Lecture 9: Linear Programming Duality

Instructor: Dieter van Melkebeek Scribe: Choudhary and Sinha

We begin by looking at some problems that can be cast as linear programming problems. The ones that we are going to see have better algorithms but it will give us an idea of the range of linear programming applications. Then we will look at the concept of duality and weak and strong duality theorems. We will end with a study of the dual of Max-flow problem.

1 Examples of problems that can be cast as linear program

1.1 Max Flow

Recall the definition of network flow problem from Lecture 4. We have a directed graph $G(V, E)$ whose edges have capacity $c_e (e \in E)$. Two vertices of G are distinguished as source, s and sink, t. If the flow through each edge is x_e, then the Max Flow problem can be rephrased as a linear programming problem as follows:

$$\max \sum_{e=(s,u)\in E} x_e \quad s.t. \quad \begin{cases} \forall v \in V\backslash\{s,t\} & \sum_{e=(u,v)\in E} x_e = \sum_{e=(v,w)\in E} x_e \\ \forall e \in E & x_e \leq c_e \\ \forall e \in E & x_e \geq 0 \end{cases} \qquad (*)$$

We try to maximize the net outward flow from the source under the constraints that the flow going into any vertex (other than source and sink) is equal to the flow coming out of it, and the flow through any edge is non-negative and less than (or equal to) the capacity of the edge.

1.2 Shortest Path

We can also convert the shortest path problem for graphs without a negative-weight cycle to a linear programming problem. Our method is based on the Bellman-Ford algorithm for computing shortest path. As before, we have a graph $G(V, E)$ with edge weights $w(u, v)$ and two vertices - source s and destination t. We can write this as a linear program as follows:

$$\max \ x_t \quad s.t. \quad \begin{cases} \forall (u,v) \in E & x_v \leq x_u + w(u,v) \\ x_s = 0 \end{cases}$$

For a proof, it is easy to see that the constraints are satisfied if, for all vertices v, x_v is substituted by the length of a shortest path from s to v. So, $x_t \geq d$ where d is the shortest distance of t from s. For the converse, simply consider any path $(x, v_1, v_2, \ldots, v_k, t)$ of minimal length d from s to t. Then, $x_{v_1} \leq w(s, v_1)$, $x_{v_2} \leq x_{v_1} + w(v_1, v_2) \leq w(x, v_1) + w(v_1, v_2)$ and so on. So, $x_t \leq w(x, v_1) + w(v_1, v_2) + \cdots + w(v_k, t) = d$. So, the optimal value of x_t is equal to the length of a shortest path from s to t.

1.3 Min-Cost Flow

Min-cost flow refers to the problem of computing the minimum cost of sending more than a threshold amount of flow through a network. In this case, each edge of the network has a cost associated with it which is linearly proportional to the amount of flow through that edge. We can cast this into linear programming as follows:

$$\min \sum d_e x_e \qquad s.t. \quad \begin{cases} (*) \text{ is satisfied and} \\ v(f) \geq goal \end{cases}$$

2 Duality

Consider the following Linear Program once again:

$$\max \quad x_1 + x_2 \qquad s.t. \quad \begin{cases} -5x_1 + 2x_2 \leq 2 \ (1) \\ 4x_1 - x_2 \leq 8 \qquad (2) \\ 2x_1 + x_2 \leq 10 \quad (3) \\ x_1 \geq 0 \\ x_2 \geq 0 \end{cases}$$

The solution to the linear program, which is the maximum value of the objective function, is 8. Note that if we take linear combinations of the three equations judiciously we can get an upper bound on the objective function. For instance:

$\frac{1}{9}(1) + 0(2) + \frac{7}{9}(3) \Rightarrow$

$x_1 + x_2 \leq 8$

This gives an upper bound on the objective function which we sought to maximize. In fact this is a tight upper bound. Is it a coincidence? Duality explores these issues.

Let us associate a non negative variable with each of the three equations, say y_1, y_2, y_3.

Now $y_1(1) + y_2(2) + y_3(3) \Rightarrow$

$(-5y_1 + 4y_2 + 2y_3)x_1 + (y_1 - y_2 + y_3)x_2 \leq 2y_1 + 8y_2 + 10y_3$

Since $y_i s$ are non-negative we can get an upper bound on the objective function whenever:

$(-5y_1 + 4y_2 + 2y_3) \geq 1$

$(y_1 - y_2 + y_3) \geq 1$

and the upper bound is precisely $2y_1 + 8y_2 + 10y_3$

So we get a new linear program:

$$\min \quad 2y_1 + 8y_2 + 10y_3 \qquad s.t. \quad \begin{cases} -5y_1 + 4y_2 + 2y_3 \geq 1 \\ y_1 - y_2 + y_3 \geq 1 \\ y_1 \geq 0 \\ y_2 \geq 0 \\ y_2 \geq 0 \end{cases}$$

As we will see the minimum, value of this linear program will give us an optimum upper bound on the optimum value of the original linear program. We will soon see relations between them.The original linear program is called the primal and the new one is called the dual. It is not difficult to see that the dual of the dual is the primal itself. Note that the dual of the linear program (A, b, c), is $(-A^T, -c, -b)$. I.e., If the **primal** is:

$$\max \; c^T x \qquad s.t. \; \begin{cases} Ax \leq b \\ x \geq 0 \end{cases}$$

Then the **dual** is:

$$\min \; b^T y \qquad s.t. \; \begin{cases} A^T y \geq c \\ y \geq 0 \end{cases}$$

It immediately follows, from the definition of dual, that the dual of the dual is the primal itself. Also note that if some variable can take negative values in the primal then in the dual we get an equality constraint corresponding to that variable. Conversely whenever there is an equality constraint in the primal then the corresponding variable in the dual can take negative values as well.

Now we prove the following.

Theorem 1. Weak Duality Theorem: *For every feasible solution x of the primal and y of the dual: $c^T x \leq b^T y$*

Proof. Since y is a feasible solution of the dual, and $x \geq 0$, $A^T y \geq c \Rightarrow c^T \leq y^T A \Rightarrow c^T x \leq (y^T A)x \Rightarrow c^T x \leq y^T (Ax)$. Since x is a feasible solution of the primal, and $y \geq 0$, $Ax \leq b \Rightarrow y^T(Ax) \leq y^T b = b^T y$. □

Corollary 1. $MaxPrimal \leq MinDual$

When can the equality hold ? We need to look at the two inequalities we encountered in the proof of Theorem 1. The first one is $(c^T - y^T A)x \leq 0$. The other one is $y^T(Ax - b) \leq 0$. The two conditions together are known as complementary slackness. If both of them are equalities we get an equality. We will show that in "good" cases, this is indeed true. In preparation, we prove the following result.

Theorem 2. *One of the following holds:*

(1) Both Primal and Dual are infeasible
(2) Primal is unbounded and Dual is infeasible
(3) Primal is infeasible and Dual is unbounded
(4) Both are feasible and bounded

Proof. Note that Theorem 1 implies that every feasible solution implies an upper bound on the solutions of the primal. As such, in the event primal is unbounded dual has to be infeasible. Similarly every feasible solution of the primal implies a lower bound on the dual. As such the unboundedness of the dual implies the infeasibility of the primal. □

To exhibit that there are actually examples for each of the above cases is not difficult. (2) and (3) are easy to demonstrate. a linear program which is infeasible and consider its dual. For (2) we can take any unbounded primal and take its dual. In (3) we just need to take case (2) and interchange the primal and the dual. We will discuss case (4) when we show that the dual of MAX-FLOW is MIN-CUT.

Theorem 3. (Strong Duality Theorem) *If both primal and dual are bounded and feasible then, Max Primal = Min Dual.*

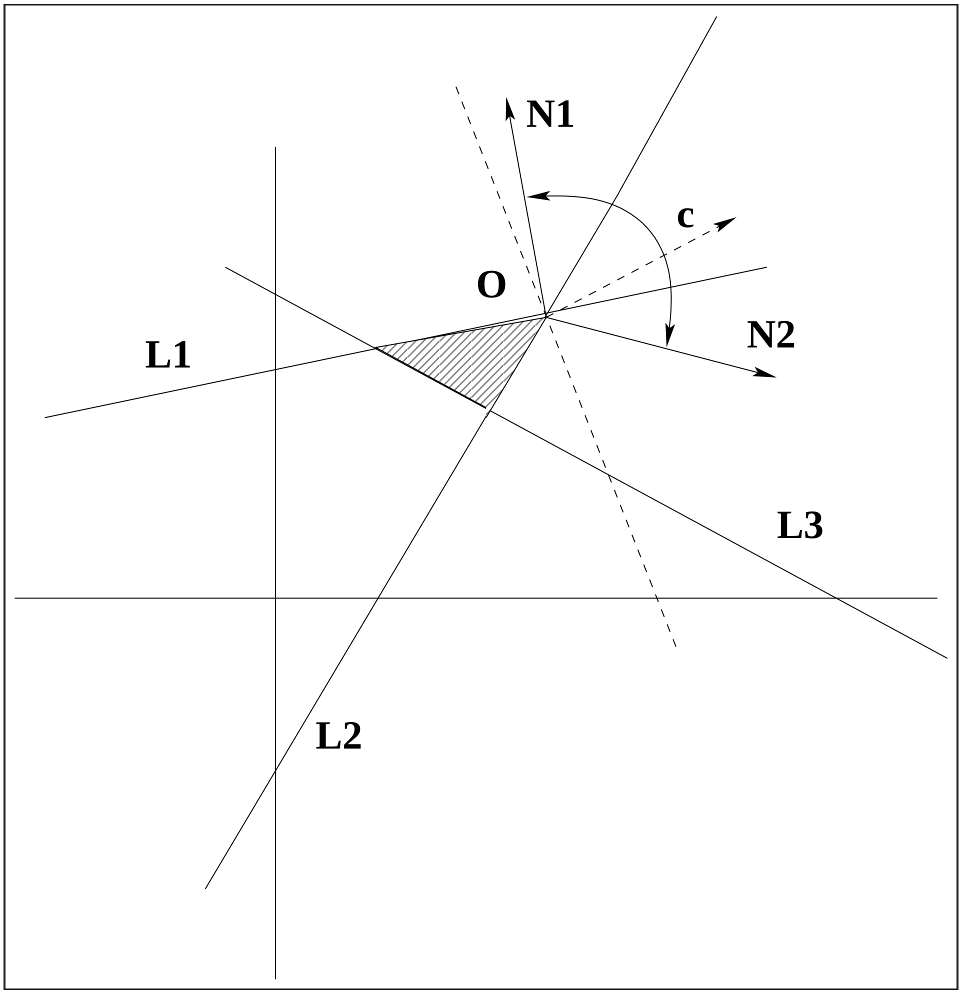

Figure 1: Outward Normals at the optimal vertex and c

Proof. As was observed, the equality $c^T x = b^T y$ holds for some feasible solution, x of Primal, and y of dual if and only if, $(c^T - y^T A)x = 0$ and $y^T(Ax - b) = 0$. Also note that both x and y are nonnegative. Therefore the equality occurs exactly when $(c_i - y^T A_{.,i})x_i = 0 \ \forall i 1 \leq i \leq m$ and $(b_j - A_{j,.}x)y_j = 0 \ \forall j 1 \leq j \leq n$ where A is an $m \times n$ matrix.

Now consider Figure 1. We claim that a vertex of the feasible region of a linear program is optimal (we have already argued in the previous lecture about optimality of vertices) exactly when the convex cone spanned by the outer normals to the binding constraints on that vertex contains c. In the figure, on vertex O, the binding constraints are L_1 and L_2. The corresponding outer normals are N_1 and N_2. Clearly the cone spanned by N_1 and N_2 spans c. We will not prove the claim; the geometric intuition of the claim however is that when we move from O to any other point in the feasible region, we move towards the inwards normals of the binding constraints on O and get a negative contribution overall in the inner-product with c, which we seek to maximize.

It follows from the above claim that given an optimal solution x of the primal, we can write c as:

$c = \sum_{(i:A_{i,\cdot}x=b_i)} y_i(A_{i,\cdot})^T - \sum_{(j:x_j=0)} z_j e_j$,

where each y_i and z_j is non-negative and e_j is the unit vector in the direction of x_j. The first term comes from the contribution of the outward normals on the binding constraints and the second one comes from the fact that we did not include the non-negativity constraints as binding constraints. In the event of any of them being binding at the optimal vertex, we have to consider positive contributions from the outwards normal in their direction, which are $-e_j s$

Now define $y_i{}^* = y_i$ if $A_{i,\cdot}x = b_i$ and 0 otherwise.

We look once again at the observation we made in the beginning about when the inequality can be an equality. Consider $(c_i - y^* A_{\cdot,i})x_i = 0$. When $x_i = 0$, we have nothing to prove. When it is not zero then the only contribution in c_i is from $\sum_{(i \ni A_{i,\cdot}x=b_i)} y^*{}_i(A_{i,\cdot})^T$. Therefore $(c_i - y^* A_{\cdot,i}) = 0$. Henceforth we get an equality.

For the second condition, $y^{*T}(Ax - b) = 0$, consider $(b_j - A_{j,\cdot}x)y_j$. By definition, $y_j{}^*$ is non-zero exactly when $b_j - A_{j,\cdot}x$ is zero. We again get an equality. Hence the result $(b^T)y^* = b^T x$. Therefore the Max Primal $\geq$ Min Dual. Combining it with Theorem 2 we get the result. □

3 The Dual of Max Flow

In this section we will study the dual of the Max Flow problem and see that the Max Flow - Min Cut theorem is a special case of the strong duality theorem. To get the dual, we have to consider linear combinations of the inequalities in $(*)$. We use the variable y_v for the vertex constraints and z_e for the edge capacity constraints. Now, the dual can be stated as:

$$\min \sum_{e \in E} c_e z_e \qquad s.t. \quad \forall (u,v) \in E \begin{cases} y_v - y_u + z_e \geq 0 & \text{if } u \neq s, v \neq t \\ y_v + z_e \geq 1 & \text{if } u = s, v \neq t \\ -y_u + z_e \geq 0 & \text{if } u \neq s, v = t \\ z_e \geq 1 & \text{if } u = s, v = t \end{cases}$$

If we add two new variables y_s and y_t with constraints that $y_s = 1$ and $y_t = 0$, all our inequalities can be written in the form $y_v - y_u + z_e \geq 0$. This is exactly a Min-Cut problem.

Let us denote the optimal value of the linear programming problem by OPT.

Claim 1. *$OPT \leq Min - Cut$.*

Proof. Given any cut (S, T), we set z_e to be 1 if e crosses the cut and zero otherwise. Also, y_v is set to 1 if v is in S and zero otherwise. It is easy to see that the constraints are satisfied in this case and the value of the objective is equal to the size of the cut. The claim follows. □

Claim 2. *$Min - Cut \leq OPT$.*

Proof. Consider any feasible solution y^*, z^*. Pick $\zeta \in (0, 1]$ uniformly at random. Define

$$S = \{v \in V | y_v \geq \zeta\}.$$

Note that this is a valid cut. Now, fix an edge e.

$$\begin{aligned} \Pr[e \text{ is in the cut}] &= \Pr(\zeta \in (y_v, y_u]) \\ &\leq \max(y_u, y_v, 0). \end{aligned}$$

Hence,

$$\begin{aligned} E(\text{capacity of the cut}) &= \sum c_e \Pr[e \text{ is in the cut}] \\ &\leq \sum c_e z_e. \end{aligned}$$

Thus, there must be a cut of capacity less that or equal to $\sum c_e z_e$. The claims follows. $\square$