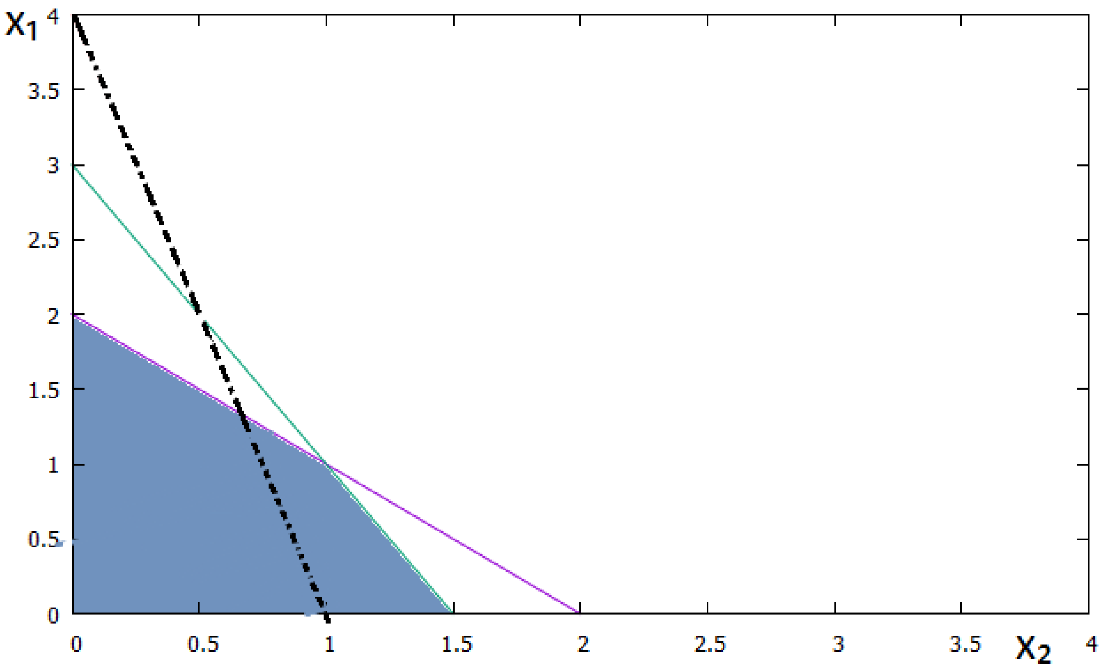
Linear Programming Problems: A Brief Overview

Griffin Dutson

Mentor: Professor Hans De Moor



Department of Mathematics, Saint Mary's College of California

May 2017

Linear Programming Problems: A Brief Overview

Linear programming problems are problems that involve optimizing one target function given a system of constraints. They are called linear programming problems not because they involve programming or computer science, but because many of these problems deal with the best way to allocate resources, which was relevant among training and logistics schedules in the government or military programs. The most famous example is probably the "diet problem," which is concerned with finding the least costly way to construct a diet that satisfies a person's nutritional requirements.

The diet problem was first considered by an economist named George Stigler in 1944. Stigler was born in Renton, Washington, on January 17, 1911 [4]. Stigler solved the general diet problem involving 77 different foods. His solution was a diet that cost only $39.93 a year (less than 11 cents/day, using 1939 prices), and consisted of wheat flour, cabbage, and dried navy beans [2]. For human nutritional requirements, Stigler considered nine common nutrients (calories, protein, calcium, iron, vitamins A, B_1, B_2, C, and niacin. Solving the diet problem requires finding a solution to the system of equations representing the nutritional constraints that also minimizes the cost function of the diet. Rather than considering 77 foods, we will look at a simple example with 5 foods in 2 equations.

Example: Let a, b, c, d, e be the number of units or servings of apples, bananas, carrots, dates, and eggs, respectively, and suppose the cost function (in cents) is given by $f(a, b, c, d, e) = 6a + 8b + 2c + 17d + 11e$. The inequalities that provide constraints for the arbitrary nutritional requirements of 70 grams of carbohydrates and 12 grams of fiber are:

$$0.4a + 1.2b + 0.6c + 0.6d + 12.2e \geq 70$$

$$0.4a + 0.6b + 0.4c + 0.2d + 2.6e \geq 12$$

These inequalities assume that apples have 0.4 grams of carbohydrates per serving, bananas have 1.2g of carbohydrates per serving, and so on. Now we have represented a simple example of the diet problem by forming inequalities based on our nutritional requirements, we must find a solution which is within the bounds of these constraints that minimizes our cost function. First, we should change the way we express the constraints to make the problem more easy. Sometimes dealing with a system of inequalities is more difficult than dealing with a system of equations. However, we can turn an inequality into an equality by introducing a *slack variable*.

Definition: A *slack variable* is a variable that is added to an inequality constraint in order to transform it into an equality constraint.

For example, the previous inequality $0.4a + 1.2b + 0.6c + 0.6d + 12.2e \geq 70$, is converted to an equality $0.4a + 1.2b + 0.6c + 0.6d + 12.2e - f = 70$, where f is a slack variable representing the amount by which one exceeds the daily value of 70 grams of carbohydrates.

Now that we have considered a simple linear programming problem, we are ready to define the general form of the problem:

Definition: The *standard form* of the linear programming problem is to determine a solution to the set of inequalities of the form:

$$a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \geq b_1$$

$$a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \geq b_2$$

$$\vdots$$

$$a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \geq b_m$$

with

$$x_j \geq 0, j = 1, 2, ..., n,$$

that minimizes the target function

$$z = c_1x_1 + c_2x_2 + \cdots + c_nx_n + z_0.$$

Now let $A = \begin{bmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$, $\mathbf{c} = \begin{bmatrix} c_1 & c_2 & \dots & c_n \end{bmatrix}$

Then we can restate the linear programming problem (Call it P) as:

$$\begin{cases} \text{minimize } \mathbf{c} \cdot \mathbf{x} \\ \text{subject to} \\ A\mathbf{x} \geq \mathbf{b} \\ x_i > 0 \ \forall \ i = 1, 2, ..., n \end{cases}$$

Where A is an $m \times n$ matrix.

A solution to the system of constraints, is called a *feasible solution*, but not all feasible solutions will optimize the target function.

Definition: A *feasible solution* is an element $\mathbf{x} \in \mathbb{R}^n$ which satisfies the constraints $A\mathbf{x} = b$, and $x_i > 0 \ \forall \ i = 1, 2, ..., n$.

Now that the problem is in matrix form, we can use Gaussian Elimination to eliminate all the variables except one in each equation. This process is called the Simplex Method. The Simplex Method, also called the Simplex Algorithm, was designed by George Bernard Dantzig.

Dantzig was born on November 8, 1914, in Portland, Oregon. [3] He graduated from the University of Maryland in 1936 and then received his M.A. in mathematics from the University of Michigan in 1938.[3] Dantzig then went on to enroll in the mathematics doctoral program at U.C. Berkeley, but left in 1941 in order to become chief of the combat analysis branch of the U.S. Air Force's statistical control headquarters. [3] During his time there, he discovered that linear programming could be applied to a variety of optimization and planning problems, since the 1940s was a time of rapid development of computing technology, Dantzig and other mathematicians were able to quickly solve complex systems of equations that would have previously taken much more time to solve, up to years per problem. Dantzig developed the simplex method for solving a linear programming problem in 1947. He then became a professor of operations and computer research at Stanford in 1966, where he contributed to many other fields, including quadratic programming, complementary pivot theory, nonlinear equations, convex programming, integer programming, stochastic programming, dynamic programming, game theory, and optimal control theory. [3].

At the end of the simplex method, the system of constraints for the linear programming problem is said to be in canonical form, and the variables that do not get set to 0 are called basic variables.

Definition: A system of m equations and n unknowns, with $m \leq n$, is in *canonical form* with a distinguished set of *basic variables* if each basic variable has coefficient 1 in one equation and 0 in the others, and each equation has exactly one basic variable with coefficient . The number of basic variables will be equal to the rank of the matrix A, and the non-basic variables serve as free variables in the system.

The simplex method can be demonstrated by the following simple example. This example illustrates how to solve a linear programming problem through the *Two Phase Simplex Method*, which is a way of implementing the Simplex Method by first finding an initial feasible solution, and then improving upon our initial solution until we find an optimal solution at the end.

Example: Suppose our problem is to maximize the function $z = 6x_1 + 5x_2$, according to the constraints

$$x_1 + x_2 \leq 5$$

$$3x_1 + 2x_2 \leq 12$$

First, we introduce slack variables s_1 and s_2 to transform these equations into inequalities, and assume that $x_1, x_2, s_1, s_2 \geq 0$. Then we have:

$$x_1 + x_2 + s_1 = 5$$

$$3x_1 + 2x_2 + s_2 = 12$$

This yields a system of three equations:

$$s_1 = 5 - x_1 - x_2$$

$$s_2 = 12 - 3x_1 - 2x_2$$

$$z = 6x_1 + 5x_2$$

In this first iteration of the algorithm, s_1, s_2 are the basic variables, and x_1, x_2 are the non-basic variables. Setting the non-basic variables to 0, the current feasible solution is then: $s_1 = 5, s_2 = 12, z = 0$. But remember that our goal is to maximize z, which we can do by either increasing x_1 or x_2. If we increase x_1 beyond $x_1 = 5$, then by the equation $s_1 = 5 - x_1 - x_2$, we can see that s_1 will become negative and violate our assumption that $s_1 \geq 0$. Similarly, the other equation $s_2 = 12 - 3x_1 - 2x_2$ implies that if we increase x_1 beyond $x_1 = 4$, we will have the same problem. Thus we take the minimum of these two constraints, and we have a limiting factor of $x_1 = 4$. Using the same method, we can find a second limiting constraint of $x_2 = 5$. Now we rewrite the equations by solving for x_1, because it becomes our pivot term. So we get:

$$x_1 = 4 - \frac{2}{3}x_2 - \frac{1}{3}x_4$$

Now we substitute this equation for x_1 in the equation $s_1 = 5 - x_1 - x_2$, to get:

$$s_1 = 1 - \frac{1}{3}x_2 + \frac{1}{3}s_2$$

Then we substitute $x_1 = 4 - \frac{2}{3}x_2 - \frac{1}{3}s_2$ back into our objective funtion $z = 6x_1 + 5x_2$, getting:

$$z = 24 + x_2 - 2s_2$$

Now we are in a situation where x_1, s_1 are basic, and x_2, s_2 are non-basic. So again setting the non-basic variables to 0, the current feasible solution is, $x_1 = 4, s_1 = 5, z = 24$. Now that the objective function has the form $z = 24 + x_2 - 2s_2$, we see that we can increase z by either increasing x_2 or decreasing s_2. However, s_2 is non-basic and was set to 0, so we cannot decrease s_2 without making it negative, which we do not allow. So our only option is to increase z by increasing x_2. Looking at the equation $x_1 = 4 - \frac{2}{3}x_2 - \frac{1}{3}s_2$, we see that the maximum value that x_2 can take while keeping x_1 positive is $x_2 = 6$. Similarly, using the equation $s_1 = 1 - \frac{1}{3}x_2 + \frac{1}{3}s_2$, we see that increasing beyond $x_2 = 3$ would cause a similar problem with s_1. Thus our limiting factor is $x_2 = 3$. So the third iteration begins by solving $s_1 = 1 - \frac{1}{3}x_2 + \frac{1}{3}s_2$ for x_2. We get:

$$x_2 = 3 - 3s_1 + s_2$$

Substituting this into $x_1 = 4 - \frac{2}{3}x_2 - \frac{1}{3}s_2$ yields:

$$x_1 = 2 + 2s_1 - s_2$$

Then by substituting $x_2 = 3 - 3s_1 + s_2$ into our objective function $z = 24 + x_2 - 2s_2$, we get:

$$z = 27 - 3s_1 - s_2$$

Now we want to increase z even more, but we cannot because we would either have to decrease s_1 or decrease s_2, or both. This is impossible since they are both already equal to 0. Now setting $s_1 = 0, s_2 = 0$ in these equations above, we obtain our final solution, which is $x_1 = 2, x_2 = 3, s_1 = 0, s_2 = 0, z = 27$, and we have successfully maximized z within 3 iterations of the simplex algorithm.

The Simplex Method can also be performed by transforming the constraints and objective function into the form of a special matrix, called a *tableau*.

Definition: A *tableau* is a matrix or array which contains the coefficients representing the constraint equations of a linear programming problem, along with its objective function in the bottom row.

After we form a tableau, we proceed by then finding pivots and performing row reduction operations on the tableau until an optimal solution for z is reached. We will follow this alternate method. First, transform the constraint inequalities into equations with the slack variables s_1 and s_2 as we did above. Then form a tableau from these equations, and perform the simplex method by using pivot operations on each row of the tableau to eventually find a solution which maximizes the objective function.

$$\left[\begin{array}{ccccc|c} x_1 & x_2 & s_1 & s_2 & z & \text{rhs} \\ \hline 1 & 1 & 1 & 0 & 0 & 5 \\ 3 & 2 & 0 & 1 & 0 & 12 \\ -6 & -5 & 0 & 0 & 1 & 0 \end{array}\right] \tag{1}$$

We find our first pivot by finding the most negative number in the bottom row, which is -6. We choose the column represented by the smallest negative number because this indicates the column where the greatest gain in our objective function can be made. We want to maximize the function $z = 6x_1 + 5x_2$, so we choose to increase x_1 instead of x_2 first. Then for each number above -6 in the associated column, we check the ratios of the number in the column labeled "rhs" divided by each number in the "x_1" column. These two ratios are $\frac{5}{1}$ and $\frac{12}{3}$. The entry which yields the smallest ratio then becomes our pivot, so our pivot is 3, in the column of x_1. This is the case because choosing the smallest ratio ensures we are increasing the variable corresponding to the lowest cost. We multiply the row containing the pivot by $\frac{1}{3}$ so that the pivot value becomes 1, then use row operations to cancel the rows above and below it, getting:

$$\left[\begin{array}{ccccc|c} x_1 & x_2 & s_1 & s_2 & z & \text{rhs} \\ \hline 0 & 1/3 & 0 & -1/3 & 0 & 1 \\ 1 & 2/3 & 0 & 1/3 & 0 & 4 \\ 0 & -1 & 0 & 3 & 1 & 24 \end{array}\right] \tag{2}$$

Then our new pivot (found through the same method as described above) becomes $\frac{1}{3}$ and after dividing the respective row by $\frac{1}{3}$ and using row operations to cancel the other rows, we get:

$$\left[\begin{array}{ccccc|c} x_1 & x_2 & s_1 & s_2 & z & \text{rhs} \\ \hline 0 & 1 & 0 & -1 & 0 & 3 \\ 1 & 0 & 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 2 & 1 & 27 \end{array}\right] \tag{3}$$

Then in the lower right corner, we see that our functions maximum value is 27, and this occurs when $x_1 = 2$ and $x_2 = 3$.

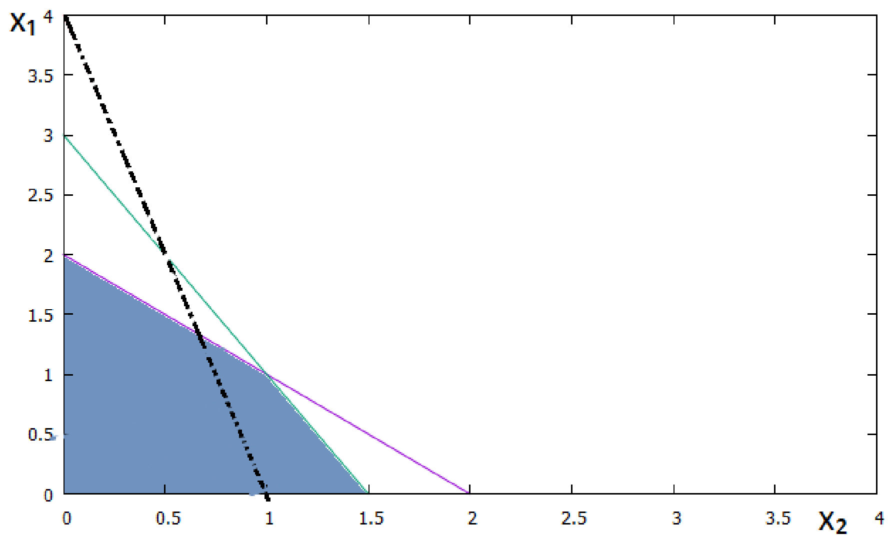
Now that we have applied the simplex method, let's look at what is happening geometrically with this method, and more importantly, whether we can determine if we have a solution or not before going through this long process. Upon graphing the constraints for any linear programming problem, you will have either an unbounded or bounded region. To properly describe this, we need to introduce another theorem.

The Corner Point Theorem:

1) If the feasible region is bounded, then the objective function has both a maximum and a minimum value and each occurs at one or more corner points.

2) If the feasible region is unbounded, the objective function may not have a maximum or minimum. But if a maximum or minimum value exists, it will occur at one or more corner points.

The Corner Point Theorem is relevant because by graphing our region before using the Simplex Method, we can determine whether or not an optimal solution will exist, and we know that if it exists, it must be at a corner point. Now let's describe what happens graphically when the simplex method is applied. First the constraints are graphed to form the feasible region. Then a *level curve* of the objective function is graphed (explained below), and upon applying the simplex method, the objective function scans back and forth through the region until it intersects with a corner point that may be the maximum or minimum we are looking for. Below is a figure of a possible feasible region for an arbitrary linear programming problem, and its objective function.



The dotted line cutting through the region is a *level curve* in the region that represents the objective function. This *level curve* is a graph of our objective function z, where we set $z = 0$ so we only deal with x_1 and x_2 as variables. We are restricted to the first quadrant because we assume that both $x_1 > 0$ and $x_2 > 0$. The pink line will shift horizontally until it intersects a corner point, at which there may or may not be a maximum or minimum. In this case, since our region is bounded, we know that both a maximum and minimum exist, and they occur at one or more corner points.

The following is another equivalent theorem which helps guarantee the existence of solutions before attempting to solve a linear programming problem.

Fundamental Theorem of Linear Programming

Given the linear programming problem P, where A is an $m \times n$ matrix of rank m:

1. If there is any feasible solution, then there is a basic feasible solution.

2. If there is any optimal solution, then there is a basic optimal solution.

The Fundamental Theorem of Linear Programming (FTLP) is simply a more mathematical, less geometric restatement of the Corner Point Theorem. We are at a corner in our feasible region when our solution only contains basic variables, thus if there is an optimal solution it must be at a corner point by the Corner Point Theorem, and then it must be a basic solution by the FTLP. Below is a proof of the Fundamental Theorem:

Proof: [6] Let $\mathbf{x}$ be a feasible solution to a linear programming problem P, where A is an $m \times n$ matrix of constraints, and we are trying to minimize our objective function of $\mathbf{c}^T\mathbf{x}$, and assume that $x_i > 0 \; \forall \; i = 1, 2, ..., n.$. Then we can write that:

$$A\mathbf{x} = \textstyle\sum_{i \in I} \mathbf{a}_i x_i = \mathbf{b}, \text{ where } I \subset \{1, ..., n\}, \text{ and } x_i > 0.$$

Deleting all indices i such that $x_i = 0$, we are left with two cases:

1) $\{\mathbf{a}_i\}_{i \in I}$ are linearly independent. If this is the case, then $\mathbf{x}$ is already a basic feasible solution, since basic solutions only consist of variables corresponding to linearly independent columns, and so both statements 1. and 2. of the fundamental theorem are proven.

2) $\{\mathbf{a}_i\}_{i \in I}$ are linearly dependent. Then $\sum_{i \in I} \mathbf{a}_i y_i = 0$ for some y_i's not all deleted. However we assume that, for at least one index $r \in I$, that $y_r > 0$. We define the vector y with the components y_i as above for $i \in I$ and components $y_i = 0$ for all the $i \notin I$. Define $x^\epsilon = x - \epsilon y$. Then, multiplying on the left by A, we have $Ax^\epsilon = Ax - \epsilon Ay$. Then, by substitution since $Ax = b$, we have $Ax - \epsilon Ay = b - \epsilon \sum_{i \in I} a_i y_i$. Since $\sum_{i \in I} a_i y_i = 0$, we are simply left with $b - \epsilon \sum_{i \in I} a_i y_i = b$. Thus $Ax^\epsilon = b$ and so by definition of a feasible solution, x^ϵ is feasible for

small enough positive or negative ϵ since, without loss of generally, if we take an arbitrary component x_p^ϵ of x^ϵ:

$$x_p^\epsilon = \begin{cases} x_p - \epsilon y_p \text{ when } p \in I \\ 0 \text{ when } p \notin I \end{cases}$$

Now for at least one index $r \in I$, $x_r^\epsilon = x_r - \epsilon y_r = 0$ for $\epsilon = x_r/y_r$. Pick the smallest such value ϵ and associated index r. Then the resulting x^ϵ is the feasible and

$$Ax^\epsilon = \sum_{i \in I^\epsilon} a_i x_i^\epsilon, \text{ where the set of indices } I^\epsilon = I \setminus \{r\}$$

Now by eliminating the indices for which $x^\epsilon = 0$, we have found another feasible solution, which has one less component, or in other words is spanned by one less column of A. We can continue eliminating columns in this manner until only linearly independent columns remain. Then, this case becomes the same as case 1), and we have a basic feasible solution, and so the first statement of the fundamental theorem of linear programming is shown for this case. To show the second part of the fundamental theorem (that we have a basic optimal solution), we must show that our solution x^ϵ is optimal if x is optimal. Since x^ϵ is a feasible solution for sufficiently small positive or negative values of ϵ, the value of our objective function: $c^T x^\epsilon = c^T x - \epsilon c^T y < c^T x$ for small ϵ as the same sign as $c^T y$ if $c^T y \neq 0$. Then since x is optimal, we must have $c^T y = 0$. By substitution, we see that $c^T x^\epsilon = c^T x$ and thus x^ϵ is basic optimal. Now both statements 1) and 2) of the Fundamental Theorem of Linear Programming were proved for each case, where the columns are linearly independent or linearly dependent, and thus we have shown all cases. QED.

Sometimes it is easier to restate a problem a different way before solving it, and then to translate the solution back when the easier version of the problem is solved. This can be done by transforming the primal linear programming problem into a dual linear programming problem:

Definition: Problems of the following forms are called *symmetric dual linear programming problems* for an $m \times n$ matrix A:

$$\text{Primal: } \begin{cases} \text{maximize } z = z_0 + c^t x \\ \text{subject to} \\ Ax \leq b \\ x \geq 0, \end{cases} \qquad \text{Dual: } \begin{cases} \text{minimize } v = z_0 + b^t y \\ \text{subject to} \\ A^t y \geq c \\ y \geq 0, \end{cases}$$

Consider the previous example of the diet problem where we want to minimize cost given nutritional requirements, and call this the *primal problem*. The *dual problem*, conversely, would be concerned with maximizing nutrition for given cost limitations. So the primal problem and the dual problem would then be written as:

$$\text{Primal:}\begin{cases} \text{minimize } f(\mathbf{x}) = 6a + 8b + 2c + 17d + 11e \\ \text{subject to} \\ 0.4a + 1.2b + 0.6c + 0.6d + 12.2e \geq 70 \\ 0.4a + 0.6b + 0.4c + 0.2d + 2.6e \geq 12 \end{cases} \qquad \text{Dual:}\begin{cases} \text{maximize } g(\mathbf{y}) = 70p + 12q \\ \text{subject to} \\ 0.4p + 0.4q \leq 6 \\ 1.2p + 0.6q \leq 8 \\ 0.6p + 0.4q \leq 2 \\ 0.6p + 0.2q \leq 17 \\ 12.2p + 2.6q \leq 11 \end{cases}$$

Where $\mathbf{x} = [a, b, c, d, e]^T$ and $\mathbf{y} = [p, q]$, and p is the number of servings of one food and q is the number of servings of another food.

The relationships between the solutions to the primal and dual problems are given by the following theorems:

Weak Duality Theorem

Suppose $\mathbf{x}_0$ is a feasible solution to the problem of maximizing $\mathbf{c} \cdot \mathbf{x}$ subject to $A\mathbf{x} \leq \mathbf{b}, \mathbf{x} \geq 0$; and $\mathbf{y}_0$ is a feasible solution to the dual problem of minimizing $\mathbf{b} \cdot \mathbf{y}$ subject to $A^t\mathbf{y} \geq \mathbf{c}$, $\mathbf{y} \geq 0$. Then

$$\mathbf{c} \cdot \mathbf{x}_0 \leq \mathbf{b} \cdot \mathbf{y}_0$$

In the context of the diet problem and its corresponding dual problem, $\mathbf{c}$ corresponds to the vector $[70, 12]$, which we can take the dot product with $[p, q]$ to form $g(\mathbf{y})$, the function we want to maximize in the dual. Similarly, $\mathbf{b}$ corresponds to the vector $[6, 8, 2, 17, 11]$ whose dot product with $[a, b, c, d, e]$ forms the function $f(\mathbf{x})$, which we want to minimize in the primal. To restate the inequality in terms of the problem, this means that the nutritional value in grams that you will get out of a given diet of foods will always be less than or equal to the cost of that diet (in cents).

Proof: [2, pg. 117] Now since $\mathbf{X_0}$ is a solution to the maximization problem, this implies that $A\mathbf{X_0} \leq \mathbf{b}$. Since $\mathbf{Y_0} \geq 0$, $\mathbf{Y_0}^t A\mathbf{X_0} \leq \mathbf{Y_0}^t\mathbf{b}$. Similarly since $\mathbf{Y_0}$ is a solution to the dual and $\mathbf{X_0} \geq 0$ implies that $A^t\mathbf{Y_0} \geq \mathbf{c}$ and $\mathbf{X_0}^t A^t\mathbf{Y_0} \geq \mathbf{X_0}^t\mathbf{c}$. Then since $\mathbf{X_0}^t A^t\mathbf{Y_0}$ is simply a real number, by the properties of transposes:

$$(\mathbf{X_0}^t A^t\mathbf{Y_0}) = (\mathbf{X_0}^t A^t\mathbf{Y_0})^t = \mathbf{Y_0}^t A\mathbf{X_0}.$$

Therefore,

$$\mathbf{c} \cdot \mathbf{X_0} = \mathbf{X_0}^t\mathbf{c} \leq \mathbf{X_0}^t A^t\mathbf{Y_0} = \mathbf{Y_0}^t A\mathbf{X_0} \leq \mathbf{Y_0}^t\mathbf{b} = \mathbf{b} \cdot \mathbf{Y_0}$$

QED.

This theorem can be extended further to state that the target functions of each problem will end up being equal, not just related in an inequality:

Strong Duality Theorem

If either the primal or dual problem has an finite optimal solution, so does the other, and $z_{max} = v_{min}$.

Proof: [5] Assume, without loss of generality, that the primal problem has a finite optimal solution. Then, by the Fundamental Theorem of Linear Programming, we know that a finite optimal basic solution exists as well. Each pivot operation required to solve a linear programming problem through the simplex tableau method can be represented by an invertible matrix that, when multiplied by our tableau on the left, performs all the necessary operations. We can represent any initial tableau with the general form:

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

Then define the matrix R to be the "record matrix" which is a matrix which contains all transformations required to obtain the final tableau of the problem. This can be done by creating matrices corresponding to each transformation, and then multiplying them together to form one final record matrix. To solve this general initial tableau represented above, we should multiply on the left by:

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

Then by performing the above mentioned matrix multiplication, we see that our final tableau will have the form:

$$\begin{bmatrix} RA & R & R\mathbf{b} \\ \mathbf{c}^T - \mathbf{y}^T A & -\mathbf{y}^T & -\mathbf{y}^T\mathbf{b} \end{bmatrix}$$

Since this is an optimal tableau, we know that $\mathbf{c} - A^T\mathbf{y} \leq 0$ and that $\mathbf{y}^T \geq 0$, so that $-\mathbf{y}^T \leq 0$, where $\mathbf{y}^T\mathbf{b}$ is the value of the objective function in the primal problem. Then adding $A^T\mathbf{y}$ to both sides of the equation $\mathbf{c} - A^T\mathbf{y} \leq 0$, we have $A^T\mathbf{y} \geq \mathbf{c}$ and $\mathbf{y} \geq 0$ so that $\mathbf{y}$ is also feasible for the dual problem. Then, by using the Weak Duality Theorem, we can see that $\mathbf{b}^T\mathbf{y} = \mathbf{c}^T\mathbf{y} \leq \mathbf{b}^T\mathbf{y}_0$ for every feasible solution $\mathbf{y}_0$ to the dual problem. Therefore $\mathbf{y}$

is a solution to the dual. QED.

When converting between the Primal and Dual problems, we are only transposing certain matrices, so we end up with a relationship between the two problems regarding the number of variables and constraints. Simply stated, after converting between problems, the number of variables in the dual is equal to the number of constraints in the primal and the number of constraints in the dual is equal to the number of variables in the primal.

Complementary Slackness Theorem [2, pg. 128]

Suppose $\mathbf{X}^* = (x_1^*, ..., x_n^*)$ is a feasible solution to the problem of

Maximizing $\mathbf{c} \cdot \mathbf{X}$ subject to $A\mathbf{X} \leq \mathbf{b}, \mathbf{X} \geq 0$ **(1)**

and $\mathbf{Y}^* = (y_1^*, ..., y_m^*)$ a feasible solution to the dual problem of

Minimizing $\mathbf{b} \cdot \mathbf{Y}$ subject to $A^t\mathbf{Y} \geq \mathbf{c}, \mathbf{Y} \geq 0$ **(2)**

Then $\mathbf{X}^*$ and $\mathbf{Y}^*$ are optimal solution points to their respective problems if and only if for each $i, 1 \leq i \leq n$, either

(slack in ith constraint of **(1)** evaluated at $\mathbf{X}^*$)$= \sum_j a_{ij} x_j^* = b_i$

or

$y_i^* = 0$ (or both)

and, for each $j, 1 \leq j \leq n$, either

(slack in jth constraint of **(2)** evaluated at $\mathbf{Y}^*$)$= \sum_i a_{ij} y_i^* = c_j$

or

$x_j^* = 0$ (or both).

This theorem attempts to depict a relationship between constraints in the primal and variables in the dual and vice versa. An interpretation of the Complementary Slackness Theorem is that if you have a primal and dual linear programming problem, then it is guaranteed that you cannot have slack in two complementary places (a constraint in the primal and a variable in the dual, or a variable in the primal and a constraint in the dual).

Proof: [1, pg. 62] Suppose that $\sum_{j=1}^{n} a_{ij}x_j^* \leq b_i$ (for $i = 1, 2, ..., m$), and that $x_j^* \geq 0$ (for $j = 1, 2, ..., n$) **(#1)**. Suppose that $\sum_{i=1}^{m} a_{ij}y_i^* \geq c_j$ (for $j = 1, 2, ..., n$), and that $y_i^* \geq 0$ (for $i = 1, 2, ..., m$) **(#2)**. Multiplying the inequality $c_j \leq \sum_{i=1}^{m} a_{ij}y_i^*$ by x_j^* on the right on both sides (we can do this since $x_j^* \geq 0$) gives us $c_jx_j^* \leq (\sum_{i=1}^{m} a_{ij}y_i^*)x_j^*$ **(# 3)**. Similarly, we can take the inequality $\sum_{j=1}^{n} a_{ij}x_j^* \leq b_i$ and multiply by y_i^* on the right on both sides to get $(\sum_{j=1}^{n} a_{ij}x_j^*)y_i^* \leq b_iy_i^*$ **(#4)**. Now we can take the summation of each inequality: so we have $\sum_{j=1}^{n} c_jx_j^* \leq \sum_{j=1}^{n}(\sum_{i=1}^{m} a_{ij}y_i^*)x_j^*$ and $\sum_{i=1}^{m}(\sum_{j=1}^{n} a_{ij}x_j^*)y_i^* \leq \sum_{i=1}^{m} b_iy_i^*$. Combining these, we know that $\sum_{j=1}^{n} c_jx_j^* \leq \sum_{j=1}^{n}(\sum_{i=1}^{m} a_{ij}y_i^*)x_j^* = \sum_{i=1}^{m}(\sum_{j=1}^{n} a_{ij}x_j^*)y_i^* \leq \sum_{i=1}^{m} b_iy_i^*$ **(#5)**. So **(#5)** holds with equalities throughout if and only if equalities hold in **(# 3)** and **(#4)**. To guarantee that we have the equality $c_jx_j^* = (\sum_{i=1}^{m} a_{ij}y_i^*)x_j^*$ is to insist that $x_j^* = 0$; failing that, we must require $c_j = \sum_{i=1}^{m} a_{ij}y_i^*$. Thus equalities hold in **(# 3)** if and only if **(#1)** is satisfied. Similarly, equalities hold in **(#4)** if and only if **(#2)** is satisfied. Thus **(#1)** and **(#2)** are necessary in order for the equation $\sum_{j=1}^{n} c_jx_j^* = \sum_{i=1} mb_iy_i^*$**(#6)** to hold. Furthermore, the Duality Theorem shows that equation **(#6)** is necessary for $\mathbf{X}^* = (x_1^*, ..., x_n^*)$ and $\mathbf{Y}^* = (y_1^*, ..., y_m^*)$ to be simultaneously optimal solutions. QED.

The Complementary Slackness theorem can be restated in a more useful way which helps us to rewrite and simplify either the primal or dual given information on the solution of either of them. This restatement of the theorem is as follows:

Complementary Slackness Theorem

- Let $\mathbf{x} = \{x_1, x_2, ...x_n\}$ be a feasible solution to the primal problem.
- Let $\mathbf{u} = \{u_1, u_2, ..., u_m\}$ be a feasible solution to the dual problem.
- Then $\mathbf{x}$ is primal optimal and $\mathbf{u}$ is dual optimal if and only if:

$$s_i u_i = 0, \text{ where } i = 1, 2, ..., m \text{ and}$$

$$e_j x_j = 0, \text{ where } j = 1, 2, ..., n.$$

where s_i is a slack variable in the primal, and e_j is a slack variable in the dual.

This restatement of the theorem makes its uses more apparent, since now we see that if we have the solution to the dual problem and we want to solve the primal, all of our nonzero values of slack variables in the dual imply that the variables in the primal problem corresponding to those slacks must equal 0. Upon substituting that these variables as 0 in the primal, the primal problem will lose one or more variables and be easier to solve. Thus knowing the solution to the dual problem, especially the values of certain slack variables, is very valuable information that we can utilize to simplify the corresponding primal problem.

The four main theorems of Linear Programming, which include the Fundamental Theorem, the Weak and Strong Duality Theorems, and the Complementary Slackness Theorem, have been outlined. These four theorems have other applications, such as determining the sensitivity of the optimal solution to small changes in certain variables. This can help with determining the most efficient way to make a gain or loss in the objective function if certain constraints are "relaxed." That is, if we can relax one constraint, which one should we relax to create the biggest gain or loss in our objective function? There are many other applications of these theorems as well, but hopefully this paper has succeeded in giving a brief overview of what linear programming problems are and provided insight into some methods and theorems that are relevant to solving them.

Works Cited

[1] Chvatal, Vasek, Linear Programming, 1983

[2] Thie, Paul R., An Introduction to Linear Programming and Game Theory 2nd Edition, 1988.

[3] "George Bernard Dantzig." World of Mathematics, Gale, 2006. Biography in Context

[4] "Stigler, George Joseph." International Encyclopedia of the Social Sciences, edited by William A. Darity, Jr., 2nd ed., vol. 8, Macmillan Reference USA, 2008, pp. 144-145. Gale Virtual Reference Library

[5] "Duality Theory" https://sites.math.washington.edu/ burke/crs/407/notes/section4.pdf, Retrieved May 2, 2017.

[6] "Basic Theorem of Linear Programming" http://www.math.udel.edu/ angell/basicth.pdf, Retrieved May 2, 2017.