 \le f_{ij} \le c_{ij}, (i,j) \in E \\
&\sum_{(i,k)\in E} f_{ik} = \sum_{(k,j)\in E} f_{kj}, \text{ for all } k \in V \setminus \{s,t\}
\end{aligned} \tag{10.5}
$$

Now, we can derive the dual of this problem. Assign dual variables a, b, x to the primal constraints, we have the following inequality:

$$
\sum_{(i,k)\in E} \big(-a_{ij} f_{ij} + b_{ij}(f_{ij} - c_{ij}) \big) + \sum_{k\in V\setminus\{s,t\}} x_k \Big(\sum_{(i,k)\in E} f_{ik} - \sum_{(k,j)\in E} f_{kj} \Big) \le 0 \tag{10.6}
$$

for any $a_{ij}, b_{ij} \ge 0, (i,k) \in E$ and $x_k, k \in V \setminus \{s,t\}$

The constraint can be rearranged as

$$
\sum_{(i,k)\in E} M_{ij}(a,b,x) f_{ij} \le \sum_{(i,k)\in E} b_{ij} c_{ij} \tag{10.7}
$$

where $M_{ij}(a,b,x)$ collects terms multiplying f_{ij}. To make the LHS in 10.7 equal to primal objective, we need to make the coefficients of f_{ij} equal to the coefficients in primal, i.e.,

$$
\begin{aligned}
M_{sj} &= b_{sj} - a_{sj} + x_j \text{ want this} = 1 \\
M_{it} &= b_{it} - a_{it} - x_i \text{ want this} = 0 \\
M_{ij} &= b_{ij} - a_{ij} + x_j - x_i \text{ want this} = 0
\end{aligned} \tag{10.8}
$$

We have shown that

$$
\text{primal optimal value} \le \sum_{(i,k)\in E} b_{ij} c_{ij} \tag{10.9}
$$

subject to a, b, x satisfying constraints. Hence the dual problem is (minimize over a, b, x to get best upper bound, a has been eliminated and incorporated into constraints)

$$
\begin{aligned}
&\min_{b\in\mathbb{R}^{|E|}, x\in\mathbb{R}^{|V|}} \sum_{(i,k)\in E} b_{ij} c_{ij} \\
&\text{subject to } b_{ij} + x_j - x_i \ge 0, \text{ for all } (i,j) \in E \\
&b \ge 0, x_s = 1, x_t = 0
\end{aligned} \tag{10.10}
$$

In the above derivation, we start from a maximization problem and end with a minimization dual problem. Next, we will show how this dual problem can be formulated as a relaxation of min cut problem.

Suppose at the the solution of 10.10, it just so happened that $x_i \in \{0,1\}$ for all $i \in V$.

Let $A = \{i : x_i = 1\}, B = \{i : x_i = 0\}$, that is, we separate the nodes into two groups. Note that $s \in A, t \in B$. Then we have

$$
b_{ij} \ge x_i - x_j \text{ for } (i,j) \in E, b \ge 0 \tag{10.11}
$$

imply that $b_{ij} = 1$ if $i \in A$ and $j \in B$, and 0 otherwise. Moreover, the objective $\sum_{(i,j)\in E} b_{ij}c_{ij}$ is the capacity of cut defined by A, B. That is, we have shown that the dual is the LP relaxation of the min cut problem:

$$\begin{aligned} \min_{b\in\mathbb{R}^{|E|}, x\in\mathbb{R}^{|V|}} & \sum_{(i,k)\in E} b_{ij}c_{ij} \\ \text{subject to } & b_{ij} \geq x_i - x_j, \\ & b_{ij}, x_i, x_j \in \{0,1\} \text{ for all } i,j \end{aligned} \tag{10.12}$$

By construction, we know value of max flow is less than or equal to that of dual. The minimization of a relaxed LP problem is smaller than that of the min cut. From what we have known so far, we have:

$$\text{value of max flow} \leq \text{optimal value for LP relaxed min cut} \leq \text{capacity of min cut} \tag{10.13}$$

A famous result called **max flow min cut theorem** states: value of max flow through a network is exactly the capacity of the min cut. Hence in 10.13, we have all the equalities. In particular, we get that the primal LP and dual LP have exactly the same optimal values, a phenomenon called **strong duality**, which will be explained in more details in the next lecture.

10.3 Another perspective on LP duality

Now let's look at the LP duality problem from another perspective, i.e., we will illustrate the correctness of the following equivalence problem:

The primal:

$$\begin{aligned} \min_{x} \ & c^T x \\ s.t. \ & Ax = b \\ & Gx \leq h \end{aligned} \tag{10.14}$$

is equivalent to the dual:

$$\begin{aligned} \max_{u,b} \ & -b^T u - h^T v \\ s.t. \ & -A^T u - G^T v = c \\ & v \geq 0 \end{aligned} \tag{10.15}$$

For any u and $v \geq 0$, and x primal feasible,

$$c^T x \geq c^T x + u^T(Ax - b) + v^T(Gx - h) := L(x,u,v) \tag{10.16}$$

and

$$f^* \geq \min_{x\in C} L(x,u,v) \geq \min_{x} L(x,u,v) := g(u,v) \tag{10.17}$$

where C denotes primal feasible set and f^* denotes the primal optimal value.

Note that $g(u,v)$ is a lower bound of f^* and we have:

$$g(u,v) = \begin{cases} -b^T u - h^T v & \text{if} \quad c = -A^T u - G^T v \\ -\infty & \text{otherwise} \end{cases} \tag{10.18}$$

By maximize $g(u, v)$ over u and $v \geq 0$, we can get the tightest bound, which is exactly the dual LP.

The primal-dual form in (10.14) and (10.15) is more general and applies to even nonconvex problems.

Example: mixed strategies for matrix games

Setup: two players, J and R, and a payout matrix P.

Game: if J chooses i amd R choose j, then J must pay R amount P_{ij}. P_{ij} can be either positive or negative.

They use *mixed strategies*: each will first specify a probability distribution:

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

The expected payout from J to R is then:

$$\sum_{i=1}^{m} \sum_{j=1}^{n} x_i y_j P_{ij} = x^T P y \tag{10.19}$$

Now consider two opposite universes:

Universe 1:

J allows R to know x ahead of time. In this case, R will choose y to maximize $x^T P y$, which results in J paying off

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

J's best strategy is then to find the x by solving

$$\begin{aligned} &\min_x \ \max_{i=1,\dots n} (P^T x)_i \\ &s.t.\ \ x \geq 0,\ \ 1^T x = 1 \end{aligned} \tag{10.20}$$

Universe 2:

R allows J to know y before time. By the same logic, R's best strategy is then to find the y by solving

$$\begin{aligned} &\max_y \ \min_{j=1,..m} (P y)_j \\ &s.t.\ \ y \geq 0,\ \ 1^T y = 1 \end{aligned} \tag{10.21}$$

Surprisingly, the expected payouts in the two universes (denoted as f_1^* and f_2^*, respectively) are actually the same, i.e., $f_1^* = f_2^*$ by *Von Neumman's minimax theorem.*

Simple proof:

Recast the problem in (10.20) as an LP:

$$\begin{aligned} &\max_{x,t} \ t \\ &s.t.\ \ x \geq 0,\ \ 1^T x = 1 \\ &\qquad P^T x \leq t \end{aligned} \tag{10.22}$$

Now form the Lagrangian:

$$L(x,t,u,v,y) = t - u^T x + v(1 - 1^T x) + y^T(P^T x - t1) \tag{10.23}$$

and the Lagrange dual function:

$$g(u,v) = \min_{x,t} L(x,t,u,v,y) = \begin{cases} v & \text{if} \quad 1 - 1^T y = 0,\ Py - u - v1 = 0 \\ -\infty & \text{otherwise} \end{cases} \tag{10.24}$$

Hence the dual problem, after eliminating slack variable u, is

$$\begin{aligned} &\max_{y,v}\ v \\ &s.t.\ \ y \geq 0,\ \ 1^T y = 1 \\ &\qquad Py \geq v \end{aligned} \tag{10.25}$$

This is exactly the problem in (10.21), and therefore *strong duality* holds. In LPs, strong duality holds unless both the primal and dual are infeasible.

10.4 Contributions

Section 10.1: Minshi Peng

Section 10.2: Chenlei Fang

Section 10.3: Qicheng Huang