CHAPTER 5

Duality Theory

Associated with every linear program is another called its dual. The dual of this dual linear program is the original linear program (which is then referred to as the primal linear program). Hence, linear programs come in primal/dual pairs. It turns out that every feasible solution for one of these two linear programs gives a bound on the optimal objective function value for the other. These ideas are important and form a subject called duality theory, which is the topic of this chapter.

1. Motivation: Finding Upper Bounds

We begin with an example:

$$
\begin{array}{ll}
\text{maximize} & 4x_1 + x_2 + 3x_3 \\
\text{subject to} & x_1 + 4x_2 \phantom{{}+x_3} \le 1 \\
& 3x_1 - x_2 + x_3 \le 3 \\
& x_1,\ x_2,\ x_3 \ge 0.
\end{array}
$$

Our first observation is that every feasible solution provides a lower bound on the optimal objective function value, ζ^*. For example, the solution $(x_1, x_2, x_3) = (1, 0, 0)$ tells us that $\zeta^* \ge 4$. Using the feasible solution $(x_1, x_2, x_3) = (0, 0, 3)$, we see that $\zeta^* \ge 9$. But how good is this bound? Is it close to the optimal value? To answer, we need to give upper bounds, which we can find as follows. Let us multiply the first constraint by 2 and add that to 3 times the second constraint:

$$
\begin{array}{rcl}
2\ \ (x_1 + 4x_2 \phantom{{}+x_3}) & \le & 2(1) \\
+3\,(3x_1 - x_2 + x_3) & \le & 3(3) \\
\hline
11x_1 + 5x_2 + 3x_3 & \le & 11.
\end{array}
$$

Now, since each variable is nonnegative, we can compare the sum against the objective function and notice that

$$
4x_1 + x_2 + 3x_3 \ \le\ 11x_1 + 5x_2 + 3x_3 \ \le\ 11.
$$

Hence, $\zeta^* \le 11$. We have localized the search to somewhere between 9 and 11. These bounds leave a gap (within which the optimal solution lies), but they are better than nothing. Furthermore, they can be improved. To get a better upper bound, we again

R. J. Vanderbei, *Linear Programming*, International Series in Operations Research & Management Science 285,
https://doi.org/10.1007/978-3-030-39415-8_5

apply the same upper bounding technique, but we replace the specific numbers we used before with variables and then try to find the values of those variables that give us the best upper bound. So we start by multiplying the two constraints by nonnegative numbers, y_1 and y_2, respectively. The fact that these numbers are nonnegative implies that they preserve the direction of the inequalities. Hence,

$$\begin{array}{rcl} y_1(\ x_1 + 4x_2 \qquad\ ) & \leq & y_1 \\ +y_2(3x_1 - \ x_2 + x_3) & \leq & 3y_2 \\ \hline (y_1 + 3y_2)x_1 + (4y_1 - y_2)x_2 + (y_2)x_3 & \leq & y_1 + 3y_2. \end{array}$$

If we stipulate that each of the coefficients of the x_i's be at least as large as the corresponding coefficient in the objective function,v

$$\begin{aligned} y_1 + 3y_2 &\geq 4 \\ 4y_1 - \ y_2 &\geq 1 \\ y_2 &\geq 3, \end{aligned}$$

then we can compare the objective function against this sum (and its bound):

$$\begin{aligned} \zeta &= 4x_1 + x_2 + 3x_3 \\ &\leq (y_1 + 3y_2)x_1 + (4y_1 - y_2)x_2 + (y_2)x_3 \\ &\leq y_1 + 3y_2. \end{aligned}$$

We now have an upper bound, $y_1 + 3y_2$, which we should minimize in our effort to obtain the best possible upper bound. Therefore, we are naturally led to the following optimization problem:

$$\begin{array}{lrcl} \text{minimize} & y_1 + 3y_2 & & \\ \text{subject to} & y_1 + 3y_2 & \geq & 4 \\ & 4y_1 - \ y_2 & \geq & 1 \\ & y_2 & \geq & 3 \\ & y_1,\ y_2 & \geq & 0. \end{array}$$

This problem is called the dual linear programming problem associated with the given linear programming problem. In the next section, we will define the dual linear programming problem in general.

2. The Dual Problem

Given a linear programming problem in standard form,

$$\begin{array}{lll} \text{maximize} & \displaystyle\sum_{j=1}^{n} c_j x_j & \\ \text{subject to} & \displaystyle\sum_{j=1}^{n} a_{ij} x_j \leq b_i & i = 1, 2, \ldots, m \\ & x_j \geq 0 & j = 1, 2, \ldots, n, \end{array} \tag{5.1}$$

the associated *dual linear program* is given by

$$\begin{aligned} \text{minimize} \quad & \sum_{i=1}^m b_i y_i \\ \text{subject to} \quad & \sum_{i=1}^m y_i a_{ij} \geq c_j \qquad & j = 1, 2, \ldots, n \\ & y_i \geq 0 \qquad & i = 1, 2, \ldots, m. \end{aligned}$$

Since we started with (5.1), it is called the *primal problem*. Our first order of business is to show that taking the dual of the dual returns us to the primal. To see this, we first must write the dual problem in standard form. That is, we must change the minimization into a maximization and we must change the first set of greater-than-or-equal-to constraints into less-than-or-equal-to. Of course, we must make these changes without altering the problem. To change a minimization into a maximization, we note that to minimize something it is equivalent to maximize its negative and then negate the answer:

$$\min \sum_{i=1}^m b_i y_i = -\max \left(-\sum_{i=1}^m b_i y_i \right).$$

To change the direction of the inequalities, we simply multiply through by minus one. The resulting equivalent representation of the dual problem in standard form then is

$$\begin{aligned} -\text{maximize} \quad & \sum_{i=1}^m (-b_i) y_i \\ \text{subject to} \quad & \sum_{i=1}^m (-a_{ij}) y_i \leq (-c_j) \qquad & j = 1, 2, \ldots, n \\ & y_i \geq 0 \qquad & i = 1, 2, \ldots, m. \end{aligned}$$

Now we can take its dual:

$$\begin{aligned} -\text{minimize} \quad & \sum_{j=1}^n (-c_j) x_j \\ \text{subject to} \quad & \sum_{j=1}^n (-a_{ij}) x_j \geq (-b_i) \qquad & i = 1, 2, \ldots, m \\ & x_j \geq 0 \qquad & j = 1, 2, \ldots, n, \end{aligned}$$

which is clearly equivalent to the primal problem as formulated in (5.1).

3. The Weak Duality Theorem

As we saw in our example, the dual problem provides upper bounds for the primal objective function value. This result is true in general and is referred to as the *weak duality theorem*:

THEOREM 5.1. *If $(x_1, x_2, \ldots, x_n)$ is feasible for the primal and $(y_1, y_2, \ldots, y_m)$ is feasible for the dual, then*

$$\sum_j c_j x_j \leq \sum_i b_i y_i.$$

PROOF. The proof is a simple chain of obvious inequalities:

$$\begin{aligned}
\sum_j c_j x_j &\leq \sum_j \left(\sum_i y_i a_{ij} \right) x_j \\
&= \sum_{ij} y_i a_{ij} x_j \\
&= \sum_i \left(\sum_j a_{ij} x_j \right) y_i \\
&\leq \sum_i b_i y_i,
\end{aligned}$$

where the first inequality follows from the fact that each x_j is nonnegative and each c_j is no larger than $\sum_i y_i a_{ij}$. The second inequality, of course, holds for similar reasons. □

Consider the subset of the real line consisting of all possible values for the primal objective function, and consider the analogous subset associated with the dual problem. The weak duality theorem tells us that the set of primal values lies entirely to the left of the set of dual values. As we shall see shortly, these sets are both closed intervals (perhaps of infinite extent), and the right endpoint of the primal set butts up against the left endpoint of the dual set (see Figure 5.1). That is, there is no gap between the optimal objective function value for the primal and for the dual. The lack of a gap between primal and dual objective values provides a convenient tool for verifying optimality. Indeed, if we can exhibit a feasible primal solution $(x_1^*, x_2^*, \ldots, x_n^*)$ and a feasible dual solution $(y_1^*, y_2^*, \ldots, y_m^*)$ for which

$$\sum_j c_j x_j^* = \sum_i b_i y_i^*,$$

then we may conclude that each of these solutions is optimal for its respective problem. To see that the primal solution is optimal, consider any other feasible solution $(x_1, x_2, \ldots, x_n)$. By the weak duality theorem, we have that

$$\sum_j c_j x_j \leq \sum_i b_i y_i^* = \sum_j c_j x_j^*.$$

Now, since $(x_1^*, x_2^*, \ldots, x_n^*)$ was assumed to be feasible, we see that it must be optimal. An analogous argument shows that the dual solution is also optimal for the dual problem. As an example, consider the solutions $x = (0, 0.25, 3.25)$ and $y = (1, 3)$

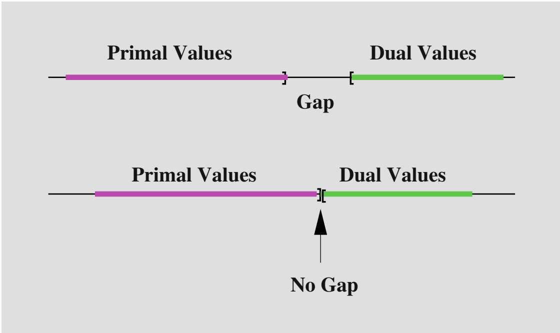

FIGURE 5.1. The primal objective values are all less than the dual objective values. An important question is whether or not there is a gap between the largest primal value and the smallest dual value.

in our example. Both these solutions are feasible, and both yield an objective value of 10. Hence, the weak duality theorem says that these solutions are optimal.

4. The Strong Duality Theorem

The fact that for linear programming there is never a gap between the primal and the dual optimal objective values is usually referred to as the *strong duality theorem*:

THEOREM 5.2. *If the primal problem has an optimal solution,*

$$x^* = (x_1^*, x_2^*, \dots, x_n^*),$$

then the dual also has an optimal solution,

$$y^* = (y_1^*, y_2^*, \dots, y_m^*),$$

such that

$$\sum_j c_j x_j^* = \sum_i b_i y_i^*. \tag{5.2}$$

Carefully written proofs, while attractive for their tightness, sometimes obfuscate the main idea. In such cases, it is better to illustrate the idea with a simple example. Anyone who has taken a course in linear algebra probably already appreciates such a statement. In any case, it is true here as we explain the strong duality theorem.

The main idea that we wish to illustrate here is that, as the simplex method solves the primal problem, it also implicitly solves the dual problem, and it does so in such a way that (5.2) holds.

To see what we mean, let us return to the example discussed in Section 5.1. We start by introducing variables w_i, $i = 1, 2$, for the primal slacks and z_j, $j = 1, 2, 3$, for the dual slacks. Since the inequality constraints in the dual problem are greater-than constraints, each dual slack is defined as a left-hand side minus the corresponding right-hand side. For example,

$$z_1 = y_1 + 3y_2 - 4.$$

Therefore, the primal and dual dictionaries are written as follows:

$$\text{(P)} \qquad \begin{array}{rcrrrrrr} \zeta & = & & + 4 x_1 & + & 1 x_2 & + & 3 x_3 \\ \hline w_1 & = & 1 & - x_1 & - & 4 x_2 & & \\ w_2 & = & 3 & - 3 x_1 & + & x_2 & - & x_3 \end{array}$$

$$\text{(D)} \qquad \begin{array}{rcrrrrr} -\xi & = & & - & 1 y_1 & - & 3 y_2 \\ \hline z_1 & = & -4 & + & y_1 & + & 3 y_2 \\ z_2 & = & -1 & + & 4 y_1 & - & y_2 \\ z_3 & = & -3 & & & + & y_2 \end{array}$$

Note that we have recorded the negative of the dual objective function, since we prefer to maximize the objective function appearing in a dictionary. Also note that the numbers in the dual dictionary are simply the negative of the numbers in the primal dictionary arranged with the rows and columns interchanged. Indeed, stripping away everything but the numbers, we have

$$\begin{bmatrix} 0 & 4 & 1 & 3 \\ 1 & -1 & -4 & 0 \\ 3 & -3 & 1 & -1 \end{bmatrix} \overset{\text{neg.–transp.}}{\longleftrightarrow} \begin{bmatrix} 0 & -1 & -3 \\ -4 & 1 & 3 \\ -1 & 4 & -1 \\ -3 & 0 & 1 \end{bmatrix}.$$

That is, as a table of numbers, the dual dictionary is the *negative transpose* of the primal dictionary.

Our goal now is to apply the simplex method to the primal problem and at the same time perform the analogous pivots on the dual problem. We shall discover that the negative-transpose property persists throughout the iterations.

Since the primal dictionary is feasible, no Phase I procedure is necessary. For the first pivot, we pick x_3 as the entering variable (x_1 has the largest coefficient, but x_3 provides the greatest one-step increase in the objective). With this choice, the leaving variable must be w_2. Since the rows and columns are interchanged in the dual dictionary, we see that "column" x_3 in the primal dictionary corresponds to "row" z_3 in the dual dictionary. Similarly, row w_2 in the primal corresponds to column y_2 in the dual. Hence, to make an analogous pivot in the dual dictionary, we select y_2 as the entering variable and z_3 as the leaving variable. While this choice of entering and leaving variable may seem odd compared to how we have chosen entering and leaving variables before, we should note that our earlier choice was guided by the desire to increase the objective function while preserving feasibility. Here, the dual dictionary is not even feasible, and so such considerations are meaningless. Once we give up those rules for the choice of entering and leaving variables, it is easy to see that a pivot can be performed with any choice of entering and leaving variables provided only that the coefficient on the entering variable in the constraint of the leaving variables does not vanish. Such is the case with the current choice. Hence, we do the pivot in both the primal and the dual. The result is

$$
\text{(P)} \qquad
\begin{array}{rcrcrcrcr}
\zeta & = & 9 & - & 5x_1 & + & 4x_2 & - & 3w_2 \\ \hline
w_1 & = & 1 & - & x_1 & - & 4x_2 & + & \\
x_3 & = & 3 & - & 3x_1 & + & x_2 & - & w_2
\end{array}
$$

$$
\text{(D)} \qquad
\begin{array}{rcrcrcr}
-\xi & = & -9 & - & 1y_1 & - & 3z_3 \\ \hline
z_1 & = & 5 & + & y_1 & + & 3z_3 \\
z_2 & = & -4 & + & 4y_1 & - & z_3 \\
y_2 & = & 3 & + & & + & z_3
\end{array}
$$

Note that these two dictionaries still have the property of being negative transposes of each other. For the next pivot, the entering variable in the primal dictionary is x_2 (this time there is no choice) and the leaving variable is w_1. In the dual dictionary, the corresponding entering variable is y_1 and the leaving variable is z_2. Doing the pivots, we get

$$
\text{(P)} \qquad
\begin{array}{rcrcrcrcr}
\zeta & = & 10 & - & 6x_1 & - & 1w_1 & - & 3w_2 \\ \hline
x_2 & = & 0.25 & - & 0.25x_1 & - & 0.25w_1 & & \\
x_3 & = & 3.25 & - & 3.25x_1 & - & 0.25w_1 & - & w_2
\end{array}
$$

$$
\text{(D)} \qquad
\begin{array}{rcrcrcr}
-\xi & = & -10 & - & 0.25z_2 & - & 3.25z_3 \\ \hline
z_1 & = & 6 & + & 0.25z_2 & + & 3.25z_3 \\
y_1 & = & 1 & + & 0.25z_2 & + & 0.25z_3 \\
y_2 & = & 3 & & & + & 1z_3
\end{array}
$$

This primal dictionary is optimal, since the coefficients in the objective row are all negative. Looking at the dual dictionary, we see that it is now feasible for the analogous reason. In fact, it is optimal too. Finally, both the primal and dual objective function values are 10.

The situation should now be clear. Given a linear programming problem, which is assumed to possess an optimal solution, first apply the Phase I procedure to get a basic feasible starting dictionary for Phase II. Then apply the simplex method to find an optimal solution. Each primal dictionary generated by the simplex method implicitly defines a corresponding dual dictionary as follows: first write down the negative transpose and then replace each x_j with a z_j and each w_i with a y_i. As long as the primal dictionary is not optimal, the implicitly defined dual dictionary will be infeasible. But once an optimal primal dictionary is found, the corresponding dual dictionary will be feasible. Since its objective coefficients are always nonpositive, this feasible dual dictionary is also optimal. Furthermore, at each iteration, the current primal objective function value coincides with the current dual objective function value.

To see why the negative-transpose property is preserved from one dictionary to the next, let us observe the effect of one pivot. To keep notations uncluttered, we consider only four generic entries in a table of coefficients: the pivot element, which we denote by a, one other element on the pivot element's row, call it b, one other in its column, call it c, and a fourth element, denoted d, chosen to make these four entries into a rectangle. A little thought (and perhaps some staring at the examples above) reveals that a pivot produces the following changes:

- the pivot element gets replaced by its reciprocal;
- elements in the pivot row get negated and divided by the pivot element;
- elements in the pivot column get divided by the pivot element; and
- all other elements, such as d, get decreased by bc/a.

These effects can be summarized on our generic table as follows:

$$\begin{array}{|c|c|c|}
\hline
 & & \\
\hline
b & a & \\
\hline
d & c & \\
\hline
\end{array}
\quad \xrightarrow{\text{pivot}} \quad
\begin{array}{|c|c|c|}
\hline
 & & \\
\hline
-\dfrac{b}{a} & \dfrac{1}{a} & \\
\hline
d - \dfrac{bc}{a} & \dfrac{c}{a} & \\
\hline
\end{array}$$

Now, if we start with a dual dictionary that is the negative transpose of the primal and apply one pivot operation, we get

$$\begin{array}{|c|c|c|}\hline & & \\ & -b & -d \\ & & \\ \hline & -a & -c \\ \hline & & \\ \hline \end{array} \quad \xrightarrow{\text{pivot}} \quad \begin{array}{|c|c|c|}\hline & & \\ & \dfrac{b}{a} & -d+\dfrac{bc}{a} \\ & & \\ \hline & -\dfrac{1}{a} & -\dfrac{c}{a} \\ \hline & & \\ \hline \end{array}.$$

Note that the resulting dual table is the negative transpose of the resulting primal table. By induction we then conclude that, if we start with this property, it will be preserved throughout the solution process.

Since the strong duality theorem is the most important theorem in this book, we present here a careful proof. Those readers who are satisfied with the above discussion may skip the proof.

PROOF OF THEOREM 5.2. It suffices to exhibit a dual feasible solution y^* satisfying (5.2). Suppose we apply the simplex method. We know that the simplex method produces an optimal solution whenever one exists, and we have assumed that one does indeed exist. Hence, the final dictionary will be an optimal dictionary for the primal problem. The objective function in this final dictionary is ordinarily written as

$$\zeta = \bar{\zeta} + \sum_{j \in \mathcal{N}} \bar{c}_j x_j.$$

But, since this is the optimal dictionary and we prefer stars to bars for denoting optimal "stuff," let us write ζ^* instead of $\bar{\zeta}$. Also, the collection of nonbasic variables will generally consist of a combination of original variables as well as slack variables. Instead of using $\bar{c}_j$ for the coefficients of these variables, let us use c_j^* for the objective coefficients corresponding to original variables, and let us use d_i^* for the objective coefficients corresponding to slack variables. Also, for those original variables that are basic we put $c_j^* = 0$, and for those slack variables that are basic we put $d_i^* = 0$. With these new notations, we can rewrite the objective function as

$$\zeta = \zeta^* + \sum_{j=1}^{n} c_j^* x_j + \sum_{i=1}^{m} d_i^* w_i.$$

As we know, ζ^* is the objective function value corresponding to the optimal primal solution:

$$\zeta^* = \sum_{j=1}^{n} c_j x_j^*. \tag{5.3}$$

Now, put

$$y_i^* = -d_i^*, \qquad i = 1, 2, \ldots, m. \tag{5.4}$$

We shall show that $y^* = (y_1^*, y_2^*, \ldots, y_m^*)$ is feasible for the dual problem and satisfies (5.2). To this end, we write the objective function in two ways:

$$\begin{aligned}
\sum_{j=1}^n c_j x_j &= \zeta^* + \sum_{j=1}^n c_j^* x_j + \sum_{i=1}^m d_i^* w_i \\
&= \zeta^* + \sum_{j=1}^n c_j^* x_j + \sum_{i=1}^m (-y_i^*) \left(b_i - \sum_{j=1}^n a_{ij} x_j \right) \\
&= \zeta^* - \sum_{i=1}^m b_i y_i^* + \sum_{j=1}^n \left(c_j^* + \sum_{i=1}^m y_i^* a_{ij} \right) x_j.
\end{aligned}$$

Since all these expressions are linear in the variables x_j, we can equate the coefficients of each variable appearing on the left-hand side with the corresponding coefficient appearing in the last expression on the right-hand side. We can also equate the constant terms on the two sides. Hence,

$$\zeta^* = \sum_{i=1}^m b_i y_i^* \tag{5.5}$$

$$c_j = c_j^* + \sum_{i=1}^m y_i^* a_{ij}, \qquad j = 1, 2, \ldots, n. \tag{5.6}$$

Combining (5.3) and (5.5), we get that (5.2) holds. Also, the optimality of the dictionary for the primal problem implies that each c_j^* is nonpositive, and hence we see from (5.6) that

$$\sum_{i=1}^m y_i^* a_{ij} \geq c_j, \qquad j = 1, 2, \ldots, n.$$

By the same reasoning, each d_i^* is nonpositive, and so we see from (5.4) that

$$y_i^* \geq 0, \qquad i = 1, 2, \ldots, m.$$

These last two sets of inequalities are precisely the conditions that guarantee dual feasibility. This completes the proof. $\square$

The strong duality theorem tells us that, whenever the primal problem has an optimal solution, the dual problem also has one and there is no duality gap. But what if the primal problem does not have an optimal solution? For example, suppose that it is unbounded. The unboundedness of the primal together with the weak duality theorem tells us immediately that the dual problem must be infeasible. Similarly, if the dual problem is unbounded, then the primal problem must be infeasible. It is natural to hope that these three cases are the only possibilities, because if they were we could then think of the strong duality theorem holding globally. That is, even if, say, the primal is unbounded, the fact that then the dual is infeasible is like saying that the primal and dual have a zero duality gap sitting out at $+\infty$. Similarly, an infeasible primal together with an unbounded dual could be viewed as a pair in which the gap is zero and sits at $-\infty$.

But it turns out that there is a fourth possibility that sometimes occurs—it can happen that both the primal and the dual problems are infeasible. For example, consider the following problem:

$$\begin{array}{ll} \text{maximize} & 2x_1 - x_2 \\ \text{subject to} & \begin{aligned} x_1 - x_2 &\le 1 \\ -x_1 + x_2 &\le -2 \\ x_1,\, x_2 &\ge 0. \end{aligned} \end{array}$$

It is easy to see that both this problem and its dual are infeasible. For these problems, one can think of there being a huge duality gap extending from $-\infty$ to $+\infty$.

Duality theory is often useful in that it provides a *certificate of optimality*. For example, suppose that you were asked to solve a really huge and difficult linear program. After spending weeks or months at the computer, you are finally able to get the simplex method to solve the problem, producing as it does an optimal dual solution y^* in addition to the optimal primal solution x^*. Now, how are you going to convince your boss that your solution is correct? Do you really want to ask her to verify the correctness of your computer programs? The answer is probably not. And in fact it is not necessary. All you need to do is supply the primal and the dual solution, and she only has to check that the primal solution is feasible for the primal problem (that is easy), the dual solution is feasible for the dual problem (that is just as easy), and the primal and dual objective values agree (and that is even easier). Certificates of optimality have also been known to dramatically reduce the amount of time certain underpaid professors have to devote to grading homework assignments!

As we have seen, the simplex method applied to a primal problem actually solves both the primal and the dual. Since the dual of the dual is the primal, applying the simplex method to the dual also solves both the primal and the dual problem. Sometimes it is easier to apply the simplex method to the dual, for example, if the dual has an obvious basic feasible solution but the primal does not. We take up this topic in the next chapter.

5. Complementary Slackness

Sometimes it is necessary to recover an optimal dual solution when only an optimal primal solution is known. The following theorem, known as the *complementary slackness theorem*, can help in this regard.

THEOREM 5.3. *Suppose that $x = (x_1, x_2, \ldots, x_n)$ is primal feasible and that $y = (y_1, y_2, \ldots, y_m)$ is dual feasible. Let $(w_1, w_2, \ldots, w_m)$ denote the corresponding primal slack variables, and let $(z_1, z_2, \ldots, z_n)$ denote the corresponding dual slack variables. Then x and y are optimal for their respective problems if and only if*

$$\begin{aligned} x_j z_j &= 0, \qquad \text{for } j = 1, 2, \ldots, n, \\ w_i y_i &= 0, \qquad \text{for } i = 1, 2, \ldots, m. \end{aligned} \tag{5.7}$$

PROOF. We begin by revisiting the chain of inequalities used to prove the weak duality theorem:

$$\sum_j c_j x_j \le \sum_j \left(\sum_i y_i a_{ij} \right) x_j \tag{5.8}$$

$$= \sum_i \left(\sum_j a_{ij} x_j \right) y_i$$

$$\le \sum_i b_i y_i. \tag{5.9}$$

Recall that the first inequality arises from the fact that each term in the left-hand sum is dominated by the corresponding term in the right-hand sum. Furthermore, this domination is a consequence of the fact that each x_j is nonnegative and

$$c_j \le \sum_i y_i a_{ij}.$$

Hence, inequality (5.8) will be an equality if and only if, for every $j = 1, 2, \ldots, n$, either $x_j = 0$ or $c_j = \sum_i y_i a_{ij}$. But since

$$z_j = \sum_i y_i a_{ij} - c_j,$$

we see that the alternative to $x_j = 0$ is simply that $z_j = 0$. Of course, the statement that at least one of these two numbers vanishes can be succinctly expressed by saying that the product vanishes.

An analogous analysis of inequality (5.9) shows that it is an equality if and only if (5.7) holds. This then completes the proof. □

Suppose that we have a nondegenerate primal basic optimal solution

$$x^* = (x_1^*, x_2^*, \ldots, x_n^*)$$

and we wish to find a corresponding optimal solution for the dual. Let

$$w^* = (w_1^*, w_2^*, \ldots, w_m^*)$$

denote the corresponding slack variables, which were probably given along with the x_j's but if not can be easily obtained from their definition as slack variables:

$$w_i^* = b_i - \sum_j a_{ij} x_j^*.$$

The dual constraints are

$$\sum_i y_i a_{ij} - z_j = c_j, \qquad j = 1, 2, \ldots, n, \tag{5.10}$$

where we have written the inequalities in equality form by introducing slack variables z_j, $j = 1, 2, \ldots, n$. These constraints form n equations in $m + n$ unknowns. But the basic optimal solution (x^*, w^*) is a collection of $n + m$ variables, many of which are positive. In fact, since the primal solution is assumed to be nondegenerate, it follows that the m basic variables will be strictly positive. The complementary slackness theorem then tells us that the corresponding dual variables must vanish. Hence, of the $m + n$ variables in (5.10), we can set m of them to zero. We are then left with just n equations in n unknowns, which we would expect to have a unique solution that can be solved for. If there is a unique solution, all the components should be nonnegative. If any of them are negative, this would stand in contradiction to the assumed optimality of x^*.

6. The Dual Simplex Method

In this section, we study what happens if we apply the simplex method to the dual problem. As we saw in our discussion of the strong duality theorem, one can actually apply the simplex method to the dual problem without ever writing down the dual problem or its dictionaries. Instead, the so-called dual simplex method is seen simply as a new way of picking the entering and leaving variables in a sequence of primal dictionaries.

We begin with an example:

$$\begin{array}{lrcr}
\text{maximize} & -x_1 - x_2 & & \\
\text{subject to} & -2x_1 - x_2 & \leq & 4 \\
& -2x_1 + 4x_2 & \leq & -8 \\
& -x_1 + 3x_2 & \leq & -7 \\
& x_1, \; x_2 & \geq & 0.
\end{array}$$

The dual of this problem is

$$\begin{array}{lrcr}
\text{minimize} & 4y_1 - 8y_2 - 7y_3 & & \\
\text{subject to} & -2y_1 - 2y_2 - y_3 & \geq & -1 \\
& -y_1 + 4y_2 + 3y_3 & \geq & -1 \\
& y_1, \; y_2, \; y_3 & \geq & 0.
\end{array}$$

Introducing variables w_i, $i = 1, 2, 3$, for the primal slacks and z_j, $j = 1, 2$, for the dual slacks, we can write down the initial primal and dual dictionaries:

$$
\text{(P)} \qquad
\begin{array}{rcrrrrr}
\zeta & = & & - & 1\,x_1 & - & 1\,x_2 \\
\hline
w_1 & = & 4 & + & 2\,x_1 & + & x_2 \\
w_2 & = & -8 & + & 2\,x_1 & - & 4\,x_2 \\
w_3 & = & -7 & + & x_1 & - & 3\,x_2
\end{array}
$$

$$
\text{(D)} \qquad
\begin{array}{rcrrrrrrr}
-\xi & = & & - & 4\,y_1 & + & 8\,y_2 & + & 7\,y_3 \\
\hline
z_1 & = & 1 & - & 2\,y_1 & - & 2\,y_2 & - & y_3 \\
z_2 & = & 1 & - & y_1 & + & 4\,y_2 & + & 3\,y_3
\end{array}
$$

As before, we have recorded the negative of the dual objective function, since we prefer to maximize the objective function appearing in a dictionary. More importantly, note that the dual dictionary is feasible, whereas the primal one is not. This suggests that it would be sensible to apply the simplex method to the dual. Let us do so, but as we go we keep track of the analogous pivots applied to the primal dictionary. For example, the entering variable in the initial dual dictionary is y_2, and the leaving variable then is z_1. Since w_2 is complementary to y_2 and x_1 is complementary to z_1, we will use w_2 and x_1 as the entering/leaving variables in the primal dictionary. Of course, since w_2 is basic and x_1 is nonbasic, w_2 must be the leaving variable and x_1 the entering variable—i.e., the reverse of what we have for the complementary variables in the dual dictionary. The result of these pivots is

$$
\text{(P)} \qquad
\begin{array}{rcrrrrr}
\zeta & = & -4 & - & 0.5\,w_2 & - & 3\,x_2 \\
\hline
w_1 & = & 12 & + & w_2 & + & 5\,x_2 \\
x_1 & = & 4 & + & 0.5\,w_2 & + & 2\,x_2 \\
w_3 & = & -3 & + & 0.5\,w_2 & - & x_2
\end{array}
$$

$$
\text{(D)} \qquad
\begin{array}{rcrrrrrrr}
-\xi & = & 4 & - & 12\,y_1 & - & 4\,z_1 & + & 3\,y_3 \\
\hline
y_2 & = & 0.5 & - & 1\,y_1 & - & 0.5\,z_1 & - & 0.5\,y_3 \\
z_2 & = & 3 & - & 5\,y_1 & - & 2\,z_1 & + & 1\,y_3
\end{array}
$$

Continuing to work on the dual, we now see that y_3 is the entering variable and y_2 leaves. Hence, for the primal we use w_3 and w_2 as the leaving and entering variable, respectively. After pivoting, we have

$$
\text{(P)} \qquad \begin{array}{rcrcrcrcr}
\zeta & = & -7 & - & 1 & w_3 & - & 4 & x_2 \\
\hline
w_1 & = & 18 & + & 2 & w_3 & + & 7 & x_2 \\
x_1 & = & 7 & + & & w_3 & + & 3 & x_2 \\
w_2 & = & 6 & + & 2 & w_3 & + & 2 & x_2
\end{array}
$$

$$
\text{(D)} \qquad \begin{array}{rcrcrcrcrcr}
-\xi & = & 7 & - & 18 & y_1 & - & 7 & z_1 & - & 6 & y_2 \\
\hline
y_3 & = & 1 & - & 2 & y_1 & - & & z_1 & - & 2 & y_2 \\
z_2 & = & 4 & - & 7 & y_1 & - & 3 & z_1 & - & 2 & y_2
\end{array}
$$

Now we notice that both dictionaries are optimal.

Of course, in each of the above dictionaries, the table of numbers in each dual dictionary is the negative transpose of the corresponding primal table. Therefore, we never need to write the dual dictionary; the dual simplex method can be entirely described in terms of the primal dictionaries. Indeed, first we note that the dictionary must be dual feasible. This means that all the coefficients of the nonbasic variables in the primal objective function must be nonpositive. Given this, we proceed as follows. First we select the leaving variable by picking that basic variable whose constant term in the dictionary is the most negative (if there are none, then the current dictionary is optimal). Then we pick the entering variable by scanning across this row of the dictionary and comparing ratios of the coefficients in this row to the corresponding coefficients in the objective row, looking for the largest negated ratio just as we did in the primal simplex method. Once the entering and leaving variable are identified, we pivot to the next dictionary and continue from there. The reader is encouraged to trace the pivots in the above example, paying particular attention to how one determines the entering and leaving variables by looking only at the primal dictionary.

7. A Dual-Based Phase I Algorithm

The dual simplex method described in the previous section provides us with a new Phase I algorithm, which if nothing else is at least more elegant than the one we gave in Chapter 2. Let us illustrate it using an example:

$$
\begin{array}{lrcr}
\text{maximize} & -x_1 + 4x_2 & & \\
\text{subject to} & -2x_1 - x_2 & \leq & 4 \\
& -2x_1 + 4x_2 & \leq & -8 \\
& -x_1 + 3x_2 & \leq & -7 \\
& x_1,\ x_2 & \geq & 0.
\end{array}
$$

The primal dictionary for this problem is

$$\text{(P)}\qquad \begin{array}{rcrrrr} \zeta & = & & -\ 1\,x_1 & +\ 4\,x_2 \\ \hline w_1 & = & 4 & +\ 2\,x_1 & +\ \ \ x_2 \\ w_2 & = & -8 & +\ 2\,x_1 & -\ 4\,x_2 \\ w_3 & = & -7 & +\ \ \ x_1 & -\ 3\,x_2 \end{array}$$

and even though at this point we realize that we do not need to look at the dual dictionary, let us track it anyway:

$$\text{(D)}\qquad \begin{array}{rcrrrr} -\xi & = & & -\ 4\,y_1 & +\ 8\,y_2 & +\ 7\,y_3 \\ \hline z_1 & = & 1 & -\ 2\,y_1 & -\ 2\,y_2 & -\ \ \ y_3 \\ z_2 & = & -4 & -\ \ \ y_1 & +\ 4\,y_2 & +\ 3\,y_3 \end{array}$$

Clearly, neither the primal nor the dual dictionary is feasible. But by changing the primal objective function, we can easily produce a dual feasible dictionary. For example, let us temporarily change the primal objective function to

$$\eta = -x_1 - x_2.$$

Then the corresponding initial dual dictionary is feasible. In fact, it coincides with the dual dictionary we considered in the previous section, so we already know the optimal solution for this modified problem. The optimal primal dictionary is

$$\begin{array}{rcrrr} \eta & = & -7 & -\ 1\,w_3 & -\ 4\,x_2 \\ \hline w_1 & = & 18 & +\ 2\,w_3 & +\ 7\,x_2 \\ x_1 & = & 7 & +\ \ \ w_3 & +\ 3\,x_2 \\ w_2 & = & 6 & +\ 2\,w_3 & +\ 2\,x_2 \end{array}$$

This primal dictionary is optimal for the modified problem but not for the original problem. However, it is feasible for the original problem, and we can now simply reinstate the intended objective function and continue with Phase II. Indeed,

$$\begin{aligned} \zeta &= -x_1 + 4x_2 \\ &= -(7 + w_3 + 3x_2) + 4x_2 \\ &= -7 - w_3 + x_2. \end{aligned}$$

Hence, the starting dictionary for Phase II is

$$
\begin{array}{rcrcrcr}
\zeta & = & -7 & - & 1\,w_3 & + & 1\,x_2 \\
\hline
w_1 & = & 18 & + & 2\,w_3 & + & 7\,x_2 \\
x_1 & = & 7 & + & w_3 & + & 3\,x_2 \\
w_2 & = & 6 & + & 2\,w_3 & + & 2\,x_2
\end{array}
$$

The entering variable is x_2. Looking for a leaving variable, we discover that this problem is unbounded. Of course, more typically one would expect to have to do several iterations of Phase II to find the optimal solution (or show unboundedness). Here we just got lucky that the game ended so soon.

It is interesting to note how we detect infeasibility with this new Phase I algorithm. The modified problem is guaranteed always to be dual feasible. It is easy to see that the primal problem is infeasible if and only if the modified problem is dual unbounded (which the dual simplex method will detect just as the primal simplex method detects primal unboundedness).

The two-phase algorithm we have just presented can be thought of as a dual–primal algorithm, since we first apply the dual simplex method to a modified dual feasible problem and then finish off by applying the primal simplex method to the original problem, starting from the feasible dictionary produced by Phase I. One could consider turning this around and doing a primal–dual two-phase algorithm. Here, the right-hand side of the primal problem would be modified to produce an obvious primal feasible solution. The primal simplex method would then be applied. The optimal solution to this primal problem will then be feasible for the original dual problem but will not be optimal for it. But then the dual simplex method can be applied, starting with this dual feasible basis until an optimal solution for the dual problem is obtained.

8. The Dual of a Problem in General Form

In Chapter 1, we saw that linear programming problems can be formulated in a variety of ways. In this section, we derive the form of the dual when the primal problem is not necessarily presented in standard form.

First, let us consider the case where the linear constraints are equalities (and the variables are nonnegative):

$$
\begin{array}{lrll}
& \text{maximize} & \displaystyle\sum_{j=1}^{n} c_j x_j & \\
(5.11) & \text{subject to} & \displaystyle\sum_{j=1}^{n} a_{ij} x_j = b_i & i = 1, 2, \ldots, m \\
& & x_j \geq 0 & j = 1, 2, \ldots, n.
\end{array}
$$

As we mentioned in Chapter 1, this problem can be reformulated with inequality constraints by simply writing each equality as two inequalities: one greater-than-or-equal-to and one less-than-or-equal-to:

$$\begin{array}{lll} \text{maximize} & \displaystyle\sum_{j=1}^{n} c_j x_j & \\ \text{subject to} & \displaystyle\sum_{j=1}^{n} a_{ij} x_j \leq b_i & i = 1, 2, \ldots, m \\ & \displaystyle\sum_{j=1}^{n} a_{ij} x_j \geq b_i & i = 1, 2, \ldots, m \\ & x_j \geq 0 & j = 1, 2, \ldots, n. \end{array}$$

Then negating each greater-than-or-equal-to constraint, we can put the problem into standard form:

$$\begin{array}{lll} \text{maximize} & \displaystyle\sum_{j=1}^{n} c_j x_j & \\ \text{subject to} & \displaystyle\sum_{j=1}^{n} a_{ij} x_j \leq b_i & i = 1, 2, \ldots, m \\ & \displaystyle\sum_{j=1}^{n} -a_{ij} x_j \leq -b_i & i = 1, 2, \ldots, m \\ & x_j \geq 0 & j = 1, 2, \ldots, n. \end{array}$$

Now that the problem is in standard form, we can write down its dual. Since there are two sets of m inequality constraints, we need two sets of m dual variables. Let us denote the dual variables associated with the first set of m constraints by y_i^+, $i = 1, 2, \ldots, m$, and the remaining dual variables by y_i^-, $i = 1, 2, \ldots, m$. With these notations, the dual problem is

$$\begin{array}{lll} \text{minimize} & \displaystyle\sum_{i=1}^{m} b_i y_i^+ - \sum_{i=1}^{m} b_i y_i^- & \\ \text{subject to} & \displaystyle\sum_{i=1}^{m} y_i^+ a_{ij} - \sum_{i=1}^{m} y_i^- a_{ij} \geq c_j & j = 1, 2, \ldots, n \\ & y_i^+, y_i^- \geq 0 & i = 1, 2, \ldots, m. \end{array}$$

A moment's reflection reveals that we can simplify this problem. If we put

$$y_i = y_i^+ - y_i^-, \qquad i = 1, 2, \ldots, m,$$

Primal	Dual
Equality constraint	Free variable
Inequality constraint	Nonnegative variable
Free variable	Equality constraint
Nonnegative variable	Inequality constraint

TABLE 5.1. Rules for forming the dual.

the dual problem reduces to

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^{m} b_i y_i \\ \text{subject to} & \sum_{i=1}^{m} y_i a_{ij} \geq c_j \qquad j = 1, 2, \ldots, n. \end{array}$$

This problem is *the* dual associated with (5.11). Note what has changed from when we were considering problems in standard form: now the dual variables are not restricted to be nonnegative. And that is the message: *equality constraints in the primal yield unconstrained variables (also referred to as free variables) in the dual, whereas inequality constraints in the primal yield nonnegative variables in the dual*. Employing the symmetry between the primal and the dual, we can say more: *free variables in the primal yield equality constraints in the dual, whereas nonnegative variables in the primal yield inequality constraints in the dual*. These rules are summarized in Table 5.1.

9. Resource Allocation Problems

Let us return to the production facility problem studied in Chapter 1. Recall that this problem involves a production facility that can take a variety of raw materials (enumerated $i = 1, 2, \ldots, m$) and turn them into a variety of final products (enumerated $j = 1, 2, \ldots, n$). We assume as before that the current market value of a unit of the ith raw material is ρ_i, that the current market price for a unit of the jth product is σ_j, that producing one unit of product j requires a_{ij} units of raw material i, and that at the current moment in time the facility has on hand b_i units of the ith raw material.

The current market values/prices are, by definition, related to each other by the formulas

$$\sigma_j = \sum_i \rho_i a_{ij}, \qquad j = 1, 2, \ldots, n.$$

These equations hold whenever the market is in equilibrium. (Of course, it is crucial to assume here that the collection of "raw materials" appearing on the right-hand side is exhaustive, including such items as depreciation of fixed assets and physical labor.) In the real world, the market is always essentially in equilibrium. Nonetheless, it continually experiences small perturbations that ripple through it and move the equilibrium to new levels.

These perturbations can be from several causes, an important one being innovation. One possible innovation is to improve the production process. This means that the values of some of the a_{ij}'s are reduced. Now, suddenly there is a windfall profit for each unit of product j produced. This windfall profit is given by

$$c_j = \sigma_j - \sum_i \rho_i a_{ij}. \tag{5.12}$$

Of course, eventually most producers of these products will take advantage of the same innovation, and once the suppliers get wind of the profits being made, they will get in on the action by raising the price of the raw materials.[1] Nonetheless, there is always a time lag; it is during this time that fortunes are made.

To be concrete, let us assume that the time lag is about 1 month (depending on the industry, this lag time could be considered too short or too long). Suppose also that the production manager decides to produce x_j units of product j and that all units produced are sold immediately at their market value. Then the total revenue during this month will be $\sum_j \sigma_j x_j$. The value of the raw materials on hand at the beginning of the month was $\sum_i \rho_i b_i$. Also, if we denote the new price levels for the raw materials at the end of the month by w_i, $i = 1, 2, \ldots, m$, then the value of any remaining inventory at the end of the month is given by

$$\sum_i w_i \left(b_i - \sum_j a_{ij} x_j \right)$$

(if any term is negative, then it represents the cost of purchasing additional raw materials to meet the month's production requirements—we assume that these additional purchases are made at the new, higher, end-of-month price). The total windfall, call it π, (over all products) for this month can now be written as

$$\pi = \sum_j \sigma_j x_j + \sum_i w_i \left(b_i - \sum_j a_{ij} x_j \right) - \sum_i \rho_i b_i. \tag{5.13}$$

Our aim is to choose production levels x_j, $j = 1, 2, \ldots, n$, that maximize this windfall. But our supplier's aim is to choose prices w_i, $i = 1, 2, \ldots, m$, so as to minimize our windfall. Before studying these optimizations, let us first rewrite the

[1] One could take the prices of raw materials as fixed and argue that the value of the final products will fall. It does not really matter which view one adopts, since prices are relative anyway. The point is simply that the difference between the price of the raw materials and the price of the final products must narrow due to this innovation.

windfall in a more convenient form. As in Chapter 1, let y_i denote the increase in the price of raw material i. That is,

$$w_i = \rho_i + y_i. \tag{5.14}$$

Substituting (5.14) into (5.13) and then simplifying notations using (5.12), we see that

$$\pi = \sum_j c_j x_j + \sum_i y_i \left(b_i - \sum_j a_{ij} x_j \right). \tag{5.15}$$

To emphasize that π depends on each of the x_j's and on the y_i's, we sometimes write it as $\pi(x_1, \ldots, x_n, y_1, \ldots, y_m)$.

Now let us return to the competing optimizations. Given x_j for $j = 1, 2, \ldots, n$, the suppliers react to minimize $\pi(x_1, \ldots, x_n, y_1, \ldots, y_m)$. Looking at (5.15), we see that for any resource i in short supply, that is,

$$b_i - \sum_j a_{ij} x_j < 0,$$

the suppliers will jack up the price immensely (i.e., $y_i = \infty$). To avoid this obviously bad situation, the production manager will be sure to set the production levels so that

$$\sum_j a_{ij} x_j \le b_i, \qquad i = 1, 2, \ldots, m.$$

On the other hand, for any resource i that is not exhausted during the windfall month, that is,

$$b_i - \sum_j a_{ij} x_j > 0,$$

the suppliers will have no incentive to change the prevailing market price (i.e., $y_i = 0$). Therefore, from the production manager's point of view, the problem reduces to one of maximizing

$$\sum_j c_j x_j$$

subject to the constraints that

$$\begin{aligned} \sum_j a_{ij} x_j &\le b_i, & i &= 1, 2, \ldots, m, \\ x_j &\ge 0, & j &= 1, 2, \ldots, n. \end{aligned}$$

This is just our usual primal linear programming problem. This is the problem that the production manager needs to solve in anticipation of adversarial suppliers.

Now let us look at the problem from the suppliers' point of view. Rearranging the terms in (5.15) by writing

$$\pi = \sum_j \left(c_j - \sum_i y_i a_{ij} \right) x_j + \sum_i y_i b_i, \tag{5.16}$$

we see that if the suppliers set prices in such a manner that a windfall remains on the jth product even after the price adjustment, that is,

$$c_j - \sum_i y_i a_{ij} > 0,$$

then the production manager would be able to generate for the facility an arbitrarily large windfall by producing a huge amount of the jth product (i.e., $x_j = \infty$). We assume that this is unacceptable to the suppliers, and so they will determine their price increases so that

$$\sum_i y_i a_{ij} \geq c_j, \qquad j = 1, 2, \ldots, n.$$

Also, if the suppliers set the price increases too high so that the production facility will lose money by producing product j, that is,

$$c_j - \sum_i y_i a_{ij} < 0,$$

then the production manager would simply decide not to engage in that activity. That is, she would set $x_j = 0$. Hence, the first term in (5.16) will always be zero, and so the optimization problem faced by the suppliers is to minimize

$$\sum_i b_i y_i$$

subject to the constraints that

$$\begin{aligned} \sum_i y_i a_{ij} &\geq c_j, \qquad j = 1, 2, \ldots, n, \\ y_i &\geq 0, \qquad i = 1, 2, \ldots, m. \end{aligned}$$

This is precisely the dual of the production manager's problem!

As we have seen earlier with the strong duality theorem, if the production manager's problem has an optimal solution, then so does the suppliers' problem, and the two objectives agree. This means than an equilibrium can be reestablished by setting the production levels and the price hikes according to the optimal solutions to these two linear programming problems.

10. Lagrangian Duality

The analysis of the preceding section is an example of a general technique that forms the foundation of a subject called *Lagrangian duality*, which we shall briefly describe.

Let us start by summarizing what we did. It was quite simple. The analysis revolved around a function

$$\pi(x_1, \ldots, x_n, y_1, \ldots, y_m) = \sum_j c_j x_j - \sum_i \sum_j y_i a_{ij} x_j + \sum_i y_i b_i.$$

To streamline notations, let x stand for the entire collection of variables $x_1, x_2, \ldots, x_n$ and let y stand for the collection of y_i's so that we can write $\pi(x, y)$ in place of $\pi(x_1, \ldots, x_n, y_1, \ldots, y_m)$. Written with these notations, we showed in the previous section that

$$\max_{x\geq 0} \min_{y\geq 0} \pi(x, y) = \min_{y\geq 0} \max_{x\geq 0} \pi(x, y).$$

We also showed that the inner optimization could in both cases be solved explicitly, that the max–min problem reduced to a linear programming problem, and that the min–max problem reduced to the dual linear programming problem.

One could imagine trying to carry out the same program for functions π that do not necessarily have the form shown above. In the general case, one needs to consider each step carefully. The max–min problem is called the primal problem, and the min–max problem is called the dual problem. However, it may or may not be true that these two problems have the same optimal objective values. In fact, the subject is interesting because one can indeed state specific, verifiable conditions for which the two problems do agree. Also, one would like to be able to solve the inner optimizations explicitly so that the primal problem can be stated as a pure maximization problem and the dual can be stated as a pure minimization problem. This, too, is often doable. There are various ways in which one can extend the notions of duality beyond the context of linear programming. The one just described is referred to as Lagrangian duality. It is perhaps the most important such extension.

Exercises

In solving the following problems, the advanced pivot tool can be used to check your arithmetic:

vanderbei.princeton.edu/JAVA/pivot/advanced.html

5.1 What is the dual of the following linear programming problem:

$$\begin{array}{lrcl}
\text{maximize} & x_1 - 2x_2 & & \\
\text{subject to} & x_1 + 2x_2 - x_3 + x_4 & \geq & 0 \\
& 4x_1 + 3x_2 + 4x_3 - 2x_4 & \leq & 3 \\
& -x_1 - x_2 + 2x_3 + x_4 & = & 1 \\
& x_2,\ x_3 & \geq & 0\ .
\end{array}$$

5.2 Illustrate Theorem 5.2 on the problem in Exercise 2.9.

5.3 Illustrate Theorem 5.2 on the problem in Exercise 2.1.

5.4 Illustrate Theorem 5.2 on the problem in Exercise 2.2.

5.5 Consider the following linear programming problem:

$$
\begin{array}{lrcl}
\text{maximize} & 2x_1 + 8x_2 - x_3 - 2x_4 & & \\
\text{subject to} & 2x_1 + 3x_2 \qquad + 6x_4 & \leq & 6 \\
& -2x_1 + 4x_2 + 3x_3 \qquad & \leq & 1.5 \\
& 3x_1 + 2x_2 - 2x_3 - 4x_4 & \leq & 4 \\
& x_1,\ x_2,\ x_3,\ x_4 & \geq & 0.
\end{array}
$$

Suppose that, in solving this problem, you have arrived at the following dictionary:

$$
\begin{array}{rcrrrrr}
\zeta & = & 3.5 & -\ 0.25\ w_1 & +\ 6.25\ x_2 & -\ 0.5\ w_3 & -\ 1.5\ x_4 \\
\hline
x_1 & = & 3 & -\ 0.5\ w_1 & -\ 1.5\ x_2 & & -\ 3.0\ x_4 \\
w_2 & = & 0 & +\ 1.25\ w_1 & -\ 3.25\ x_2 & -\ 1.5\ w_3 & +\ 13.5\ x_4 \\
x_3 & = & 2.5 & -\ 0.75\ w_1 & -\ 1.25\ x_2 & +\ 0.5\ w_3 & -\ 6.5\ x_4\ .
\end{array}
$$

(a) Write down the dual problem.
(b) In the dictionary shown above, which variables are basic? Which are nonbasic?
(c) Write down the primal solution corresponding to the given dictionary. Is it feasible? Is it degenerate?
(d) Write down the corresponding dual dictionary.
(e) Write down the dual solution. Is it feasible?
(f) Do the primal/dual solutions you wrote above satisfy the complementary slackness property?
(g) Is the current primal solution optimal?
(h) For the next (primal) pivot, which variable will enter if the largest-coefficient rule is used? Which will leave? Will the pivot be degenerate?

5.6 Solve the following linear program:

$$
\begin{array}{lrcr}
\text{maximize} & -x_1 - 2x_2 & & \\
\text{subject to} & -2x_1 + 7x_2 & \leq & 6 \\
& -3x_1 + \ \ x_2 & \leq & -1 \\
& 9x_1 - 4x_2 & \leq & 6 \\
& x_1 - \ \ x_2 & \leq & 1 \\
& 7x_1 - 3x_2 & \leq & 6 \\
& -5x_1 + 2x_2 & \leq & -3 \\
& x_1,\ x_2 & \geq & 0.
\end{array}
$$

5.7 Solve the linear program given in Exercise 2.3 using the dual–primal two-phase algorithm.

5.8 Solve the linear program given in Exercise 2.4 using the dual–primal two-phase algorithm.

5.9 Solve the linear program given in Exercise 2.6 using the dual–primal two-phase algorithm.

5.10 Using today's date (MMYY) for the seed value, solve 10 problems using the dual phase I primal phase II simplex method:

vanderbei.princeton.edu/JAVA/pivot/dp2phase.html

5.11 Using today's date (MMYY) for the seed value, solve 10 problems using the primal phase I dual phase II simplex method:

vanderbei.princeton.edu/JAVA/pivot/pd2phase.html

5.12 For x and y in $\mathbb{R}$, compute

$$\max_{x\geq 0}\ \min_{y\geq 0}\ (x-y) \quad \text{and} \quad \min_{y\geq 0}\ \max_{x\geq 0}\ (x-y)$$

and note whether or not they are equal.

5.13 Consider the following process. Starting with a linear programming problem in standard form,

$$\begin{array}{lll} \text{maximize} & \displaystyle\sum_{j=1}^{n} c_j x_j & \\ \text{subject to} & \displaystyle\sum_{j=1}^{n} a_{ij} x_j \leq b_i & i = 1, 2, \ldots, m \\ & x_j \geq 0 & j = 1, 2, \ldots, n, \end{array}$$

first form its dual:

$$\begin{array}{lll} \text{minimize} & \displaystyle\sum_{i=1}^{m} b_i y_i & \\ \text{subject to} & \displaystyle\sum_{i=1}^{m} y_i a_{ij} \geq c_j & j = 1, 2, \ldots, n \\ & y_i \geq 0 & i = 1, 2, \ldots, m. \end{array}$$

Then replace the minimization in the dual with a maximization to get a new linear programming problem, which we can write in standard form as follows:

$$\begin{aligned} \text{maximize} \quad & \sum_{i=1}^{m} b_i y_i \\ \text{subject to} \quad & \sum_{i=1}^{m} -y_i a_{ij} \leq -c_j \qquad j = 1, 2, \ldots, n \\ & \qquad\qquad y_i \geq 0 \qquad\qquad i = 1, 2, \ldots, m. \end{aligned}$$

If we identify a linear programming problem with its data, (a_{ij}, b_i, c_j), the above process can be thought of as a transformation T on the space of data defined by

$$(a_{ij}, b_i, c_j) \xrightarrow{T} (-a_{ji}, -c_j, b_i).$$

Let $\zeta^*(a_{ij}, b_i, c_j)$ denote the optimal objective function value of the standard-form linear programming problem having data (a_{ij}, b_i, c_j). By strong duality together with the fact that a maximization dominates a minimization, it follows that

$$\zeta^*(a_{ij}, b_i, c_j) \leq \zeta^*(-a_{ji}, -c_j, b_i).$$

Now if we repeat this process, we get

$$\begin{aligned} (a_{ij}, b_i, c_j) & \xrightarrow{T} (-a_{ji}, -c_j, b_i) \\ & \xrightarrow{T} (a_{ij}, -b_i, -c_j) \\ & \xrightarrow{T} (-a_{ji}, c_j, -b_i) \\ & \xrightarrow{T} (a_{ij}, b_i, c_j) \end{aligned}$$

and hence that

$$\begin{aligned} \zeta^*(a_{ij}, b_i, c_j) & \leq \zeta^*(-a_{ji}, -c_j, b_i) \\ & \leq \zeta^*(a_{ij}, -b_i, -c_j) \\ & \leq \zeta^*(-a_{ji}, c_j, -b_i) \\ & \leq \zeta^*(a_{ij}, b_i, c_j). \end{aligned}$$

But the first and the last entry in this chain of inequalities are equal. Therefore, all these inequalities would seem to be equalities. While this outcome could happen sometimes, it certainly is not always true. What is the error in this logic? Can you state a (correct) nontrivial theorem that follows from this line of reasoning? Can you give an example where the four inequalities are indeed all equalities?

5.14 Consider the following variant of the resource allocation problem:

$$
\begin{aligned}
&\text{maximize} && \sum_{j=1}^{n} c_j x_j \\
&\text{subject to} && \sum_{j=1}^{n} a_{ij} x_j \le b_i && i = 1, 2, \ldots, m \\
&&& 0 \le x_j \le u_j && j = 1, 2, \ldots, n.
\end{aligned}
\tag{5.17}
$$

As usual, the c_j's denote the unit prices for the products and the b_i's denote the number of units on hand for each raw material. In this variant, the u_j's denote upper bounds on the number of units of each product that can be sold at the set price. Now, let us assume that the raw materials have not been purchased yet and it is part of the problem to determine the b_i's. Let p_i, $i = 1, 2, \ldots, m$ denote the price for raw material i. The problem then becomes an optimization over both the x_j's and the b_i's:

$$
\begin{aligned}
&\text{maximize} && \sum_{j=1}^{n} c_j x_j - \sum_{i=1}^{m} p_i b_i \\
&\text{subject to} && \sum_{j=1}^{n} a_{ij} x_j - b_i \le 0 && i = 1, 2, \ldots, m \\
&&& 0 \le x_j \le u_j && j = 1, 2, \ldots, n \\
&&& b_i \ge 0 && i = 1, 2, \ldots, m.
\end{aligned}
$$

(a) Show that this problem always has an optimal solution.

(b) Let $y_i^*(b)$, $i = 1, 2, \ldots, m$, denote optimal dual variables for the original resource allocation problem (5.17). Note that we have explicitly indicated that these dual variables depend on the b's. Also, we assume that problem (5.17) is both primal and dual nondegenerate so the $y_i^*(b)$ is uniquely defined. Show that the optimal value of the b_i's, call them b_i^*'s, satisfy

$$y_i^*(b^*) = p_i.$$

Hint: You will need to use the fact that, for resource allocation problems, we have $a_{ij} \ge 0$ for all i, and all j.

5.15 Consider the following linear program:

$$
\begin{aligned}
&\text{maximize} && \sum_{j=1}^{n} p_j x_j \\
&\text{subject to} && \sum_{j=1}^{n} q_j x_j \le \beta \\
&&& x_j \le 1 && j = 1, 2, \ldots, n \\
&&& x_j \ge 0 && j = 1, 2, \ldots, n.
\end{aligned}
$$

Here, the numbers p_j, $j = 1, 2, \ldots, n$ are positive and sum to one. The same is true of the q_j's:

$$\sum_{j=1}^{n} q_j = 1$$
$$q_j > 0.$$

Furthermore, assume that

$$\frac{p_1}{q_1} < \frac{p_2}{q_2} < \cdots < \frac{p_n}{q_n}$$

and that the parameter β is a small positive number. Let $k = \min\{j : q_{j+1} + \cdots + q_n \leq \beta\}$. Let y_0 denote the dual variable associated with the constraint involving β, and let y_j denote the dual variable associated with the upper bound of 1 on variable x_j. Using duality theory, show that the optimal values of the primal and dual variables are given by

$$x_j = \begin{cases} 0 & j < k \\ \frac{\beta - q_{k+1} - \cdots - q_n}{q_k} & j = k \\ 1 & j > k \end{cases}$$

$$y_j = \begin{cases} \frac{p_k}{q_k} & j = 0 \\ 0 & 0 < j \leq k \\ q_j \left(\frac{p_j}{q_j} - \frac{p_k}{q_k} \right) & j > k \end{cases}$$

See Exercise 1.3 for the motivation for this problem. (Note: The set of indices defining the integer k is never empty. To see this, note that for $j = n-1$ the condition is $q_n \leq \beta$, which may or may not be true. But, for $j = n$, the sum on the left-hand side contains no terms and so the condition is $0 \leq \beta$, which is always true. Hence, the sum always contains at least one element...the number n.)

5.16 *Diet Problem.* An MIT graduate student was trying to make ends meet on a very small stipend. He went to the library and looked up the National Research Council's publication entitled "Recommended Dietary Allowances" and was able to determine a minimum daily intake quantity of each essential nutrient for a male in his weight and age category. Let m denote the number of nutrients that he identified as important to his diet, and let b_i for $i = 1, 2, \ldots, m$ denote his personal minimum daily requirements. Next, he made a list of his favorite foods (which, except for pizza and due mostly to laziness and ineptitude in the kitchen, consisted almost entirely of frozen prepared meals). He then went to the local grocery store and made a list of the unit price for each of his favorite foods. Let us denote these prices as c_j

for $j = 1, 2, \ldots, n$. In addition to prices, he also looked at the labels and collected information about how much of the critical nutrients are contained in one serving of each food. Let us denote by a_{ij} the amount of nutrient i contained in food j. (Fortunately, he was able to call his favorite pizza delivery service and get similar information from them.) In terms of this information, he formulated the following linear programming problem:

$$\begin{array}{lll} \text{minimize} & \displaystyle\sum_{j=1}^{n} c_j x_j & \\ \text{subject to} & \displaystyle\sum_{j=1}^{n} a_{ij} x_j \geq b_i & i = 1, 2, \ldots, m \\ & x_j \geq 0 & j = 1, 2, \ldots, n. \end{array}$$

Formulate the dual to this linear program. Can you introduce another person into the above story whose problem would naturally be to solve the dual?

5.17 *Saddle points*. A function $h(y)$ defined for $y \in \mathbb{R}$ is called *strongly convex* if

- $h''(y) > 0$ for all $y \in \mathbb{R}$,
- $\lim_{y \to -\infty} h'(y) = -\infty$, and
- $\lim_{y \to \infty} h'(y) = \infty$.

A function h is called *strongly concave* if $-h$ is strongly convex. Let $\pi(x, y)$, be a function defined for $(x, y) \in \mathbb{R}^2$ and having the following form

$$\pi(x, y) = f(x) - xy + g(y),$$

where f is strongly concave and g is strongly convex. Using elementary calculus

1. Show that there is one and only one point $(x^*, y^*) \in \mathbb{R}^2$ at which the gradient of π,

$$\nabla \pi = \begin{bmatrix} \partial\pi/\partial x \\ \partial\pi/\partial y \end{bmatrix},$$

vanishes. *Hint: From the two equations obtained by setting the derivatives to zero, derive two other relations having the form $x = \phi(x)$ and $y = \psi(y)$. Then study the functions ϕ and ψ to show that there is one and only one solution.*

2. Show that

$$\max_{x \in \mathbb{R}} \min_{y \in \mathbb{R}} \pi(x, y) = \pi(x^*, y^*) = \min_{y \in \mathbb{R}} \max_{x \in \mathbb{R}} \pi(x, y),$$

where (x^*, y^*) denotes the "critical point" identified in Part 1 above. (Note: Be sure to check the signs of the second derivatives for both the inner and the outer optimizations.)

Associated with each strongly convex function h is another function, called the *Legendre transform* of h and denoted by L_h, defined by

$$L_h(x) = \max_{y \in \mathbb{R}}(xy - h(y)), \qquad x \in \mathbb{R}.$$

3. Using elementary calculus, show that L_h is strongly convex.
4. Show that

$$\max_{x \in \mathbb{R}} \min_{y \in \mathbb{R}} \pi(x, y) = \max_{x \in \mathbb{R}}(f(x) - L_g(x))$$

   and that

$$\min_{y \in \mathbb{R}} \max_{x \in \mathbb{R}} \pi(x, y) = \min_{y \in \mathbb{R}}(g(y) + L_{-f}(-y)).$$

5. Show that the Legendre transform of the Legendre transform of a function is the function itself. That is,

$$L_{L_h}(z) = h(z) \qquad \text{for all } z \in \mathbb{R}.$$

   Hint: This can be proved from scratch but it is easier to use the result of Part 2 above.

Notes

The idea behind the strong duality theorem can be traced back to conversations between G.B. Dantzig and J. von Neumann in the fall of 1947, but an explicit statement did not surface until the paper of Gale et al. (1951). The term *primal problem* was coined by G.B. Dantzig's father, T. Dantzig. The dual simplex method was first proposed by Lemke (1954).

The solution to Exercise 5.13 (which is left to the reader to supply) suggests that a random linear programming problem is infeasible with probability $1/4$, unbounded with probability $1/4$, and has an optimal solution with probability $1/2$.