Math 340 Lecture 16

A word on notation. To try to minimize confusion here, I’m deciding that I’ll stick with the following notation for variables for the primal and dual.

- Primal problem with \( n \) decision variables \( x_1, \ldots, x_n \) and \( m \) constraints, giving \( m \) slack variables \( x_{n+1}, \ldots, x_{n+m} \).
- Dual problem with \( m \) decision variables that I’ll call \( y_1, \ldots, y_m \) and \( n \) slack variables \( y_{m+1}, \ldots, y_{m+n} \).

In the end it seems like the best choice is to always use \( x_1, \ldots, x_n \) and \( y_1, \ldots, y_m \) for the decision variables, and add on the slack variables afterwards. The downside is that this makes it harder to see how the variables pair up (a decision variable \( x_i \) will pair with a slack variable \( y_{m+i} \) and a slack variable \( x_{m+i} \) will pair with a decision variable \( y_i \)). If I want to emphasize how the pairings work I might write \( \bar{x}_1, \ldots, \bar{x}_m \) instead of \( x_{n+1}, \ldots, x_{n+m} \) for the slack variables in the primal and \( \bar{y}_1, \ldots, \bar{y}_n \) instead of \( y_{m+1}, \ldots, y_{m+n} \) in the dual, so \( x_i \) pairs with \( \bar{y}_i \) and \( y_i \) pairs with \( \bar{x}_i \). (Vanderbei’s book uses \( w_1, \ldots, w_m \) and \( z_1, \ldots, z_n \) for these two sets of slack variables but I’ve already used \( z \) and \( w \) for the objective functions for the primal and dual so I want to stay away from those letters).

In any case: The decision variables for the primal will always start at \( x_1 \), and the decision variables for the dual will always start at \( y_1 \). The last few notes have been updated to reflect this.

Complementary slackness. Theorem (complementary slackness): Let \( \mathbf{x} \) and \( \mathbf{y} \) be feasible solutions for the primal and for the dual. Then they are both optimal if and only if the following two things hold:

- For each decision variable \( x_i \) of the primal, either \( x_i = 0 \) or the corresponding slack variable \( y_{m+i} \) of the dual is zero (i.e. the \( i \)-th constraint for the dual is an equality for \( \mathbf{y} \)).
- For each decision variable \( y_i \) of the dual, either \( y_i = 0 \) or the corresponding slack variable \( x_{n+i} \) of the primal is zero (i.e. the \( i \)-th constraint for the primal is an equality for \( \mathbf{x} \)).

You can also reformulate this to get rid of the slack variables entirely, at the expense of having to write out more of the problem explicitly. So suppose we have the usual primal and dual setup:

- Maximize \( \bar{c} \cdot \mathbf{x} \) subject to \( A\mathbf{x} \leq \bar{b} \) and \( \mathbf{x} \geq 0 \).
- Minimize \( \bar{b} \cdot \mathbf{y} \) subject to \( A^\top \mathbf{y} \geq \bar{c} \) and \( \mathbf{y} \geq 0 \).

If we let \( A = (a_{ij}) \), \( b = (b_i) \), and \( c = (c_i) \) then by definition the slack variables for the primal are

\[
x_{n+i} = b_i - \sum_{j=1}^{n} a_{ij} x_j,
\]

so saying \( x_{n+i} = 0 \) is equivalent to saying \( b_i = \sum_{j=1}^{n} a_{ij} x_j \). Similarly saying \( y_{n+j} = 0 \) is equivalent to saying \( c_j = \sum_{i=1}^{m} a_{ij} y_i \). So, we can rewrite this as:

Theorem (complementary slackness, again): Suppose \((x_1, \ldots, x_n)\) and \((y_1, \ldots, y_m)\) are feasible solutions to the primal and dual (consisting of just the decision variables). Then they are both optimal for their problems if and only if we have:

- For each decision variable \( x_j \) of the primal, either \( x_j = 0 \) or \( \sum_{i=1}^{m} a_{ij} y_i = c_j \) (the slack of the \( j \)-th constraint of the dual is zero).
- For each decision variable \( y_i \) of the dual, either \( y_i = 0 \) or \( \sum_{j=1}^{n} a_{ij} x_j = b_i \) (the slack in the \( i \)-th constraint of the primal is zero).
Proof of complementary slackness. Why is this true? For one direction, it sort of followed from the
proof-by-dictionary: how I set things up, if \( x_i \) was basic in the primal then \( y_i \) was nonbasic in the dual, and
vice-versa. If you read off the solution from the optimal dictionaries, any nonbasic variable gets set to zero,
so you’re done.

Proof in general: We’ll probe the last statement, that doesn’t mention slack variables at all. It turns
out all we really need to do is chain some inequalities together for our feasible solutions \((x_1^*, \ldots, x_n^*)\)
and \((y_1^*, \ldots, y_m^*)\) (using all of these as decision variables, and just not writing down any slack variables anywhere):

\[
\sum_{j=1}^{n} c_j x_j^* \leq \sum_{j=1}^{n} \left( \sum_{i=1}^{m} a_{ij} y_i^* \right) x_j^* = \sum_{i=1}^{m} \left( \sum_{j=1}^{n} a_{ij} x_j^* \right) y_i^* \leq \sum_{i=1}^{m} b_i y_i^*.
\]

The left-hand side is the objective value of the primal at \( \mathbf{x}^* \) and the right-hand side is the objective value
of the dual at \( \mathbf{y}^* \). So okay, basically I’ve just written down a proof of weak duality again, but let’s look more
carefully at the inequalities and see when they’re strict or not.

The left-hand inequality holds because for each \( j \), \( c_j \leq \sum_{i=1}^{m} a_{ij} y_i^* \) - this is the \( i \)-th constraint on the
dual. When can there be equality? We need equality \( c_j x_j^* = \left( \sum_{i=1}^{m} a_{ij} y_i^* \right) x_j^* \) for all \( i \). This happens exactly
when either the \( i \)-th constraint on the dual is an equality (we have \( c_j = \sum_{i=1}^{m} a_{ij} y_i^* \)), or we have \( x_j^* = 0 \) and
don’t care what’s going on with the constraint. So we’ve shown that the left-hand inequality is an equality
if and only if

- For each decision variable \( x_j \) of the primal, either \( x_j = 0 \) or \( \sum_{i=1}^{m} a_{ij} y_i = c_j \) (the slack of the \( j \)-th
constraint of the dual is zero).

The right-hand inequality is similar; it comes from the constraint \( \sum_{j=1}^{n} a_{ij} x_j \leq b_i \) on the primal, and equality
holds if and only if we have equality for all of the terms either because of an equality in this constraint or
because \( y_i^* = 0 \):

- For each decision variable \( y_i \) of the dual, either \( y_i = 0 \) or \( \sum_{j=1}^{n} a_{ij} x_j = b_i \) (the slack in the \( i \)-th
constraint of the primal is zero).

So we’re done! This really was just taking a proof of weak duality and saying “hmm, when do all of the
inequalities actually give me equalities?”

Applying complementary slackness to verify optimality and find dual optimal solutions. The
way the complementary slackness theorem is stated involves solutions \( \mathbf{x}^* \) and \( \mathbf{y}^* \) for both the dual and
primal, and checking if they are simultaneously dual. This is great if I hand you a solution to both and you
want to certify both are optimal, but more often you might be given a solution \( \mathbf{x}^* \) to the primal without
also having one for the dual. So we can rearrange the complementary slackness theorem to deal with this
case as follows:

Theorem: A feasible solution \( \mathbf{x}^* = (x_1^*, \ldots, x_n^*) \) of the primal is optimal if and only if there exist numbers
\((y_1^*, \ldots, y_m^*)\) satisfying

- \( \sum_{i=1}^{m} a_{ij} y_i^* = c_j \) whenever \( x_j^* > 0 \), and \( \sum_{i=1}^{m} a_{ij} y_i^* \geq c_j \) always.

- \( y_i^* = 0 \) whenever \( \sum_{j=1}^{n} a_{ij} x_j^* < b_i \) and \( y_i^* \geq 0 \) always.

Moreover if there is such a \((y_1^*, \ldots, y_m^*)\) then it’s optimal for the dual.

Here the “...and (inequality) always” parts at the end are just the conditions for \( \mathbf{y}^* \) to actually be feasible
for the dual. But if we’re starting with our purported optimal solution \( \mathbf{x}^* \) for the primal, complementary
slackness forces a bunch of these inequalities to actually be equalities, which gives you a good head start on
finding a solution! (If the dictionary for \( \mathbf{x}^* \) is nondegenerate you’ll actually get \( m \) equalities, which most
of the time will let you solve for the \( m \) variables \( y_1^*, \ldots, y_m^* \) directly via linear algebra, no simplex method
required).
Examples: Using this to check optimality, and to find dual optimal solutions. Okay, to start out, I know I was asked a few times “what is duality useful for”? I don’t think I emphasized enough one of the important applications of it: verifying optimal solutions that someone else has found, or providing a certificate of optimality that your solution is actually correct. Just saying “I did the simplex method and the optimal solution is this” doesn’t let someone else check that you’re right so easily - maybe they could go back and reconstruct the final dictionary you got but that can be a pretty unpleasant computation. But if you say “here’s my optimal solution for the primal and my optimal solution for the dual” it’s much easier for someone to check that you’re right - just check both solutions are feasible, then check the optimal values are equal.

The above proposition is kind of in the same spirit - if I give you something I claim is an optimal solution \(x^*\) for the primal, it gives you a good place to start with constructing an optimal solution for the dual (and thus you either certify I’m right, or run into a problem and conclude I’m wrong).

So for an example consider the following primal LP:

- Maximize \(-4x_1 + 8x_2 + 7x_3\) subject to
  - \(2x_1 + 2x_2 + x_3 \leq 1\)
  - \(x_1 - 4x_2 - 3x_3 \leq 1\)
  - \(x_1, x_2, x_3 \geq 0\).

We can then write down the dual:

- Minimize \(y_1 + y_2\) subject to
  - \(2y_1 + y_2 \geq -4\).
  - \(2y_1 - 4y_2 \geq 8\).
  - \(y_1 - 3y_2 \geq 7\).

Okay, so we could just go and solve this and find the answer to both. But what if someone comes along and tells you \((x^*_1, x^*_2, x^*_3) = (0, 1/2, 0)\) is an optimal solution for the primal? You can check it’s feasible just fine, and then we can go to our proposition above. If this solution was optimal then there would be an optimal solution \((y^*_1, y^*_2)\) to the dual, which satisfies \(y^*_2 = 0\) (because the second constraint for the primal is a 
\(<\) and \(2y^*_1 - 4y^*_2 = 8\) (because \(x^*_2 > 0\)). So you can then solve that you have to have \(y^*_1 = 4\) and the optimal solution for the dual has to be \((y^*_1, y^*_2) = (4, 0)\). So, does this actually work? The answer is no because it’s not feasible; the last constraint is violated! What we’ve actually shown is that there’s no choice of \((y^*_1, y^*_2)\) satisfying the conditions in the theorem. This proves the \((x^*_1, x^*_2, x^*_3)\) we started with was not optimal, and whoever told you that in the first place was wrong.

So then what if someone more trustworthy tells you the optimal solution is actually \((x^*_1, x^*_2, x^*_3) = (0, 0, 1)\)? We can do the same thing as above: the second constraint has some slack so \(y^*_2 = 0\), and the third decision variable is positive so the third dual constraint must be an equality \(y^*_1 - 3y^*_2 = 7\). So the dual solution we get has to be \((y^*_1, y^*_2) = (7, 0)\). Now, is this one feasible? You just check all of the constraints and find yes, it is. So we’ve proved that there exists a solution \(y^*\) as in the theorem, so the theorem lets us conclude that the \(x^*\) we were given was optimal for the primal, and the \(y^*\) we found was optimal for the dual.