Computability, Complexity, and Algorithms

Charles Brubaker and Lance Fortnow

Duality - (Udacity)

Introduction - (Udacity, Youtube)

Every linear program, it turns out, has a dual program which mirrors the behaviour of the original. In this lesson, we will examine this phenomenon to give us a chance to apply some of the knowledge we gained about linear programs, as well as to deepen our understanding of some other problems that we've already studied. See if you can guess which problems as the lesson goes along.

Bounding an LP - (Udacity)

I want to start off our discussion with a little exercise where we try to find an upper bound on the value of a linear program. We'll start with this linear program here,

Question

Consider the LP(*).

$$(*)\quad \begin{aligned} \max\ & 6x_1 - 2x_2 \\ \text{s.t } & 2x_1 + x_2 \le 2 \\ & -x_1 + 2x_2 \le 3 \\ & x_1, x_2 \ge 0 \end{aligned}$$

and we're going to take a linear combination of these inequality constraints to obtain a bound on the objective function.

Multiplying the first inequality by y_1 and the second by y_2, and adding them together, we obtain this inequality here. Note that it is important that the y's be non-negative to avoid reversing the inequality.

Question

Consider the LP(*). For $y_1, y_2 \ge 0$, observe

$$(*)\quad \begin{aligned} \max\ & 6x_1 - 2x_2 \\ \text{s.t } & 2x_1 + x_2 \le 2 \\ & -x_1 + 2x_2 \le 3 \\ & x_1, x_2 \ge 0 \end{aligned}$$

$$\begin{array}{rl} & y_1(2x_1 + x_2) \le y_1 2 \\ + & y_2(-x_1 + 2x_2) \le y_2 3 \\ \hline & (2y_1 - y_2)x_1 + (y_1 + 2y_2)x_2 \le 2y_1 + 3y_2 \end{array}$$

If we chose y_1 and y_2 such that $6 \le 2y_1 - y_2$ and $-2 \le y_1 + 2y_2$, then the objective function can be at most the left hand side of our new inequality, which can be at most the right.

Question

Consider the LP(*). For $y_1, y_2 \ge 0$, observe

$\max\ 6x_1 - 2x_2$

$$(*)\quad \begin{aligned} \text{s.t. } & 2x_1 + x_2 \le 2 \\ & -x_1 + 2x_2 \le 3 \\ & x_1, x_2 \ge 0 \end{aligned}$$

$$\begin{aligned} & y_1(2x_1 + x_2) \le y_1 2 \\ + \; & y_2(-x_1 + 2x_2) \le y_2 3 \\ \hline & \le (2y_1 - y_2)x_1 + (y_1 + 2y_2)x_2 \le 2y_1 + 3y_2 \end{aligned}$$

If
$6 \le 2y_1 - y_2$
$-2 \le y_1 + 2y_2$,
then
$6x_1 - 2x_2$

The quantity $2y_1 + 3y_2$ then becomes an upper bound on our objective function.

For this exercise, I want you to choose y_1 and y_2 to make this bound as tight as possible.

Question

Consider the LP(*). For $y_1, y_2 \ge 0$, observe

$$\max \; 6x_1 - 2x_2$$

$$(*)\quad \begin{aligned} \text{s.t. } & 2x_1 + x_2 \le 2 \\ & -x_1 + 2x_2 \le 3 \\ & x_1, x_2 \ge 0 \end{aligned}$$

$$\begin{aligned} & y_1(2x_1 + x_2) \le y_1 2 \\ + \; & y_2(-x_1 + 2x_2) \le y_2 3 \\ \hline & \le (2y_1 - y_2)x_1 + (y_1 + 2y_2)x_2 \le 2y_1 + 3y_2 \end{aligned}$$

If
$6 \le 2y_1 - y_2$
$-2 \le y_1 + 2y_2$,
then
$6x_1 - 2x_2$

Choose $y_1, y_2 \ge 0$ to make the bound as tight as possible. $y_1 = \boxed{}$ $y_2 = \boxed{}$

Dual Programs - (Udacity, Youtube)

Associated with every linear program is a so-called dual program, which is also a linear program. This definition is most elegant when stated in terms of the symmetric form. Indeed, now you see why this form gets the name symmetric.

Dual Programs

Primal	Dual
$\max c^T x$	$\min b^T y$
s.t. $Ax \le b$	s.t. $A^T y \ge c$
$x \ge 0$	$y \ge 0$

As we saw in the exercise, the dual program can be thought of as the problem of minimizing an upper bound on the primal. Note that for all feasible y, we have $b^T y$ is at most $y^T Ax$ using the contraint from the primal and the nonnegativity of y. And this is at most $c^T x$, using the contraint from the dual and nonnegativity of x.

Dual Programs

Primal	Dual
$\max c^T x$	$\min b^T y$
s.t. $Ax \le b$	s.t. $A^T y \ge c$
$x \ge 0$	$y \ge 0$

$\forall$ feasible y, $b^T y = y^T b \ge y^T A x \ge c^T x$

$(b \ge Ax, y^T \ge 0)$ $(y^T A \ge c^T, x \ge 0)$

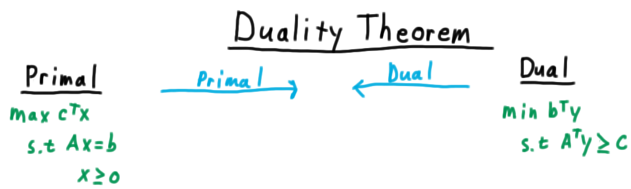
Weak Duality Lemma: If x is feasible for the primal and y is feasible for the dual, then $c^T x \leq b^T y$.

In fact, we just proved the Weak Duality Lemma, which states that if x is feasible for the primal problem and y is feasible for the dual problem, then $c^T x$ is at most $b^T y$.

Another thing to note here, is that if your primal problem isn't in this exact form, you can always convert it, then look at the corresponding dual and simplify. Often, however, it is easier just to remember that the dual is the problem of bounding the primal as tightly as possible. For instance, if we change the inequality in the primal to equality, then we can proceed by the same argument, only this first inequality becomes an equality, and I don't have to rely on y being non-negative. Everything else is the same.

Duality Theorem - (Udacity, Youtube)

Here is the picture so far. We have primal programs over here trying to be maximized, and we have our dual program here, trying to be minimized, and the obvious question is "Do they ever meet?"



Well, the answer is ``Yes, they always do." More precisely, we state this as follows in the Duality Theorm.

> *If either the primal problem or the dual has a feasible optimal solution, then so does the other, and the optimal objective values are equal. If either problem has an unbounded objective value, then the other is infeasible.*

We'll start the proof by showing the second part. Suppose the primal is unbounded and y is a feasible for the dual. (We're going to show that both of these can't be true.) By weak duality, $b^T y \geq c^T x$ for all feasible x. Since the primal is unbounded, however, I can find x that gives me a value as high as I want. Whatever, the value of $b^T y$ is, I can find a feasible x such that $c^T x$ is larger, which creates a contradiction. The case where the dual is unbounded, is analogous.

Now, we return to the first part: "If either the primal problem or the dual has a feasible optimal solution, then so does the other, and the optimal objective values are equal." Let's start with the primal having a finite optimal solution. From this it follows that there is a finite *basic* optimal solution by the Fundamental Theorem of LP. Let's let the basis be the first m columns of the matrix A as usual and divide x and c up accordingly. (As usual B stands for basic here)

Recall then from the simplex algorithm that the vector r_D which represented the effects of moving along on of the directions in x_D had to be nonpositive. I.e.

$$0 \geq r_D^T = c_D^T - c_B^T B^{-1} D.$$

Otherwise, this basic solution wasn't optimal. Now, we're going to actually construct a solution for the dual. Letting

$$y^T = c_B^T B^{-1}$$

, we have that $y^T D \geq c_D^T$ from the nonpositivity of r. Therefore,

$$y^T A = [y^T B, y^T D] \geq [y^T B, c_B^T B^{-1} D] \geq [c_B^T, c_D^T] = c^T.$$

We conclude that y is feasible for the dual.

Moreover, by substitution, we see that

$$y^T b = c_B^T B^{-1} b = c_B^T x_B.$$

where x is the basic optimal solution. By weak duality, this is the best we can do, so both y also is optimal.

Duality Theorem

Primal

$\max c^T x$
s.t $Ax = b$
$x \geq 0$

Primal — Dual

Dual

$\min b^T y$
s.t $A^T y \geq c$

Duality Theorem: If either the primal problem or the dual has a finite optimal solution, then so does the other, and the optimum objective values are equal. If either problem has an unbounded objective value, then the other is infeasible.

Dual Optimal Solutions - (Udacity, Youtube)

With this proof, we actually have shown something even stronger than the Duality theorem we set out to show, because we have actually given a way to determine a dual optimal solution. We'll start with the linear program in standard form, as usual, and we'll let the columns of the matrix Bform an optimal basis, meaning that it generates an optimal basic feasible solution.
Then y^T defined at $c_B^T B^{-1}$ is an optimal solution to the dual problem by our previous argument. Moreover, the optimal values are equal.

Theorem:
Let the linear program

$\max c^T x$
s.t $Ax = b$
$x \geq 0$

have an optimal basic feasible solution corresponding to the basis formed by the columns of B. (Note $x_B = B^{-1}b$).
Then $y^T = c_B^T B^{-1}$ is an optimal solution to the dual problem

$\min y^T b$
s.t $y^T A \geq c^T$

Moreover, the optimum values of the problems are equal.

Dual Solution Calculation - (Udacity)

Let's do an exercise on this idea of a dual optimal solution. Given that x as shown here is an optimal basic solution to the linear program below, find the dual optimal solution.

We'll let y_1 correspond to this first constraint and y_2 correspond to the second.

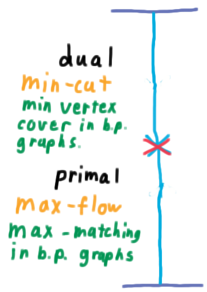
Question

Given that $x_1 = 2, x_2 = 2, x_3 = 0, x_4 = 0$ is an optimal solution to the linear program below find an optimal solution for the dual.

$\max\ x_1 + 4x_2 - 2x_3 - x_4$
s.t. $3x_1 + x_3 + 2x_4 = 6 \leftarrow y_1$
$-x_1 + 2x_2 - x_3 - x_4 = 2 \leftarrow y_2$
$x_i \geq 0$

Duality of Max Matching - (Udacity, Youtube)

By now, we have seen this picture several times, where we one quantity that we are trying to maximize and another which serves as an upper bound that we are trying to minimize, and … luckily the two meet at some point that is optimal for both. We have just seen this with our primal and dual linear programs
but we saw it earlier in the semester as well with our max-flow/min-cut problem
and also with our max-matching and vertex cover problems in bipartite graphs.



It's natural to ask, are these phenomena all related? Well, yes they are and probably the easiest way to see that is to realize that all of these can be characterized as linear programs and their duals.

Let's take a look at the duality of maximum matching in bipartite graphs first.
We'll let the variable x_{ij} indicate whether x_{ij} should be included in the matching. Then as a linear programming the problem becomes to maximize the number of matched edges subject to the constraints that no vertex in L can be matched more than once and no vertex in R can be matched more than once. Of course, we can't have negatively matched edges.

Duality of Max Matching

Let x_{ij} indicate whether (i,j) is included in the matching.

$$\max \sum_{i\in L, j\in R} x_{ij}$$

$$\text{s.t.} \quad \sum_{i\in L} x_{ij} \le 1 \quad \forall j \in R$$

$$\sum_{j\in R} x_{ij} \le 1 \quad \forall i \in L$$

$$x_{ij} \ge 0$$

To build the dual program, we let y_i and y_j be the variables corresponding to these contraints, and we want to minimize their sum because the constraint vector here is just all ones.

Duality of Max Matching

Let x_{ij} indicate whether (i,j) is included in the matching.

$$\max \sum_{i\in L, j\in R} x_{ij} \qquad\qquad \min \sum_{i\in L} y_i + \sum_{j\in R} y_j$$

$$\text{s.t.} \quad \sum_{i\in L} x_{ij} \le 1 \quad \forall j \in R \quad y_j$$

$$\sum_{j\in R} x_{ij} \le 1 \quad \forall i \in L \quad y_i$$

$$x_{ij} \ge 0$$

For the contraints, observe that the coefficients in the objective function are 1 and that any x_{ij} appears once in the equation for i and once in the equation for j.

Hence $y_i + y_j \geq 1$. And of course y_i and y_j can't be negative.

Duality of Max Matching

Let x_{ij} indicate whether (i,j) is included in the matching.

$$\max \sum_{i\in L, j\in R} x_{ij} \qquad\qquad \min \sum_{i\in L} y_i + \sum_{j\in R} y_j$$

$$\text{s.t.} \quad \sum_{i\in L} x_{ij} \leq 1 \quad \forall\, j\in R \quad y_j \qquad\qquad \text{s.t.} \quad y_i + y_j \leq 1 \quad \forall\, \begin{smallmatrix} i\in L \\ j\in R \end{smallmatrix}$$

$$\sum_{j\in R} x_{ij} \leq 1 \quad \forall\, i\in L \quad y_i \qquad\qquad y_i, y_j \geq 0 \quad \forall\, \begin{smallmatrix} i\in L \\ j\in R \end{smallmatrix}$$

$$x_{ij} \geq 0$$

The interpretation here is straightforward: vertex i is in the cover if and only if $y_i = 1$ and similarly vertex j is in the cover if and only if $y_j = 1$.

Every edge must have at least one vertex in the cover and we are trying to minimize the size of the cover.

So we have just seen how maximum bipartite matching can be expressed as a linear program and it's dual also turned out to have a natural interpretation as the vertex cover problem. This is really neat. Every decision problem in P can be converted to a linear program ultimately, just because linear programming is P-complete, but not every conversion will result in variables and a dual program that have such intuitive interpretations. When this happens, it often gives a way to gain deeper insight into a problem and its structure.

As you might have guessed, this happens also for the max-flow/min-cut problem and we'll explore that next.

Duality of Max Flow - (Udacity, Youtube)

For completeness, we'll go ahead and explore the duality in the maximum flow problem as well. We can cast it as a linear programming problem by letting f_{uv} be the flow and letting c_{uv} be the capacity across an edge (u, v).

Our goal is to maximize the flow out of the source s subject to the conservation of flow constraint and the capacity contraint. Of course, flows must be nonnegative as well.

Duality of Max-flow

Let f_{uv} be the flow and let c_{uv} be the capacity along $(u,v) \in E$.

$$\max \sum_{v:(s,v)\in E} f_{sv}$$

$$\text{s.t.} \quad \sum_{v:(v,u)\in E} f_{vu} - \sum_{v:(u,v)} f_{uv} = 0 \quad \forall\, u \in V - \{s,t\} \quad \text{(Conservation of flow)}$$

$$f_{uv} \leq c_{uv} \qquad \forall\, (u,v) \in E \quad \text{(capacity constraints)}$$

$$f_{uv} \geq 0$$

To express the dual we'll use y_u for conservation contraint at vertex u and y_{uv} for capacity contraint at edge (u, v). Two subscripts mean a capacity contraint, one subscript means a conservation constraint.

The dual problem is to minimize the sum over all edges of $c_{uv}y_{uv}$. Note that the y_u's have no role in the objective function because their coefficients are zero.

Duality of Max-flow

Let f_{uv} be the flow and let c_{uv} be the capacity along $(u,v) \in E$.

$$\max \sum_{v:(s,v)\in E} f_{sv} \qquad\qquad \min \sum_{u,v \in E} y_{uv} c_{uv}$$

$$\text{s.t.} \quad \sum_{v:(v,u)\in E} f_{vu} - \sum_{v:(u,v)} f_{uv} = 0 \qquad y_u$$

$$f_{uv} \le c_{uv} \qquad y_{uv}$$

$$f_{uv} \ge 0$$

The constraints for the dual involve several cases. We'll consider first those arrising from the objective function coefficients being one for edges out of the source. The flows appear onces in the capacity constraint and once in the conservation equation for the receiving vertex.

Duality of Max-flow

Let f_{uv} be the flow and let c_{uv} be the capacity along $(u,v) \in E$.

The case for edges going into the sink is analgous. The flow is present in the capacity constraint and in the conservation of flow equation for the sending vertex. These must be at least one because the objective function coefficient is zero.

Duality of Max-flow

Let f_{uv} be the flow and let c_{uv} be the capacity along $(u,v) \in E$.

$$\max \sum_{v:(s,v)\in E} f_{sv} \qquad\qquad \min \sum_{u,v \in E} y_{uv} c_{uv}$$

$$\text{s.t.} \quad \sum_{v:(v,u)\in E} f_{vu} - \sum_{v:(u,v)} f_{uv} = 0 \qquad y_u$$

$$f_{uv} \le c_{uv} \qquad y_{uv}$$

$$f_{uv} \ge 0$$

$$\begin{aligned}
\text{s.t.} \quad & y_{sv} + y_v \ge 1 \quad \forall (s,v) \in E \\
& y_{vt} - y_v \ge 0 \quad \forall (v,t) \in E \\
& y_{uv} + y_v - y_u \ge 0 \quad \text{all other edges} \\
& y_{uv} \ge 0
\end{aligned}$$

For all other edges, the the flow appear in the capacity constraint and BOTH conservation of flow equations. Again, the coefficient in the objective function is zero so that becomes the contraint. And these dual variables have to be nonnegative.

The interpretation of these dual variable can be a little tricky so, I'm going to rearrange the constraints to isolate the capacity variables on the left-hand side.

Duality of Max-flow

Let f_{uv} be the flow and let c_{uv} be the capacity along $(u,v) \in E$.

$$\max \sum_{v:(s,v)\in E} f_{sv} \qquad\qquad \min \sum_{u,v \in E} y_{uv} c_{uv}$$

$$\text{s.t.} \quad y_u \rightarrow \sum_{v:(v,u)\in E} f_{vu} - \sum_{v:(u,v)} f_{uv} = 0$$

$$y_{uv} \rightarrow f_{uv} \le c_{uv}$$

$$f_{uv} \ge 0$$

$$\text{s.t.} \quad y_{sv} \ge 1 - y_v \quad \forall (s,v)\in E$$

$$y_{vt} \ge y_v - 0 \quad \forall (v,t)\in E$$

$$y_{uv} \ge y_u - y_v \quad \text{all other edges}$$

$$y_{uv} \ge 0$$

This makes it a little easier to see what is going on. Actually, I think this would make a good exercise.

Interpretation of y - (Udacity)

Suppose that y is a basic optimal feasible solution for the given LP. Which statements are part of an interpretation of y as an s-t cut, say (A, B)?

Conclusion - (Udacity, Youtube)

In this lesson, we defined the dual of a linear program and showed how this dual program can be see as the problem of making a certain kind of bound on the primal program as tight as possible. Then, we saw how maximum flow and maximum bipartite matching can be expressed as linear programs and how the minimun s-t cut and vertex cover problems were their duals.