Supplementary notes on duality
Math 354-01, Spring 2016

This is a supplement to Sections 3.1, 3.2, and 3.3. It gives more details
about the Duality Theorem as well as the important idea of complementary
slackness.

Let’s say that we have a LPP \((P)\) in standard form, and its dual \((D)\):

\[
\begin{align*}
(P) & : \max z_P = c^T x \\
& \text{subject to } Ax \leq b, \quad x \geq 0
\end{align*}
\]

\[
\begin{align*}
(D) & : \min z_D = b^T w \\
& \text{subject to } A^T w \geq c, \quad w \geq 0
\end{align*}
\]

Here \(A\) is any \(m \times n\) matrix, \(x = [x_1 \cdots x_n]^T\) is the column vector
of primal variables and \(w = [w_1 \cdots w_m]^T\) is the column vector of dual
variables. Also \(c\) is a column vector in \(\mathbb{R}^n\) and \(b\) a column vector in \(\mathbb{R}^m\),
each playing the role of “cost vector” in one problem and “resource vector”
in the other.

1. Standard Form, Slack Variables, Canonical Form

In this totally pedestrian section we prepare for further analysis by
putting \((D)\) in standard form, and then putting both problems in canonical
form. Nothing new is happening here, but keep your eye on the notation.
The reason is that slack variables will turn out to be a key connector between
\((P)\) and \((D)\), so we will use carefully chosen notation for them that is a bit
unorthodox.

First, putting \((D)\) in standard form means changing the minimization
problem to a maximization problem, and reversing the inequality sign
in the constraint \(Aw \geq c\). Our problems then read as follows. The resulting
problem \((D')\) is equivalent to \((D)\).

\[
\begin{align*}
(P) & : \max z_P = c^T x \\
& \text{subject to } Ax \leq b, \quad x \geq 0
\end{align*}
\]

\[
\begin{align*}
(D') & : \max -z_D = (-b)^T w \\
& \text{subject to } (-A^T)w \leq (-c), \quad w \geq 0
\end{align*}
\]

Next, canonical form. For \((P)\), introduce slack variables \(w_1^*, \ldots, w_m^*\),
for the constraints numbered 1, \ldots, \(m\), respectively. (The names are an
important aid in keeping track of the situation.) Likewise introduce slack
variables for \((D')\), called \(x_1^*, \ldots, x_n^*\). In canonical form, \((P)\) and \((D')\) are

\[
\begin{align*}
(P) & : \max z_P = c^T x = C^T X \\
& \text{subject to } [A | I]X = b \\
x \geq 0
\end{align*}
\]

\[
\begin{align*}
(D') & : \max -z_D = -b^T w = -B^T W \\
& \text{subject to } [I | -A^T]W = -c \\
W \geq 0
\end{align*}
\]

\[
X = \begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix}, \quad B = \begin{bmatrix} 0 \end{bmatrix}, \quad W = \begin{bmatrix} w_1^* & \cdots & w_m^* \end{bmatrix}, \quad C = \begin{bmatrix} c \end{bmatrix}
\]

The unorthodox thing is that we have put the slack variables for
\((D')\) first in \(W\), instead of last. This is very helpful for this theoretical
discussion.

2. Duality Gap; the Weak Duality Theorem

Suppose that \(X\) and \(W\) as above are solutions to the canonical form
versions of \((P)\) and \((D')\). Then the slack variables \(w_1^*, \ldots, w_m^*\) make up the difference between \(Ax\) and \(b\). Therefore

\[
\begin{bmatrix} w_1^* \\ \vdots \\ w_m^* \end{bmatrix} = b - Ax, \quad \text{and so } X = \begin{bmatrix} x \\ b - Ax \end{bmatrix}.
\]

Similarly \(x_1^*, \ldots, x_n^*\) make up the difference between \(c\) and \(A^T w\). Therefore

\[
\begin{bmatrix} x_1^* \\ \vdots \\ x_n^* \end{bmatrix} = A^T w - c, \quad \text{and so } W = \begin{bmatrix} A^T w - c \end{bmatrix}.
\]
This implies the useful “Duality Gap” identity:
\[
X \cdot W = W \cdot X = (-c + A^T w)^T x + w^T (b - Ax)
= -c^T x + w^T (A^T x) + w^T b - w^T A x
= -c^T x + w^T (A^T x) + w^T b - w^T A x = w^T b - c^T x
\]

That is,
\[
X \cdot W = b^T w - c^T x.
\]

This is the gap between the value \( z_D = b^T w \) of the objective function of \((D)\) at \( w \) and the value of \( z_P = c^T x \) of the objective function of \((P)\) at \( x \).

Furthermore, \( x \) is a feasible solution of \((P)\) if and only if \( X \geq 0 \), and \( w \) is a feasible solution of \((D)\) if and only if \( W \geq 0 \). An immediate consequence of these observations is the Weak Duality Theorem, which in turn has some important corollaries.

**Weak Duality Theorem.** Let \( x \) and \( w \) be feasible solutions of \((P)\) and \((D)\), respectively. Then
\[
z_D(w) \geq z_P(x), \text{ that is, } b^T w \geq c^T x.
\]

**Proof.** Since \( x \) and \( w \) are feasible, \( X \geq 0 \) and \( W \geq 0 \). Hence, \( X \cdot W \geq 0 \). By the Duality Gap, \( b^T w \geq c^T x \).

**Corollary 1.** If either \((P)\) or \((D)\) is an unbounded LPP, then the other one is an infeasible LPP.

**Corollary 2.** Suppose that \( x_0 \) is a feasible solution of \((P)\), and that \( w_0 \) is a feasible solution of \((D)\). Suppose also that \( b^T w_0 = c^T x_0 \). Then \( x_0 \) and \( w_0 \) are optimal solutions of \((P)\) and \((D)\), respectively.

**Proof.** Let’s show that \( x_0 \) is an optimal solution for \((P)\). We know it’s a feasible solution. Let \( x \) be any feasible solution for \((P)\). Then
\[
c^T x_0 = b^T w_0 \geq c^T x,
\]

by assumption (=) and the Weak Duality Theorem (≥). Since this holds for any feasible solution \( x \), \( x_0 \) is an optimal solution of \((P)\).

The proof for \( w_0 \) is similar.

### 3. Complementary Slackness

Again given a pair of vectors \( X \) and \( W \) as above, they may or may not have an important property known as **complementary slackness**, or just **complementarity**. This property stipulates that
\[
x_i x_i^* = 0 \text{ and } w_j^* w_j = 0 \text{ for all } i = 1, \ldots, n \text{ and } j = 1, \ldots, m.
\]

In other words, this property holds if and only if for each of the \( n + m \) positions, one or both of the vectors \( X \) and \( W \) has a 0 entry. Each position corresponds to a variable \( x_i \) or \( w_j \) for \((P)\) or \((D)\), and the slack variable \( x_i^* \) or \( w_j^* \) for the corresponding constraint in the dual problem.

So complementarity means that if some variable for (the standard form problem) \((P)\) or \((D)\) is nonzero, then both sides of the corresponding constraint inequality in \((D)\) or \((P)\) are equal, i.e., the corresponding slack variable for \((D)\) or \((P)\) equals 0.

Complementarity clearly implies that \( X \cdot W = 0 \). On the other hand if \( X \cdot W = 0 \), then complementarity need not hold, because \( X \) and \( W \) can have negative entries. However, if \( X \geq 0 \) and \( W \geq 0 \) and \( X \cdot W = 0 \), then \( X \) and \( W \) are complementary. In this case the dot product is a sum of nonnegative terms, and so can only be zero if all the terms are zero.

Complementarity and optimality are closely related.

**Complementary Slackness Theorem.** Let \( x, w \), \( X \) and \( W \) be as above. Suppose that \( x \) and \( w \) are feasible solutions for \((P)\) and \((D)\), respectively. Then
\[
x \text{ and } w \text{ are optimal solutions for } (P) \text{ and } (D) \text{ (resp.)}
\iff
X \text{ and } W \text{ satisfy complementary slackness.}
\]

Furthermore, when this occurs, \( c^T x = b^T w \).

**Proof of \( \iff \).** Complementary slackness implies \( X \cdot W = 0 \). By the Duality Gap \( c^T x = b^T w \). By Corollary 2 to the Weak Duality Theorem, \( x \) and \( w \) are optimal solutions of \((P)\) and \((D)\), respectively.

The proof of \( \implies \) is harder. It is part (c) of the Duality Theorem, just below.
4. The Duality Theorem

This is generally agreed to be the most important theorem in the basic theory of linear programming.

**Duality Theorem.** Consider a LPP \((P)\) and its dual \((D)\).

(a) If either problem is unbounded, then the other is infeasible.

(b) If either problem has an optimal solution, then so does the other, and the optimal values of the objective functions for the two problems are the same.

(c) Any optimal solutions of \((P)\) and \((D)\) satisfy complementary slackness.

Let’s prove this theorem. Part (a) is Corollary 1 above; it followed from the Weak Duality Theorem.

Consider (b). Since \((P)\) is also the dual of \((D)\), it’s enough to show that if \((P)\) has an optimal solution, say \(x\), then so does \((D)\), say \(w\), and the duality gap \(b^T w - c^T x\) between them equals 0.

To see this, imagine having solved \((P)\) by the simplex method. The initial tableau was

\[
\begin{array}{ccc}
 x & w^* \\
 A & I & b \\
 -c^T & 0 & 0 \\
\end{array}
\]

The effect of all the row operations applied to the constraint rows, as we go from initial to final tableau, is the same as left multiplication of the upper part of the initial tableau by some invertible matrix, usually called \(B^{-1}\) in this context. So in the final tableau we have \(B^{-1}A\) in place of \(A\), \(B^{-1}I\) in place of \(I\), and \(B^{-1}b\) in place of \(b\).

But the objective row changes as well, on the way to the final tableau. It is always changed by adding multiples of the upper rows to it. These upper rows, in any tableau, are linear combinations of the rows of \([A | I | b]\). Call those rows \(a_1, \ldots, a_m\). Then the new objective row is \([-c^T 0 0|0] + w_1 \cdot a_1 + w_2 a_2 + \cdots + w_m a_m\), for some scalars \(w_1, \ldots, w_m\). This can be described compactly in terms of matrix notation. Let \(w = \begin{bmatrix} w_1 \\ \vdots \\ w_m \end{bmatrix}\). Then

\[
-b^T \begin{bmatrix} 0 & I \end{bmatrix} \begin{bmatrix} w \\ I \end{bmatrix} = b^T w - c^T x
\]

in the objective row, \(-c^T\) changes to \(-c^T + w^T A\); the 0, under \(I\), changes to \([w_1 \cdots w_m]\) = \(w^T\); and the 0 in the lower right corner changes to \(w^T b\). Therefore the final tableau is:

\[
\begin{array}{ccc}
 x & w^* \\
 B^{-1}A & B^{-1} & B^{-1}b \\
 (A^T w - c)^T & w^T & w^T b \\
\end{array}
\]

The whole objective row is determined by the part \(w^T\) of it in the slack variable columns (and by \(A, b,\) and \(c\)). The row vector that appears there, which we have called \(w^T\), has wonderful and surprising properties.

Since the final tableau satisfies the optimality criterion,

\[
A^T w - c \geq 0 \text{ and } w \geq 0
\]

so \(w\) is a feasible solution of \((D)\)!!!

Moreover, the corresponding value of \(z_D\), which is \(b^T w\), equals \(w^T b\) so equals the optimal value of \(z_P\)!!!.

Then because of Corollary 2 above, \(w\) is an optimal solution of \((D)\)!!!.

And as a result the optimal values of \(z_P\) and \(z_D\) are equal.

Finally, consider (c). Suppose \(x\) and \(w\) are optimal solutions to \((P)\) and \((D)\), respectively. By part (b) the duality gap is 0. Therefore

\[
X \cdot W = 0.
\]

But since \(x\) is feasible for \((P)\), \(X \geq 0\), and since \(w\) is feasible for \((D)\), \(W \geq 0\). As \(X\) and \(W\) have only nonnegative entries, the only way to have \(X \cdot W = 0\) is to have complementary slackness. Q.E.D.

Notice that the final tableau for \((P)\) contains an optimal solution for \((D)\); \(w\) is in the slack variable columns of the final tableau, and the optimal value of \(z_D\) equals the optimal value of \(z_P\), in the lower right corner.

5. Corresponding Basic Solutions of \((P)\) and \((D)\)

A basic solution \(X\) to \((P)\) has \(m\) basic variables and \(n\) non-basic variables; a basic solution \(W\) to \((D)\) has \(n\) basic variables and \(m\) nonbasic variables. Since nonbasic variables are 0, \(X\) and \(W\) could be a perfect fit for complementary slackness, if \(X\) has basic variables just where \(W\) has nonbasic variables (and vice-versa).
Definition. Let \( X \) and \( W \) be basic solutions to \((P)\) and \((D)\), respectively. We will say that \( X \) and \( W \) \textit{correspond} if and only if for each \( i = 1, \ldots, n \) and each \( j = 1, \ldots, m \), exactly one of \( x_i \) and \( x_i^* \) is a basic variable and exactly one of \( w_j \) and \( w_j^* \) is a basic variable.

Corresponding basic solutions are very closely related, in the following ways.

**PROPERTY 1.** Corresponding basic solutions \( X \) and \( W \) of \((P)\) and \((D)\) satisfy complementary slackness.

Property (1) is not hard to see, because for each \( i \), either \( x_i \) or \( x_i^* \) is nonbasic, so either \( x_i = 0 \) or \( x_i^* = 0 \), so \( x_i x_i^* = 0 \).

**PROPERTY 2.** For each basic solution \( X \) of \((P)\) (and choice of basic variables), there is a unique corresponding basic solution \( W \) of \((D)\) (and choice of basic variables).

Indeed the basic variables of \( X \) being given, the basic variables of \( W \) are then determined, and so \( W \) is determined.

**PROPERTY 3.** Let \( X \) and \( W \) be corresponding basic solutions of \((P)\) and \((D)\). Let \( T_X \) and \( T_W \) be tableaux for \((P)\) and \((D)\) corresponding to \( X \) and \( W \), respectively. Delete (or ignore) the basic columns of \( T_X \) and \( T_W \) to obtain arrays \( T'_X \) and \( T'_W \). Then

(a) The values of \( z_0 \) in \( T'_X \) and \( T'_W \) are negatives of each other.

(b) The resource column of \( T'_X \) is “identical” to the transpose of the objective row of \( T'_W \), and vice-versa.

(c) The central portion of \( T'_X \) is “identical” to the negative transpose of the central portion of \( T'_W \).

The word “identical” is in quotation marks because (b) and (c) are not literally true. Let’s use the following convention: if \( g \) is one of the variable names for \( X \), then \( g' \) is the corresponding variable name for \( W \) — obtained either by adding or removing the asterisk. So if \( g = x_i \) then \( g' = x_i^* \); if \( g = w_j^* \) then \( g' = w_j \). In (c), every entry \( a_{gh} \) of \( T'_X \) is in a row labelled by a variable \( g \) in \( X \), and a column labelled by a variable \( h \) in \( X \). A similar labelling system holds for entries \( a'_{h'g'} \) in \( T'_W \). What (c) asserts is that

\[
a_{gh} = -a'_{h'g'}, \text{ for all basic } g \text{ and nonbasic } h \text{ in } X.
\]

Notice that \( g' \) is nonbasic in \( W \) and \( h' \) is basic in \( W \).

Similarly (b) means that the objective row entry of \( T'_X \) for a nonbasic variable \( g \) equals the resource column entry of \( T'_W \) for the (basic!) variable \( g' \). And a similar connection holds between the objective row of \( T'_W \) and the resource column of \( T'_X \).

Properties (2) and (3) are not obvious. Their proofs are given in exercises below.

Here is an example. It is only to illustrate the above properties; we are not actually going to solve a LPP. Consider the following primal-dual pair:

\[
\begin{align*}
(P) & \quad \text{Maximize } 2x_1 - 3x_2 \\
& \quad \text{subject to } 4x_1 - x_2 \leq 3 \\
& \quad \quad 3x_1 + 5x_2 \leq -2 \\
& \quad \quad -x_1 + 2x_2 \leq -1 \\
& \quad \quad x_1 \geq 0 \\
& \quad \quad x_2 \geq 0
\end{align*}
\]

\[
\begin{align*}
(D') & \quad \text{Maximize } -3w_1 + 2w_2 + w_3 \\
& \quad \text{subject to } -4w_1 - 3w_2 + w_3 \leq -2 \\
& \quad \quad w_1 - 5w_2 - 2w_3 \leq 3 \\
& \quad \quad w_1 \geq 0 \\
& \quad \quad w_2 \geq 0 \\
& \quad \quad w_3 \geq 0
\end{align*}
\]

After introducing the slack variables, we can consider the initial tableaux:

\[
\begin{array}{cccccccc}
(P) & & & & & & (D') \\
& x_1 & x_2 & w_1^* & w_2^* & w_3^* & x_1^* & x_2^* & w_1 & w_2 & w_3 \\
\hline
w_1^* & 4 & -1 & 1 & 0 & 0 & 3 & x_1^* & 1 & 0 & -4 & -3 & 1 & -2 \\
w_2^* & 3 & 5 & 0 & 1 & 0 & -2 & x_2^* & 0 & 1 & 1 & -5 & -2 & 3 \\
w_3^* & -1 & 2 & 0 & 0 & 1 & -1 & & 0 & 0 & 3 & -2 & -1 & 0 \\
\hline
-2 & 3 & 0 & 0 & 0 & 0 & & & & & & & \\
\end{array}
\]

Notice that if you ignore the basic columns and ignore stars, the tableaux are transposes of each other, with minus signs inserted in the core and the objective function value \( z_0 \). This is what (3) says.

Now let’s consider another basic solution for \((P)\), say the one with basic variables \( x_1, x_2, w_3^* \). The corresponding basic solution for \((D')\) has basic variables \( w_1, w_2 \). Using row operations to create a \( 3 \times 3 \) identity matrix in the \( x_1, x_2, w_3^* \) columns of the initial \((P)\) tableau, we get the \((P)\) tableau for the new basic solution. Similarly, the new \((D')\) tableau comes from pivoting in the initial \((D')\) tableau, to make a \( 2 \times 2 \) identity matrix in the \( w_1, w_2 \) columns. The new pair of corresponding tableaux is
In the same rows. Show that

\[ T = T' \]

Let

\[ T, T' \]

be tableaux satisfying the U.C., but for infeasibility any tableau satisfying the I.C. will do. For this reason it is possible that the dual of an infeasible problem may be infeasible as well.

The text gives a nice chart of the possible outcomes for a primal and dual problem.

7. The Dual Simplex Method

A couple of remarks here. The ordinary ("primal") simplex method and dual simplex method apply to different kinds of starting tableaux for a LPP. Both aim to find an optimal solution of the LPP, but if no optimal solution exists, the two methods fail in different ways.

SIMPLEX: any feasible tableau \( \rightarrow \) an optimal solution, or LPP is unbounded

DUAL SIMPLEX: any dual-feasible tableau \( \rightarrow \) an optimal solution, or LPP is infeasible

When one carries out the dual simplex method on tableaux \( T_X \) for a LPP, the corresponding tableaux \( T_W \) for the dual problem are going through the steps of the simplex method.

8. Exercises

Ex. 1-2: Duality

1. For a certain LPP in standard form, slack variables have been introduced and the simplex method is applied. At a certain point the following tableau is reached. What is the tableau for the corresponding basic solution of the dual problem? What are the current values of the dual variables \( w_1, w_2, w_3 \)? Is the current basic solution optimal? optimal for the dual? feasible? feasible for the dual?

2. Can you recover the original constraints \( Ax \leq b \) and the objective coefficient vector \( c \) from the tableau in the previous problem?

Ex. 3-6: Tableaux for the primal and dual

3. Let \( A \) be \( m \times n \). Form \( B = [A \mid I] \) and \( C = [I \mid -A^T] \), of sizes \( m \times (m + n) \) and \( n \times (m + n) \), respectively. Suppose that \( S \subset \{1, \ldots, m + n\} \) and \( S' \) is the complementary subset. Suppose that the columns \( b_i, i \in S \), are linearly independent. Then prove that the columns \( c_j, j \in S' \) of \( C \) are also linearly independent.

4. Let \( T \) and \( T' \) be tableaux for the same LPP with the same basic variables, and in the same rows. Show that \( T = T' \).
5. Let $X$ and $W$ be corresponding basic solutions of $(P)$ and $(D')$, with corresponding tableaux $T_X$ and $T_W$. Show that the objective row of $T_X$, with the last entry removed, is $W^T$. Show also that the lower right corners of $T_W$ and $T_X$ are negatives of each other.

6. Let $X$ and $W$ and $T_X$ and $T_W$ be as in the previous problem. Let $T^c_X$ and $T^c_W$ be the cores of these two tableaux. Let $g$ be a basic variable and $h$ a nonbasic variable for $(P)$. Let $g'$ and $h'$ be the corresponding variables for $(D)$. Show that if $a$ is the $(g, h)$ entry of $T^c_X$ and $b$ is the $(h', g')$ entry of $T^c_W$, then $a + b = 0$. (Hint: Show that $a + b$ is an entry of $(T^c_X(T^c_W)^T$, which is the 0 matrix. It is useful to recall from Chapter 3 that $T^c_X = B^{-1}[A|I]$ and $T^c_W = (B')^{-1}[I - A^T]$, for some invertible matrices $B$ and $B'$.)