LP Duality Theory

Recall from the introductory notes on Linear Programming that the dual to an LP in standard form

$$(\mathcal{P}) \qquad \begin{array}{ll} \text{maximize} & c^T x \\ \text{subject to} & Ax \le b, \ 0 \le x \end{array}$$

is the LP

$$(\mathcal{D}) \qquad \begin{array}{ll} \text{minimize} & b^T y \\ \text{subject to} & A^T y \ge c, \ 0 \le y. \end{array}$$

These two LPs are related via the Weak Duality Theorem.

Theorem 0.1 (WDT) *If x is feasible for $\mathcal{P}$ and y is feasible for $\mathcal{D}$, then*

$$c^T x \le y^T A x \le y^T b.$$

In particular, if $\mathcal{P}$ is unbounded, then $\mathcal{D}$ is necessarily infeasible, and if $\mathcal{D}$ is unbounded, then $\mathcal{P}$ is necessarily unfeasible.

The Weak Duality Theorem yields the following elementary corollary.

Corollary 0.1 *Let $\bar{x}$ be feasible for $\mathcal{P}$ and $\bar{y}$ feasible for $\mathcal{D}$ if $c^T\bar{x} = b^T\bar{y}$, then $\bar{x}$ solves $\mathcal{P}$ and $\bar{y}$ solves $\mathcal{D}$.*

Proof: Let x be any other vector feasible for $\mathcal{P}$. Then, by the WDT,

$$c^T x \le b^T \bar{y} = c^T \bar{x}.$$

Therefore,

$$\begin{array}{lll} \text{maximize} & c^T x & \le \ c^T \bar{x} \\ \text{subject to} & Ax \le b, \ 0 \le x & \end{array}$$

But $A\bar{x} \le b$, $0 \le \bar{x}$, so $\bar{x}$ solves $\mathcal{P}$. Similarly, if y is any other vector feasible for $\mathcal{D}$, then

$$b^T \bar{y} = c^T \bar{x} \le b^T y.$$

Therefore

$$\begin{array}{llll} b^T \bar{y} & \le & \text{minimize} & b^T y \\ & & \text{subject to} & A^T y \ge c, \ 0 \le y, \end{array}$$

so that $\bar{y}$ solves $\mathcal{D}$. ■

THEOREM 0.2 (THE STRONG DUALITY THEOREM) *If either $\mathcal{P}$ or $\mathcal{D}$ has a finite optimal value, then so does the other and these optimal values coincide, and, in addition, optimal solutions to both $\mathcal{P}$ and $\mathcal{D}$ exist.*

Observe that this result states that the finiteness of the optimal value implies the existence of a solution. This is not always the case for nonlinear optimization problems. Indeed, consider the problem

$$\min_{x\in\mathbb{R}} e^x.$$

This problem has a finite optimal value, namely zero; however, this value is not attained by any point $x \in \mathbb{R}$. That is, it has a finite optimal value, but a solution does not exist. The existence of solutions when the optimal value is finite is one of the many special properties of linear programs.

Before proving the Strong Duality Theorem, we first build up some machinery that will be useful in the next section. Since pivoting is just a special form of Gaussian elimination, it can be realized by matrix multiplication. We now show how this is done.

Consider the vector $v \in \mathbb{R}^m$ block decomposed as

$$v = \begin{bmatrix} a \\ \alpha \\ b \end{bmatrix}$$

where $a \in \mathbb{R}^s$, $\alpha \in \mathbb{R}$, and $b \in \mathbb{R}^t$ with $m = s + 1 + t$. Assume that $\alpha \neq 0$. We wish to determine a matrix G such that

$$Gv = e_{s+1}$$

where for $j = 1, \ldots, n$, e_j is the unit coordinate vector having a one in the jth position and zeros elsewhere. We claim that the matrix

$$G = \begin{bmatrix} I_{s\times s} & -\alpha^{-1}a & 0 \\ 0 & \alpha^{-1} & 0 \\ 0 & -\alpha^{-1}b & I_{t\times t} \end{bmatrix}$$

does the trick. Indeed,

$$Gv = \begin{bmatrix} I_{s\times s} & -\alpha^{-1}a & 0 \\ 0 & \alpha^{-1} & 0 \\ 0 & -\alpha^{-1}b & I_{t\times t} \end{bmatrix} \begin{pmatrix} a \\ \alpha \\ b \end{pmatrix} = \begin{bmatrix} a - a \\ \alpha^{-1}\alpha \\ -b + b \end{bmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = e_{t+1}.$$

The matrix G is called a *Gaussian Pivot Matrix*. Note that G is invertible since

$$G^{-1} = \begin{bmatrix} I & a & 0 \\ 0 & \alpha & 0 \\ 0 & b & I \end{bmatrix},$$

and that for any vector of the form $w = \begin{pmatrix} x \\ 0 \\ y \end{pmatrix}$ where $x \in \mathbb{R}^s$ $y \in \mathbb{R}^t$, we have

$$Gw = w.$$

The Gaussian pivot matrices perform precisely the operations required in order to execute a simplex pivot. That is, each simplex pivot can be realized as left multiplication of the simplex tableau by the appropriate Gaussian pivot matrix.

For example, consider the tableau

$$\left[\begin{array}{cccccc|c} 1 & 4 & 2 & 1 & 0 & 0 & 11 \\ 3 & \textcircled{2} & 1 & 0 & 1 & 0 & 5 \\ 4 & 2 & 2 & 0 & 0 & 1 & 8 \\ \hline 4 & 5 & 3 & 0 & 0 & 0 & 0 \end{array}\right]$$

where the $(2,2)$ element is chosen as the pivot element. The corresponding Gaussian pivot matrix is

$$G_1 = \begin{bmatrix} 1 & -2 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & \frac{-5}{2} & 0 & 1 \end{bmatrix}.$$

We have

$$\begin{bmatrix} 1 & -2 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & \frac{-5}{2} & 0 & 1 \end{bmatrix} \left[\begin{array}{cccccc|c} 1 & 4 & 2 & 1 & 0 & 0 & 11 \\ 3 & 2 & 1 & 0 & 1 & 0 & 5 \\ 4 & 2 & 2 & 0 & 0 & 1 & 8 \\ \hline 4 & 5 & 3 & 0 & 0 & 0 & 0 \end{array}\right] = \left[\begin{array}{cccccc|c} -5 & 0 & 0 & 1 & -2 & 0 & 1 \\ \frac{3}{2} & 1 & \frac{1}{2} & 0 & \frac{1}{2} & 0 & \frac{5}{2} \\ 1 & 0 & \textcircled{1} & 0 & -1 & 1 & 3 \\ \hline -\frac{7}{2} & 0 & \frac{1}{2} & 0 & \frac{-5}{2} & 0 & \frac{-25}{2} \end{array}\right].$$

Repeating this process with the new pivot element in the $(3,3)$ position yields the Gaussian pivot matrix

$$G_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \frac{-1}{2} & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{-1}{2} & 1 \end{bmatrix}$$

so that

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \frac{-1}{2} & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{-1}{2} & 1 \end{bmatrix} \left[\begin{array}{cccccc|c} -5 & 0 & 0 & 1 & -2 & 0 & 1 \\ \frac{3}{2} & 1 & \frac{1}{2} & 0 & \frac{1}{2}0 & \frac{5}{2} & \\ 1 & 0 & 1 & 0 & -1 & 1 & 3 \\ \hline \frac{-7}{2} & 0 & \frac{1}{2} & 0 & \frac{-5}{2} & 0 & \frac{-25}{2} \end{array}\right] = \left[\begin{array}{cccccc|c} -5 & 0 & 0 & 1 & -2 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & \frac{-1}{2} & 1 \\ 1 & 0 & 1 & 0 & -1 & 1 & 3 \\ \hline -4 & 0 & 0 & 0 & \frac{-3}{2} & \frac{-1}{2} & -14 \end{array}\right]$$

which is the optimal tableau. If

$$\begin{bmatrix} A & I & b \\ c^T & 0 & 0 \end{bmatrix} \tag{0.4}$$

is the initial tableau, then

$$G_2G_1 \begin{bmatrix} A & I & b \\ c^T & 0 & 0 \end{bmatrix} = \begin{bmatrix} -5 & 0 & 0 & 1 & -2 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & \frac{-1}{2} & 1 \\ 1 & 0 & 1 & 0 & -1 & 1 & 3 \\ -4 & 0 & 0 & 0 & -2 & \frac{-1}{2} & -14 \end{bmatrix}$$

That is, we would be able to go directly to the optimal tableau if we knew the matrix

$$G_2G_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \frac{-1}{2} & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{-1}{2} & 1 \end{bmatrix} \begin{bmatrix} 1 & -2 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & \frac{-5}{2} & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -2 & 0 & 0 \\ 0 & 1 & \frac{-1}{2} & 0 \\ 0 & -1 & 1 & 0 \\ 0 & -2 & \frac{-1}{2} & 1 \end{bmatrix}.$$

By partitioning this matrix in a manner that is conformal to the block structure of the initial tableau (0.4), we write

$$G_2G_1 = \begin{bmatrix} R & 0 \\ -y^T & 1 \end{bmatrix}$$

where

$$R = \begin{bmatrix} 1 & -2 & 0 \\ 0 & 1 & \frac{-1}{2} \\ 0 & -1 & 1 \end{bmatrix} \quad \text{and} \quad y = \begin{pmatrix} 0 \\ 2 \\ \frac{1}{2} \end{pmatrix}.$$

The matrix R is called the *record matrix.* This matrix is necessarily invertible since each G_i is invertible.

The process described above can just as well be applied to any initial tableau (0.4). After pivoting is completed, the optimal tableau (if there is one) necessarily has the factored representation

$$\begin{bmatrix} R & 0 \\ -y^T & 1 \end{bmatrix} \begin{bmatrix} A & I & b \\ c^T & 0 & 0 \end{bmatrix}$$

where the matrix $\begin{bmatrix} R & 0 \\ -y^T & 1 \end{bmatrix}$ is simply the product of all of the Gaussian pivot matrices that take us to the solution. This matrix will always have a one in the lower right-hand corner with a vector of zeros above it since simplex pivots never take their pivot element from the objective row of the tableau. With this notation the proof of the Strong Duality Theorem easily follows.

PROOF:[Strong Duality Theorem] Since the dual of the dual is the primal, we may as well assume that the primal has a finite optimal value. In this case, the Fundamental Theorem of Linear Programming says that an optimal basic feasible solution exists. Let

$$\begin{bmatrix} R & 0 \\ -y^T & 1 \end{bmatrix} \begin{bmatrix} A & I & b \\ c^T & 0 & 0 \end{bmatrix}$$

be the factored form representation of an optimal tableau associated with an optimal basic feasible solution. Then

$$\begin{bmatrix} R & 0 \\ -y^T & 1 \end{bmatrix} \begin{bmatrix} A & I & b \\ c^T & 0 & 0 \end{bmatrix} = \begin{bmatrix} RA & R & Rb \\ c^T - y^T A & -y^T & -y^T b \end{bmatrix}.$$

Since this is an optimal tableau we know that

$$c - A^T y \leq 0, \qquad\qquad -y^T \leq 0$$

and $y^T b$ equal the optimal value in the primal problem. But then $A^T y \geq c$ and $0 \leq y$ so that y is feasible for the dual problem $\mathcal{D}$. In addition, the Weak Duality Theorem implies that

$$\begin{array}{rcll} b^T y & = & \text{maximize} & c^T x \qquad\qquad \leq \; b^T \widehat{y} \\ & & \text{subject to} & Ax \leq b, \; 0 \leq x \end{array}$$

for every vector $\widehat{y}$ that is feasible for $\mathcal{D}$. Therefore, y solves $\mathcal{D}$. ■

This is an amazing fact! Our method for solving the primal problem $\mathcal{P}$, the simplex algorithm, simultaneously solves the dual problem $\mathcal{D}$! This fact will be of enormous practical value when we study sensitivity analysis.

By the Strong Duality Theorem, we now know that at optimality we must have equality in the Weak Duality Theorem. That is, if x solves $\mathcal{P}$ and y solves $\mathcal{D}$, then

$$c^T x = y^T A x = b^T y. \tag{0.5}$$

Let us examine the consequences of this a bit further. First note that the equation $c^T x = y^T A x$ implies that

$$0 = x^T (A^T y - c) = \sum_{j=1}^{n} x_j (\sum_{i=1}^{m} a_{ij} y_i - c_j).$$

Since feasibility implies that

$$0 \leq x_j \qquad \text{and} \qquad 0 \leq \sum_{i=1}^{m} a_{ij} y_i - c_j$$

for $j = 1, \ldots, m$, we must have that

$$x_j (\sum_{i=1}^{m} a_{ij} y_i - c_j) = 0 \qquad \text{for} \qquad j = 1, \ldots, m,$$

or equivalently,

$$x_j = 0 \qquad \text{or} \qquad \sum_{i=1}^{m} a_{ij} y_i = c_j \qquad \text{or both} \tag{0.6}$$

for $j = 1, \ldots, n$. Similarly, (0.5) implies that

$$0 = y^T(b - Ax) = \sum_{i=1}^{m} y_i(b_i - \sum_{j=1}^{n} a_{ij} x_j).$$

Again, feasibility implies that

$$0 \le y_i \qquad \text{and} \qquad 0 \le b_i - \sum_{j=1}^{n} a_{ij} x_j$$

for $i = 1, \ldots, m$. Thus, we must have

$$y_i(b_i - \sum_{j=1}^{n} a_{ij} x_j) = 0 \qquad \text{for} \qquad j = 1, \ldots, n,$$

or equivalently,

$$y_i = 0 \qquad \text{or} \qquad \sum_{j=1}^{n} a_{ij} x_j = b_i \qquad \text{or both} \tag{0.7}$$

for $i = 1, \ldots, m$. The two observations (0.6) and (0.7) combine to yield the following theorem.

Theorem 0.8 (The Complementary Slackness Theorem) *The vector $x \in \mathbb{R}^n$ solves $\mathcal{P}$ and the vector $y \in \mathbb{R}^m$ solves $\mathcal{D}$ if and only if x is feasible for $\mathcal{P}$ and y is feasible for $\mathcal{D}$ and*

(i) either $0 = x_j$ or $\sum_{i=1}^{m} a_{ij} y_i = c_j$ or both for $j = 1, \ldots, n$, and

(ii) either $0 = y_i$ or $\sum_{j=1}^{n} a_{ij} x_j = b_i$ or both for $i = 1, \ldots, m$.

Proof: If x solves $\mathcal{P}$ and y solves $\mathcal{D}$, then by the Strong Duality Theorem we have equality in the Weak Duality Theorem. But we have just observed that this implies (0.6) and (0.7) which are equivalent to (i) and (ii) above.

Conversely, if (i) and (ii) are satisfied, then we get equality in the Weak Duality Theorem. Therefore, by Corollary 0.2, x solves $\mathcal{P}$ and y solves $\mathcal{D}$. ■

The Complementary Slackness Theorem can be used to develop a test of optimality for a putative solution to $\mathcal{P}$ (or $\mathcal{D}$). We state this test as a corollary.

Corollary 0.2 *The vector $x \in \mathbb{R}^n$ solves $\mathcal{P}$ if and only if x is feasible for $\mathcal{P}$ and there exists a vector $y \in \mathbb{R}^m$ feasible for $\mathcal{D}$ and such that*

(i) if $\sum_{j=1}^{n} a_{ij}x_j < b$, then $y_i = 0$, for $i = 1, \ldots, m$ and

(ii) if $0 < x_j$, then $\sum_{i=1}^{m} a_{ij}y_i = c_j$, for $j = 1, \ldots, n$.

Proof: (i) and (ii) implies equality in the Weak Duality Theorem. The primal feasibility of x and the dual feasibility of y combined with Corollary 0.1 yield the result. ■