6.854 Advanced Algorithms

Lecture 15: October 15, 2003 Lecturer: David Karger and Erik Demaine
Scribes: Nelson Lai

15.1 Addendum from last lecture

Theorem 1 *If the primal P (primal) or D (dual) are feasible, then they have the same value.*

15.2 Rules for Taking Duals

In general we construct the primal P as a minimization problem and, conversely, the dual D as a maximization problem. If P is a linear program in standard form given by:

$$\begin{aligned} z &= \min(c^T x) \\ Ax &\geq b \\ x &\geq 0 \end{aligned}$$

then the dual, D is given by:

$$\begin{aligned} w &= \max(b^T y) \\ A^T y &\leq c \\ y &\geq 0 \end{aligned}$$

In general, the form of the dual will depend on the form of the primal. If one is given a primal linear program P in mixed form:

$$\begin{aligned} x &= \min(c_1 x_1 + c_2 x_2 + c_3 x_3) \\ A_{11}x_1 + A_{12}x_2 + A_{13}x_3 &= b_1 \\ A_{21}x_1 + A_{22}x_2 + A_{23}x_3 &\geq b_2 \\ A_{31}x_1 + A_{32}x_2 + A_{33}x_3 &\leq b_3 \\ x_1 &\geq 0 \\ x_2 &\leq 0 \\ x_3 & \quad \text{unrestricted in sign (UIS)} \end{aligned}$$

then the dual D is given by:

$$
\begin{aligned}
w &= \max(b_1y_1 + b_2y_2 + b_3y_3) \\
y_1A_{11} + y_2A_{21} + y_3A_{31} &\leq c_1 \\
y_1A_{12} + y_2A_{22} + y_3A_{32} &\geq c_2 \\
y_1A_{13} + y_2A_{23} + y_3A_{33} &= c_3 \\
y_1 & \quad \text{unrestricted in sign (UIS)} \\
y_2 &\geq 0 \\
y_3 &\leq 0
\end{aligned}
$$

By simple transformations, we can confirm that this is consistent with the dual for the standard form of the primal and that in fact the dual of the dual is the primal.

We can summarize these results with the following table which states the rules for taking duals. Note that each variable in the primal corresponds to a variable in the dual and each constraint in the primal corresponds to a variable in the dual.

PRIMAL	minimize	maximize	DUAL
constraints	$\geq b_i$ $\leq b_i$ $= b_i$	≥ 0 ≤ 0 unrestricted	variables
variables	≥ 0 ≥ 0 unrestricted	$\leq c_j$ $\leq c_j$ $= c_j$	constraints

Note that this makes intuitive sense. For example, the primal minimization problem has lower bounds as the natural constraints. This corresponds to a positive variable in the dual maximization problem. Conversely, the primal maximization problem has upper bounds as natural constraints. The dual minimization problem now has a negative variable.

To develop an intuition for these relationships, we consider the effect of the sign of a variable in the minimization problem on the type of the corresponding constraint in the maximization problem. We know from weak duality that $c^Tx \geq yb = yAx$. Consider the case where $x_1 \geq 0$. Then in order to have $yAx_1 \leq c_1x_1$, we must have $yA_{11} \leq c_1$ for any y. Similarly, if $x_2 \leq 0$, then we must have $yA_{12} \geq c_2$ in order for $c^Tx \geq yAx$. Finally, for x_3 unrestricted, we must have $yA_{13} = c_3$ since multiplying both sides by x might or might not change the direction of any inequality. In general, tighter constraints in the primal lead to looser constraints in the dual. An equal constraint leads to an unrestricted variable and adding new constraints creates new variables and more flexibility.

15.3 Shortest Paths

We now examine an example showing the relationship between the primal and dual problems. We consider formulating the shortest paths problem as a linear program. Given a graph G, we wish to find the shortest path from any one point (the source) to any other point. We formulate the problem as a dual (or maximization) linear program.

$$\begin{aligned} w &= \max(d_t - d_s) \\ d_j - d_i &\leq c_{ij} \\ d_j & \quad \text{unrestricted} \end{aligned}$$

Each variable d_i represents the distance to vertex i and each constraint represents the triangle inequality — that is, the the distance to vertex i should always be less than or equal to the distance to vertex j plus the distance from vertex j to vertex i. Any feasible solution to this would find a lower bound to the shortest path distances — the maximization objective makes sure these shortest path distances are valid. You can imagine physically holding up the source and the sink and pulling them apart slowly. The first time we cannot pull any further, this indicates the shortest path has been reached.

The constraint matrix A has n^2 rows and n columns of ± 1 or 0. Each row ij has a 1 in column i, -1 in column j, and 0 in all others. Thus we can write the primal as follows:

$$\begin{aligned} z &= \min(c^T x) \\ &= \sum_{i,j} c_{ij} x_{ij} \\ \sum_{j=1}^{n} x_{js} - x_{sj} &= -1 \\ \sum_{j=1}^{n} x_{jt} - x_{tj} &= 1 \\ \sum_{j=1}^{n} x_{ji} - x_{ij} &= 0 \qquad \forall i \neq s, t \end{aligned}$$

But this is simply a linear program for a minimum cost unit-flow! The constraints represent the conservation of flow with one unit of flow going into the sink and one unit coming out from the source. All other vertices are constrained to have the same amount of flow coming in as going out. Thus any feasible solution to the linear program will be a feasible flow. The objective function simply tries to minimize the cost of this flow. We see that often the dual of a LP allows us to understand the problem from a different (but equivalent) perspective.

15.4 The Gravitational Model

Consider a linear program $\min\{cx \mid Ax \geq b\}$. We consider a hollow polytope defined by a set of constraints. Let c be the gravitation vector, pointing straight up. We can put a ball in the polytope, and let it fall to the bottom.

At equilibrium point x^*, the forces exerted by the floors are balanced by the gravitational force. The normal forces by the floors are aligned with the A_i's. Therefore, we have $c = \sum y_i A_i$ for some **nonnegative** force coefficients y_i. In particular, y^* is a feasible solution for $\max\{yb \mid yA = c, y \geq 0\}$. Since the forces can be only be exerted by those walls touching the ball, we have $y_i = 0$ if $A_i x > b_i$. Therefore, we have

$$y_i(a_i x - b_i) = 0,$$

thus,

$$yb = \sum y_i(a_i x_i) = cx,$$

which means that y^* is dual optimal.

15.5 Complementary Slackness

The above example leads to the idea of *complementary slackness.* Given feasible solutions x and y, $cx - by \geq 0$ is called the *duality gap.* The solutions are optimal if and only if the gap is zero. Therefore, the gap is a good measure of "how far off" we are from the optimum.

Going back to original primal and dual forms, we can rewrite the dual: $yA + s = c$ for some $s \geq 0$ (that is, $s = c_j - yA_j$).

Theorem 2 *The followings are equivalent for feasible x and y:*

- *x and y are optimal*
- $sx = 0$
- $x_j s_j = 0$ *for all j*
- $s_j > 0$ *implies* $x_j = 0$

Proof: First, $cx = by$ if and only if

$$(yA + s)x = (Ax)y,$$

thus $sx = 0$. If $sx = 0$, then since $s, x \geq 0$, we have have $s_j x_j = 0$, so of course $s_j > 0$ forces $x_j = 0$. The converse is easy. ■

The basic idea of complementary slackness is that an optimum solution cannot have a variable x_j and corresponding dual constraint s_j slack at same time — one must be tight.

This can be stated in another way:

Theorem 3 *In arbitrary form LPs, feasible points optimal if:*

$$\begin{aligned} y_i(a_i x - b_i) &= 0 \quad \forall i \\ (c_j - yA_j)x_j &= 0 \quad \forall j \end{aligned}$$

Proof: Note that in the definition of primal/dual, feasibility means $y_i(a_i x - b_i) \geq 0$ (since $\geq$ constraint corresponds to nonnegative y_i). Also, $(c_j - yA_j)x_j \geq 0$, thus

$$\begin{aligned} \sum y_i(a_i x - b_i) + (c_j - yA_j)x_j &= yAx - yb + cx - yAx \\ &= cx - yb \\ &= 0 \end{aligned}$$

at optimum. But since all terms are nonnegative, they must be all 0. ∎

15.6 Examples Using Complementary Slackness

In some linear optimization problems, we can gain new insight by investigating its primal and dual optimal solutions using complementary slackness. We are going to give two examples. In the first example, we will consider the LP formulation of the maximum flow problem. Using complementary slackness, we derive the *Max-Flow Min-Cut Theorem*. In the second example, we consider the minimum cost circulation problem. Using the linear programming framework, we give an alternative proof of the complementary slackness property introduced in lecture 13 (the lecture on minimum cost flow).

15.6.1 Max-flow Min-Cut Theorem

In the maximum flow problem, we can imagine the network has an arc (t, s) with infinite capacity. And we are maximizing the flow on that arc. Therefore, the max flow problem can be written as follows (in the gross flow form):

$$\begin{aligned} \max x_{ts} & \\ \sum_w x_{vw} - x_{wv} &= 0 \\ x_{vw} &\leq u_{vw} \\ x_{vw} &\geq 0 \end{aligned}$$

In the dual problem, for each node v there is a conservation constraint. Besides, for each edge (v, w) there is a capacity constraint. Therefore, in the primal formulation, we have a variable z_v for each conservation constraint and a variable y_{vw} for each capacity constraint. The primal formulation is therefore:

$$\min \sum_{vw} u_{vw} y_{vw}$$
$$\begin{aligned} z_v - z_w + y_{vw} &\geq 0 \\ z_t - z_s + y_{ts} &\geq 1 \\ y_{vw} &\geq 0 \end{aligned}$$

We rewrite the first set of constraints as $y_{vw} \geq z_w - z_v$. Besides, the second constraint can be written as $z_t - z_s \geq 1$. This is because $y_{ts} = 0$ in any optimal solution. If $y_{ts} > 0$ in some optimal solution, the fact that $u_{ts} = \infty$ implies that $u_{ts} y_{ts} = \infty$ and therefore the optimal value is unbounded. This is impossible since the max flow problem is never infeasible (in particular, the zero flow is a feasible solution).

If we consider y_{vw} as the edge length of (v, w) and z_v as the distance from s to v, we can interpret the dual problem as follows: *Minimize the volume of the network by tuning the edge lengths, subject to the constraint that the distance from s to t is at least 1.* Here the volume of network is defined as the sum of edge volumes, which is the product of edge capacity u_{vw} and edge length y_{vw}.

Note that the optimal solution in this primal problem is at most the min-cut value of the network, as we can assign length 1 to the min-cut edges and 0 otherwise. This satisfies the s-t distance constraint (because any s-t path has to traverse some edge of a cut.) The value of this solution is the sum of min-cut edge capacities. By strong duality this implies max-flow $\leq$ min-cut. We now prove the other direction.

Denote z_v^*, y_{vw}^* as an optimal solution for the primal problem and similarly x_{vw}^* for the dual problem. Since z_v^* are distances, we can always rescale z_s^* to 0. Let $T = \{v | z_v^* \geq 1\}$. Note that $s \notin T$ and $t \in T$. Therefore T is a $s - t$ cut.

Now consider any edge (v, w) crossing the cut:

1. if $v \in T$ and $w \notin T$, then $z_v^* \geq 1$ and $z_w^* < 1$. Therefore, $z_v^* - z_w^* + y_{vw}^* \geq z_v^* - z_w^* > 0$. Therefore, the *constraint* for edge (v, w) in the primal problem is slack. By complementary slackness, the *variable* x_{vw}^* in the dual problem has to be tight, i.e., $x_{vw}^* = 0$.
2. if $v \notin T$ and $w \in T$, then $z_w^* \geq 1$ and $z_v^* < 1$. It follows that the *variable* $y_{vw}^* \geq z_w^* - z_v^* > 0$. Again, by complementary slackness, the *constraint* in the dual problem $x_{vw} \leq u_{vw}$ is tight. Therefore, $x_{vw}^* = u_{vw}$.

In other words, in a max flow, all edges entering T is saturated and all edge leaving T is empty. Therefore, in a max flow, the net flow entering T equals the cut value of T. Since the flow value equals the amount of net flow entering any $s - t$ cut, the max-flow value equals the cut value of T, which is at least the min-cut value. As a result, we have shown that max-flow $\geq$ min-cut, which completes the proof of the Max-Flow Min-Cut Theorem.