

Duality Notes

18.310, Fall 2005, Prof. Peter Shor

One of the most important aspects of linear programming is the duality theorem. Let's consider a linear program in the standard form we talked about last time.

$$\begin{aligned} \max \sum_k v_k x_k \quad & \text{subject to} \\ \sum_k A_{jk} x_k \leq b_j \quad & \forall j \quad \text{and} \\ x_k \geq 0 \quad & \forall k \end{aligned}$$

Now, how might we come up with an upper bound for the objective function $\sum_k v_k x_k$?

Recall the example we did at the end of last class, where we took a sum of some multiple of the inequalities and found that this gave us an upper bound on the objective function? We'll use the same trick, in greater abstraction.

So what we'll try to do is to take all these inequalities

$$\sum_k A_{jk} x_k \leq b_j,$$

multiply each of them by some value y_j , and add all the equations up. We want to choose these values of y_j so that this is a bound on x_k . What happens when we sum these inequalities? We get

$$\sum_j y_j \sum_k A_{jk} x_k \leq \sum_j b_j y_j.$$

We need that all the $y_j \geq 0$, because if we multiplied by a negative y_j , we would reverse the sign of the inequality.

However, if all the y_j are positive, then the above equation must hold for any feasible point $\{x_1, x_2, \dots, x_n\}$, since a feasible point satisfies all the inequalities. How can we make sure that the resulting sum is a bound on $\sum v_k x_k$? We need to make sure that

$$\sum_k v_k x_k \leq \sum_j \sum_k y_j A_{jk} x_k$$

for all feasible points x_k . One way to ensure this is to make sure that the coefficient on x_k on the left is less than the coefficient on x_k on the right. Thus, we would need

$$v_k \leq \sum_j y_j A_{jk} \quad \forall k.$$

Note that here we are using the fact that $x_k \geq 0$.

Any set of y_j satisfying the above equations gives us an upper bound $\sum_j b_j y_j$ on the objective function. How do we get the best bound on the objective function? We need to minimize $\sum_j b_j y_j$.

When we minimize $\sum_j b_j y_j$, we get another linear program, although this one isn't in our standard form. We have the linear program

$$\begin{aligned} \min \sum_j b_j y_j \quad & \text{subject to} \\ \sum_j A_{jk} y_j & \geq v_k \quad \forall k \quad \text{and} \\ y_j & \geq 0 \quad \forall j. \end{aligned}$$

What have we done? The right-hand side of the constraints, variables b_j , have switched places with the constants in the objective, variables v_k . We've effectively transposed the matrix A_{jk} , so instead of $\sum_k A_{jk} x_k$ we have $\sum_j A_{jk} y_j$. We've swapped min for max, and \leq for \geq .

Now, we've got a new linear program, which we call the *dual*. We call the original linear program the *primal*. The dual of the dual will be the primal. We have seen that if we have two feasible sets of variables x_k and y_j (for the primal and dual, respectively), the objective function of the dual is always at least the objective function of the primal. This shows that the optimal value of the dual is always at least the optimal value of the primal.

Now, a truly amazing fact about linear programming, and the source of a lot of its effectiveness, is that these two values are equal. This is known as **the duality theorem of linear programming**. We will prove it in a little bit. The theorem actually says a bit more, about infeasible and unbounded linear programs. We won't prove this in class, but it's not hard to generalize the proof to handle these cases.

Theorem: (The duality theorem for linear programming)

If both the primal and the dual are feasible and unbounded, the optimal value of the primal is equal to the optimal value of the dual. The primal is infeasible if the dual is unbounded. The dual is infeasible if the primal is unbounded.

It is possible for both the dual and the primal to be infeasible. One can get this situation by combining a linear program with an infeasible primal and unbounded dual by one with an unbounded primal and infeasible dual.

There is a recipe for taking a linear program (whether or not it's in standard form) and finding its dual. Inequalities turn into variables y_j with the constraint $y_j \geq 0$. Equations turn into variables y_j which can be either positive or negative, and so

on. However, the recipe is complicated enough that if you memorize it, you're likely to forget it or remember it wrong when you have to use it, especially if you try to memorize the recipe for general linear programs and not just standard form. The best way of remembering how to find the dual is remembering the proof above that the dual optimum is at least the primal optimum, and reproducing it whenever you need it.

We now want to prove the duality theorem. That is, we want to show that the optimal value of the primal is equal to the optimal value of the dual. We've shown one inequality already (this was the easy one), so now we need to show the other inequality. We will be able to do this by looking at the final tableau in the simplex algorithm, and show that the objective function it gives is not only a feasible solution for the primal but also a feasible solution for the dual.

So let's recall the simplex algorithm. We started by taking all the inequalities that weren't of the form $x_k \geq 0$ and adding a slack variable s_j to them, so that we get inequalities

$$\begin{aligned} x_k &\geq 0 & \forall k \\ s_j &\geq 0 & \forall j \end{aligned}$$

and equalities of the form

$$\sum_k A_k x_k = b_k.$$

We then let all these equalities be a row in a tableau, and put the objective function on the bottom row, and performed row operations on the tableau. We stopped (and claimed we had the optimum value of the objective function) when the bottom row had all non-positive entries. Suppose that at the end of the algorithm, the last row is

$$w_1, w_2, w_3, \dots, w_n, r_1, r_2, \dots, r_m \mid -F$$

where w_k is in the column whose variable is x_k , and r_k is in the column whose variable is s_k . Assuming that the simplex algorithm found an optimum, all entries in the last row will be non-positive, and there will be at least m zeros.

Now, we know this last row is obtained by taking the original objective function

$$v_1, v_2, v_3, \dots, v_n, 0, 0, \dots, 0 \mid 0$$

and adding some linear combination of the rows to it. A typical row is

$$A_{j,1}, A_{j,2}, A_{j,3}, \dots, A_{j,n}, 0, \dots, 0, 1, 0, \dots, 0 \mid 0$$

where the 1 is in the column belonging to s_j . Now, suppose we take the last row to be the original last row minus y_j times the j 'th row. This means that

$$w_k = v_k - \sum_j y_j A_{j,k}$$

and

$$r_j = -y_j.$$

Since $w_k \leq 0$, we see that

$$v_k \leq \sum_j A_{j,k} y_j \quad \forall k.$$

Since $r_j \leq 0$, we see that

$$y_j \geq 0.$$

And finally, from the last column, we have that the optimum value of the objective function, F satisfies

$$F = \sum_j y_j b_j.$$

We have shown that the optimal value of the primal is equal to the optimal value of the dual.

There's more information we can learn from the solution of the dual. If an equation in the primal LP is satisfied with strict inequality, then the corresponding dual variable y_j must be 0 in the optimal dual solution, because otherwise when we multiply this equation by y_j , we introduce an inequality, and the primal and dual optima would not be equal. Similarly, if a variable in the primal LP is non-zero in the optimum solution, the corresponding equation in the dual LP must be satisfied with equality in the optimal solution to the dual.

Often, if there is some intuitive interpretation of the linear program (for example, for maximum flow in a graph), there will also be some intuitive interpretation of the dual linear program (in this case for minimum cut in the graph). The equality of these two linear programs then may correspond to a combinatorial theorem.