Optimization I | **Summer 2011** | **Lecture 8**

Linear Programming Duality

1 Duality recap

We motivated the dual of a linear program by thinking about the best possible lower bound on the optimal value we can achieve. We found the following form for a primal LP and its dual:

$$
\text{Primal:}\quad
\begin{array}{lccccc|c|l}
\min & c_1x_1 & +\dots & +c_jx_j & +\dots & +c_nx_n & & \\
\text{s.t.} & a_{11}x_1 & +\dots & +a_{1j}x_j & +\dots & +a_{1n}x_n & \geq \text{ or } \leq \text{ or } = & b_1 \\
 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
 & a_{i1}x_1 & +\dots & +a_{ij}x_j & +\dots & +a_{in}x_n & & b_i \\
 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
 & a_{m1}x_1 & +\dots & +a_{mj}x_j & +\dots & +a_{mn}x_n & & b_m
\end{array}
$$

$$\boxed{\text{sign constraints on } x_1, \dots, x_n}$$

$$
\text{Dual:}\quad
\begin{array}{lccccc|c|l}
\max & b_1y_1 & +\dots & +b_iy_i & +\dots & +b_my_m & & \\
\text{s.t.} & a_{11}y_1 & +\dots & +a_{i1}y_i & +\dots & +a_{m1}y_m & \geq \text{ or } \leq \text{ or } = & c_1 \\
 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
 & a_{1j}y_1 & +\dots & +a_{ij}y_i & +\dots & +a_{mj}y_m & & c_j \\
 & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
 & a_{1n}y_1 & +\dots & +a_{in}y_i & +\dots & +a_{mn}y_m & & c_n
\end{array}
$$

$$\boxed{\text{sign constraints on } y_1, \dots, y_m}$$

Note that we can also write the primal and dual in matrix notation as

$$
\text{Primal:}\quad
\begin{array}{llcl}
\min & c^Tx & & \\
\text{s.t.} & Ax & \boxed{\begin{array}{c}\geq \text{ or } \leq \text{ or } = \\ \text{(different for} \\ \text{each row)}\end{array}} & \text{b}
\end{array}
\qquad
\text{Dual:}\quad
\begin{array}{llcl}
\min & b^Ty & & \\
\text{s.t.} & A^Ty & \boxed{\begin{array}{c}\geq \text{ or } \leq \text{ or } = \\ \text{(different for} \\ \text{each row)}\end{array}} & \text{c}
\end{array}
$$

$$\boxed{\text{sign constraints on } x_1, \dots, x_n} \qquad \boxed{\text{sign constraints on } y_1, \dots, y_m}$$

The signs of the constraints in the dual are determined by the sign constraint of the corresponding variables in the primal:

$$\text{Primal variable } x_j \geq 0 \Rightarrow \text{ Dual constraint } \leq c_j,$$

and if $x_j \leq 0$, then the dual constraint is $\geq c_j$, and if x_j is free then the dual constrain is $= c_j$.

The sign constraints of the variables in the dual are determined by the sign of the corresponding constraint in the primal:

$$\text{Primal constraint } \geq b_i \Rightarrow \text{ Dual variable } y_i \geq 0,$$

and if the primal constraint requires $\leq b_i$, we get $y_i \leq 0$ and if the primal constraint is $= b_i$, then y_i is free.

Theorem 1 (Weak duality). *If x is a feasible solution to the primal LP and y is a feasible solution to the dual LP, then $c^T x \geq b^T y$.*

Proof. We have

$$c^T x = \sum_{j=1} c_j x_j \geq \sum_{j=1}^{n} (\sum_{i=1}^{m} a_{ij} y_i) x_j = \sum_{i=1}^{m} y_i \sum_{j=1}^{n} a_{ij} x_j \geq \sum_{i=1}^{m} y_i b_i = b^T y.$$

The first inequality follows from how we determined the sign of the constraints in the dual: we made sure that

$$c_j x_j \geq \sum_{i=1}^{m} a_{ij} y_i x_j.$$

The second inequality follows from how we determined the sign of the variables in the dual: we made sure that

$$y_i \sum_{j=1}^{n} a_{ij} x_j \geq y_i b_i.$$

□

Weak duality tells us the following: We want to minimize the primal objective, but as soon as we find a feasible solution to the dual LP with some objective z, then we know the primal objective cannot be less than that. On the other hand, we want to maximize the dual objective, but as soon as we find a feasible solution to the primal LP with some objective z, then we know the dual objective cannot be higher than that.

Last time we discussed the following theorem:

Theorem 2. *If we transform a minimization linear program into another equivalent minimization problem, then their duals are also equivalent.*

Using that, we can prove the following theorem:

Theorem 3. *"The dual of the dual is the primal": If we transform the dual linear program into an equivalent minimization problem, and take its dual, then we obtain a problem that is equivalent to the original problem.*

Proof. By Theorem 2, it is enough to prove it if the primal LP is in standard form. We then have

$$\begin{array}{c} \min c^T x \\ \text{s.t.} Ax = b \\ x \geq 0 \end{array} \overset{\text{take dual}}{\rightarrow} \begin{array}{c} \max b^T y \\ \text{s.t.} A^t y \leq c \\ \end{array} \overset{\text{transform}}{\rightarrow} \begin{array}{c} \min -b^T y \\ \text{s.t.} -A^T y \geq -c \\ \end{array} \overset{\text{take dual}}{\rightarrow} \begin{array}{c} \max -c^T x \\ \text{s.t.} -Ax = -b \\ x \geq 0 \end{array} \overset{\text{transform}}{\rightarrow} \begin{array}{c} \min c^T x \\ \text{s.t.} Ax = b \\ x \geq 0 \end{array}$$

□

Now, a major result in linear programming says that if both the primal LP and the dual LP have a feasible solution, then their optimal values are equal. This is called strong duality.

Theorem 4 (Strong duality). *If a linear program has an optimal solution, then so does its dual and the respective objective values are equal.*

The proof of this can be done in two ways: there is a direct proof using separating hyperplanes which we will not go into, but you can find it in the book in Section 4.7. We repeat the proof from last time:

Proof. Convert the primal into standard form, i.e. $\min c^T x$ subject to $Ax = b, x \geq 0$. The dual is $\max b^T y$ subject to $A^T y \leq c$, which we note is the same as $\max y^T b$ subject to $c^T - y^T A \geq 0$. The simplex method will terminate with an optimal solution in which all reduced costs are non-negative, i.e. $c^T - c_B^T \mathbf{B}^{-1} A \geq 0$. So $y^T = c_B^T \mathbf{B}^{-1}$ is a feasible dual solution, and it has dual objective value $y^T b = c_B^T \mathbf{B}^{-1} b$, which is the same as the primal objective value. □

We now know that the following are the only possibilities for the primal and dual:

- both have the same finite optimal value,
- one is unbounded and the other infeasible,
- both are infeasible.

2 An Application of Duality: Shortest Path

We are given a directed graph $G = (V, E)$. Let $|V| = m, |E| = n$. We assume the nodes are labelled $v_1, \ldots, v_m$, and the edges are labelled $e_1, \ldots, e_n$. There is a cost/length $c_j \geq 0$ associated with each edge $e_j \in E$. Suppose we want to find the shortest path from v_1 to v_m.

To formulate the problem of finding the shortest path as a linear program, we think of sending a unit of flow out of v_1 to v_m. For every other node, the flow into the node must be equal to the flow out of the node. We can represent this as follows: We let $x = [x_1, \ldots, x_n]^T$ where x_j is the amount of flow on edge e_j. We define an $m \times n$ matrix A which has the following entries:

$$a_{ij} = \begin{cases} 1 & \text{if } e_j \text{ leaves } v_i \\ -1 & \text{if } e_j \text{ enters } v_i \\ 0 & \text{otherwise} \end{cases}$$

Let a_i^T denote the i-th row of A, then $a_i^T x$ is equal to $\sum_{j:e_j \text{ leaves } v_i} x_j - \sum_{j:e_j \text{ enters } v_i} x_j$. We define

$$b_i = \begin{cases} 1 & \text{if } i = 1 \\ -1 & \text{if } i = m \\ 0 & \text{otherwise} \end{cases}$$

So we want to solve a problem in standard form: $\min c^T x$ subject to $Ax = b, x \geq 0$.

Note that we are interested in finding a **integral solution** to this problem, and we now allow the linear program to find fractional solutions as well. So we are not guaranteed that the optimal value of the LP is not strictly lower than the length of the shortest path from v_1 to v_m (why can it not be strictly higher?). Later in the course, you will learn that, in fact, this LP has integral basic feasible solutions, but for today, we will use duality to prove that the optimal value of this LP is actually equal to the length of the shortest path.

The dual of this LP is $\max b^T y$ subject to $A^T y \leq c$. Since b_i is 0 everywhere except for $i = 1, m$, we get that $b^T y = y_1 - y_m$. The j-th constraint is equal to $\sum_{i=1}^n a_{ij} y_i \leq c_j$. Suppose edge $e_j = (v_k, v_\ell)$, and let $c_{(v_k, v_\ell)} = c_j$. Then the j-th constraint is equal to $y_k - y_\ell \leq c_{(v_k, v_\ell)}$. So the dual LP is the following:

$$\begin{array}{ll} \max & y_1 - y_m \\ \text{s.t.} & y_k - y_\ell \leq c_{(v_k, v_\ell)} \text{ for every edge } (v_k, v_\ell) \in E \end{array}$$

Let $d(v_k, v_m)$ be the length of the shortest path from v_k to v_m. Then, we can obtain a feasible dual solution by setting $y_k = d(v_k, v_m)$: we need to verify that

$$y_k \leq c_{(v_k, v_\ell)} + y_\ell.$$

But of course this is true: the length of the shortest path from v_k to v_m cannot be longer than if we first go from v_k to v_ℓ and then take the shortest path from there to v_m. So we have found a dual feasible solution, and its objective value is $d(v_1, v_m) - d(v_m, v_m) = d(v_1, v_m)$.

Since we also have a primal solution with the same objective value (set $x_j = 1$ for the edges on the shortest path from v_1 to v_m), by strong duality, both solutions are optimal.

3 Interpreting duals

In the example we saw, the dual had a very nice interpretation, but it seems hard to interpret duals in general. Here is how I always think about primals and duals:

I think of primal LPs as "I am doing some activity (for example, making a product, going from A to B) and I am trying to minimize the cost of doing it". Then, I think of the dual LP as "a company wants to offer me the service of doing the activity for me, and wants to maximize the amount of profit they can make".

The company cannot raise its prices arbitrarily, because if it becomes more expensive than if I do it myself, then I will not buy their service. So, in the example of the shortest path, the company may be a bus company that can take me from v_1 to v_m. In fact, they have bus services all over the graph: I can leave from any location and get off in any other location. They have a pricing scheme that gives a price y_k for a ticket from location v_k, and if you get off at an intermediate location v_ℓ then you don't have to pay the full price, but instead, you pay $y_k - y_\ell$. The company tries to maximize the amount they get from people going from v_1 to v_m, and they cannot make any journey on the network more expensive than if a customer would go there by other means, so $y_k - y_\ell \leq c_{(v_k,v_\ell)}$.

Another example, given in the book, is the diet problem: I want to satisfy all my nutritional requirements and minimize the cost of food. Some company sells pills for each of the nutrients, and wants to maximize the amount of money they get for selling me the pills that satisfy all my requirements. However, they cannot make the pills more expensive than if I just buy my groceries and satisfy the nutritional requirements that way.

Finally, we can think of dual variables as "marginal costs" – more about that when we talk about sensitivity analysis.

4 Complementary Slackness

If x and y are optimal solutions to the primal and dual problem respectively, then in our proof of weak duality, we know that all inequalities have to hold *at equality*. This means that we know the following:

$$c_j x_j = \sum_{j=1}^{n} a_{ij} y_i x_j \quad \Leftrightarrow \quad (\sum_{i=1}^{m} a_{ij} y_i - c_j) x_j = 0$$

$$y_i \sum_{j=1}^{n} a_{ij} x_j = y_i b_i \quad \Leftrightarrow \quad (\sum_{j=1}^{n} a_{ij} x_j - b_i) y_i = 0$$

So, we need, that if $x_j \neq 0$, then $\sum_{i=1}^{m} a_{ij} y_i = c_j$, and similarly, if $y_i \neq 0$ then $\sum_{j=1}^{n} a_{ij} x_j = b_i$.

We thus have the following theorem:

Theorem 5. *Suppose the primal and dual LP are both feasible. Let x be a solution for the primal LP, y a solution for the dual LP. Then x is optimal for the primal and y is optimal for the dual, if and only if*

- *x is feasible for the primal LP,*
- *y is feasible for the dual LP,*
- *x and y satisfy complementary slackness:*
  - *If $x_j \neq 0$, then the j-th constraint of the dual holds at equality, and*
  - *if $y_i \neq 0$, then the i-th constraint of the primal holds at equality.*

Now, suppose we have a primal problem in standard form. Then the constraints of the primal are all equality constraints, so the second type of complementary slackness conditions hold for any pair of feasible solutions x and y. The first type of complementary slackness conditions require that, if $x_j > 0$ then the i-th dual constraint holds at equality, i.e., $c_j - \sum_{i=1}^{m} a_{ij} y_i = 0$. In other words, if $x_j > 0$ then $c_j - y^T A_j = 0$.

Now, suppose x is a basic feasible solution, and let B be the indices of the basic variables. Let $y^T = c_B^T \mathbf{B}^{-1}$ as in the proof of the Strong Duality theorem. Then complementary slackness requires that if $x_j > 0$ then $c_j - c_B^T \mathbf{B}^{-1} A_j = 0$, i.e., the reduced cost of j must be zero if x_j is basic. But this is exactly what the Simplex Method does! Note that y is not feasible until the final iteration, when $c^T - y^T A = c^T - c_B^T \mathbf{B}^{-1} A_j \geq 0$.

We can thus conclude the following:

> *The Simplex Method maintains a solution x for the primal LP and y for the dual LP that satisfy complementary slackness. The solution x is feasible for the primal in all iterations, and the Simplex Method works toward making the solution y dual feasible.*

The Simplex Method is therefore called a **primal algorithm** – it always maintains a primal feasibility and works toward dual feasibility. We can also develop an algorithm which maintains two solutions x and y that satisfy complementary slackness and which maintains dual feasibility and works toward primal feasibility – such an algorithm is called a **dual algorithm**, and we will see the Dual Simplex Method next week. Finally, we could also have a feasible primal and dual solution and work toward achieving complementary slackness. This is what happens in interior point methods, which you will see later in the course.