Linear Programming Review

1 Introduction

1.1 What is Linear Programming?

A mathematical optimization problem is one in which some function is either maximized or minimized relative to a given set of alternatives. The function to be minimized or maximized is called the *objective function* and the set of alternatives is called the feasible region (or constraint region). In this course, the feasible region is always taken to be a subset of $\mathbb{R}^n$ (real n-dimensional space) and the objective function is a function from $\mathbb{R}^n$ to $\mathbb{R}$.

We further restrict the class of optimization problems that we consider to linear programming problems (or LPs). An LP is an optimization problem over $\mathbb{R}^n$ wherein the objective function is a linear function, that is, the objective has the form

$$c_1x_1 + c_2x_2 + \cdots + c_nx_n$$

for some $c_i \in \mathbb{R}$ $i = 1, \ldots, n$, and the feasible region is the set of solutions to a finite number of linear inequality and equality constraints, of the form

$$a_{i1}x_i + a_{i2}x_2 + \cdots + a_{in}x_n \leq b_i \qquad i = 1, \ldots, s$$

and

$$a_{i1}x_i + a_{i2}x_2 + \cdots + a_{in}x_n = b_i \qquad i = s+1, \ldots, m.$$

Linear programming is an extremely powerful tool for addressing a wide range of applied optimization problems. A short list of application areas is resource allocation, production scheduling, warehousing, layout, transportation scheduling, facility location, flight crew scheduling, parameter estimation, ...

1.2 An Example

To illustrate some of the basic features of LP, we begin with a simple two-dimensional example. In modeling this example, we will review the four basic steps in the development of an LP model:

1. Determine and label the *decision variables.*
2. Determine the objective and use the decision variables to write an expression for the *objective function.*
3. Determine the *explicit constraints* and write a functional expression for each of them.
4. Determine the *implicit constraints.*

PLASTIC CUP FACTORY

> A local family-owned plastic cup manufacturer wants to optimize their production mix in order to maximize their profit. They produce personalized beer mugs and champaign glasses. The profit on a case of beer mugs is \$25 while the profit on a case of champaign glasses is \$20. The cups are manufactured with a machine called a plastic extruder which feeds on plastic resins. Each case of beer mugs requires 20 lbs. of plastic resins to produce while champaign glasses require 12 lbs. per case. The daily supply of plastic resins is limited to at most 1800 pounds. About 15 cases of either product can be produced per hour. At the moment the family wants to limit their work day to 8 hours.

We will model the problem of maximizing the profit for this company as an LP. The first step in our modeling process is to determine the *decision variables*. These are the variables that represent the quantifiable decisions that must be made in order to determine the daily production schedule. That is, we need to specify those quantities whose values completely determine a production schedule and its associated profit. In order to determine these quantities, one can ask the question "If I were the plant manager for this factory, what must I know in order to implement a production schedule?" The best way to determine the decision variables is to put oneself in the shoes of the decision maker and then ask the question "What do I need to know in order to make this thing work?" In the case of the plastic cup factory, everything is determined once it is known how many cases of beer mugs and champaign glasses are to be produced each day.

Decision Variables:

$B = \#$ of cases of beer mugs to be produced daily.

$C = \#$ of cases of champaign glasses to be produced daily.

You will soon discover that the most difficult part of any modeling problem is the determination of decision variables. Once these variables are correctly determined then the remainder of the modeling process usually goes smoothly.

After specifying the decision variables, one can now specify the problem objective. That is, one can write an expression for the objective function.

Objective Function:

Maximize profit where profit $= 25B + 20C$

The next step in the modeling process is to express the feasible region as the solution set of a finite collection of linear inequality and equality constraints. We separate this process into two steps:

1. determine the explicit constraints, and
2. determine the implicit constraints.

The explicit constraints are those that are explicitly given in the problem statement. In the problem under consideration, there are explicit constraints on the amount of resin and the number of work hours that are available on a daily basis.

Explicit Constraints:

resin constraint: $20B + 12C \leq 1800$

work hours constraint: $\frac{1}{15}B + \frac{1}{15}C \leq 8$.

This problem also has other constraints called implicit constraints. These are constraints that are not explicitly given in the problem statement but are present nonetheless. Typically these constraints are associated with "natural" or "common sense" restrictions on the decision variable. In the cup factory problem it is clear that one cannot have negative cases of beer mugs and champaign glasses. That is, both B and C must be non-negative quantities.

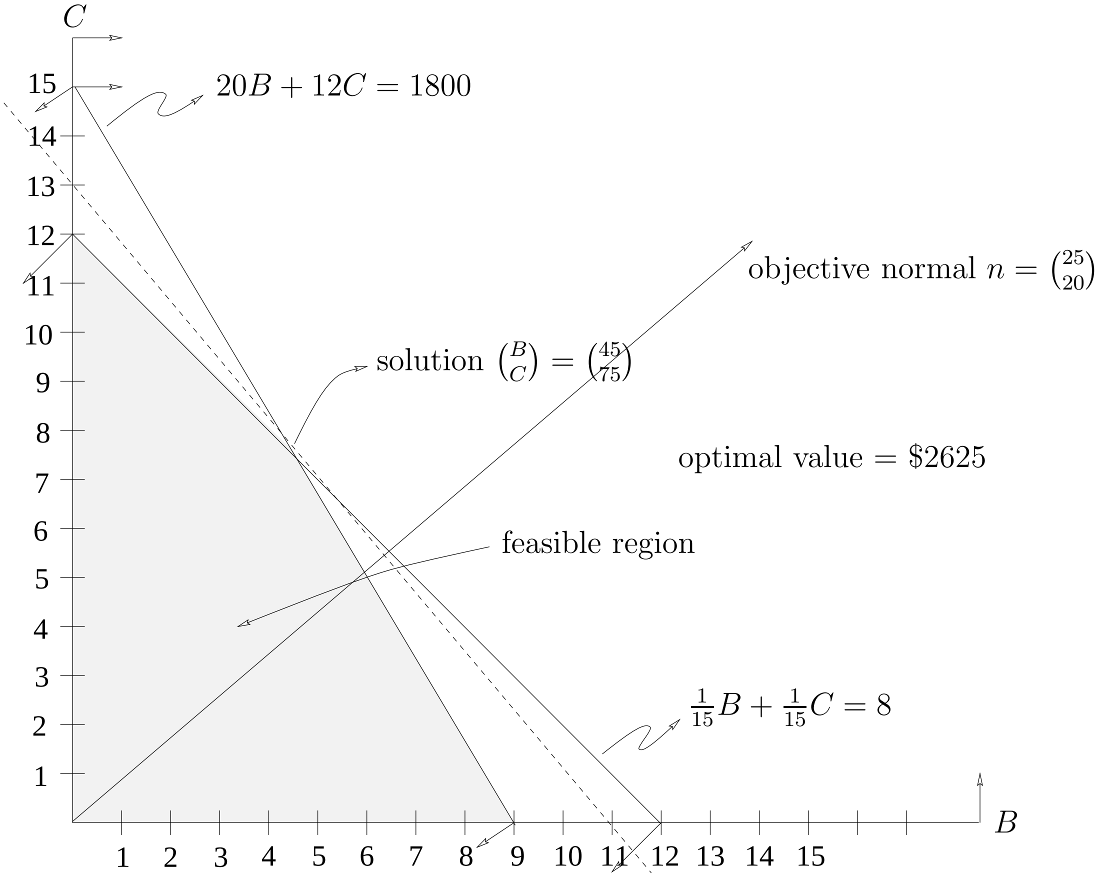
Implicit Constraints:

$$0 \leq B, \qquad 0 \leq C.$$

The entire model for the cup factory problem can now be succinctly stated as

$$\begin{array}{rl} \mathcal{P} : & \max 25B + 20C \\ & \text{subject to} \;\; \begin{array}{rcl} 20B + 12C & \leq & 1800 \\ \frac{1}{15}B + \frac{1}{15}C & \leq & 8 \\ 0 & \leq & B, C \end{array} \end{array}$$

Since this problem is two dimensional it is possible to provide a graphical solution. The first step toward a graphical solution is to graph the feasible region. To do this, first graph



the line associated with each of the linear inequality constraints. Then determine on which side of each of these lines the feasible region must lie (don't forget the implicit constraints!). Once the correct side is determined it is helpful to put little arrows on the line to remind yourself of the correct side. Then shade in the resulting feasible region.

The next step is to draw in the vector representing the gradient of the objective function at the origin. Since the objective function has the form

$$f(x_1, x_2) = c_1x_1 + c_2x_2,$$

the gradient of f is the same at every point in $\mathbb{R}^2$;

$$\nabla f(x_1, x_2) = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}.$$

Recall from calculus that the gradient always points in the direction of increasing function values. Moreover, since the gradient is constant on the whole space, the level sets of f associated with different function values are given by the lines perpendicular to the gradient. Consequently, to obtain the location of the point at which the objective is maximized we simply set a ruler perpendicular to the gradient and then move the ruler in the direction of the gradient until we reach the last point (or points) at which the line determined by the ruler intersects the feasible region. In the case of the cup factory problem this gives the solution to the LP as $\binom{B}{C} = \binom{45}{75}$

We now recap the steps followed in the solution procedure given above:

Step 1: Graph each of the linear constraints indication on which side of the constraint the feasible region must lie. Don't forget the implicit constraints!

Step 2: Shade in the feasible region.

Step 3: Draw the gradient vector of the objective function.

Step 4: Place a straightedge perpendicular to the gradient vector and move the straightedge either in the direction of the gradient vector for maximization, or in the opposite direction of the gradient vector for minimization to the last point for which the straightedge intersects the feasible region. The set of points of intersection between the straightedge and the feasible region is the set of solutions to the LP.

The solution procedure described above for two dimensional problems reveals a great deal about the geometric structure of LPs that remains true in n dimensions. We will explore this geometric structure more fully as the course evolves. But for the moment, we continue to study this 2 dimensional LP to see what else can be revealed about the structure of this problem.

Before leaving this section, we make a final comment on the modeling process described above. We emphasize that there is not one and only one way to model the Cup Factory problem, or any problem for that matter. In particular, there are many ways to choose the decision variables for this problem. Clearly, it is sufficient for the shop manager to know how many hours each days should be devoted to the manufacture of beer mugs and how many hours to champaign glasses. From this information everything else can be determined. For example, the number of cases of beer mugs that get produced is 15 times the number of hours devoted to the production of beer mugs. Therefore, as you can see there are many ways to model a given problem. But in the end, they should all point to the same optimal process.

1.3 Duality Theory

We now briefly discuss how the "hidden hand of the market place" gives rise to a theory of dual linear programs. Think of the cup factory production process as a black box through which the resources flow. Raw resources go in one end and exit the other. When they come out the resources have a different form, but whatever comes out is still comprised of the entering resources. However, something has happened to the value of the resources by passing through the black box. The resources have been purchased for one price as they enter the box and are sold in their new form as they leave. The difference between the entering and exiting prices is called the profit. Assuming that there is a positive profit the resources have increased in value as they pass through the production process. The marginal value of a resource is precisely the increase in the per unit value of the resource due to the production process.

Let us now consider how the market introduces pressures on the profitability and the value of the resources available to the market place. We take the perspective of the cup factory *vs* the market place. The market place does not want the cup factory to go out of business. On the other hand, it does not want the cup factory to see a profit. It wants to keep all the profit for itself and only let the cup factory just break even. It does this by setting the price of the resources available in the market place. That is, the market sets the price for plastic resin and labor and it tries to do so in such a way that the cup factory sees no profit and just breaks even. Since the cup factory is now seeing a profit, the market must figure out by how much the sale price of resin and labor must be raised to reduce this profit to zero. This is done by minimizing the value of the available resources over all price increments that guarantee that the cup factory either loses money or sees no profit from both of its products. If we denote the per unit price increment for resin by R and that for labor by L, then the profit for beer mugs is eliminated as long as

$$20R + \frac{1}{15}L \geq 25$$

since the left hand side represents the increased value of the resources consumed in the production of one case of beer mugs and the right hand side is the current profit on a case of beer mugs. Similarly, for champaign glasses, the market wants to choose R and L so that

$$12R + \frac{1}{15}L \geq 20.$$

Now in order to maintain equilibrium in the market place, that is, not drive the cup factory out of business (since then the market realizes no profit at all), the market chooses R and L so as to minimize the increased value of the available resources. That is, the market chooses R and L to solve the problem

$$\begin{array}{rl} \mathcal{D} : & \text{minimize}\, 1800R + 8L \\ & \begin{array}{rrcl} \text{subject to} & 20R + \frac{1}{15}L & \geq & 25 \\ & 12R + \frac{1}{15}L & \geq & 20 \\ & 0 & \leq & R, L \end{array} \end{array}$$

This is just another LP. It is called the "dual" to the LP $\mathcal{P}$ in which the cup factory tries to maximize profit. Observe that if $\binom{B}{C}$ is feasible for $\mathcal{P}$ and $\binom{R}{L}$ is feasible for $\mathcal{D}$, then

$$\begin{array}{rcl} 25B + 20C & \leq & [20R + \frac{1}{15}L]B + [12R + \frac{1}{15}L]C \\ & = & R[20B + 12C] + L[\frac{1}{15}B + \frac{1}{15}C] \\ & \leq & 1800R + 8L. \end{array}$$

Thus, the value of the objective in $\mathcal{P}$ at a feasible point in $\mathcal{P}$ is bounded above by the objective in $\mathcal{D}$ at any feasible point for $\mathcal{D}$. In particular, the optimal value in $\mathcal{P}$ is bounded above by the optimal value in $\mathcal{D}$. The "strong duality theorem" states that if either of these problems has a finite optimal value, then so does the other and these values coincide. In addition, we claim that the solution to $\mathcal{D}$ is given by the marginal values for $\mathcal{P}$. That is, $\binom{R}{L} = \begin{bmatrix} 5/8 \\ 375/2 \end{bmatrix}$ is the optimal solution for $\mathcal{D}$. In order to show this we need only show that $\binom{R}{L} = \begin{bmatrix} 5/8 \\ 375/2 \end{bmatrix}$ is feasible for $\mathcal{D}$ and that the value of the objective in $\mathcal{D}$ at $\binom{R}{L} = \begin{bmatrix} 5/8 \\ 375/2 \end{bmatrix}$ coincides with the value of the objective in $\mathcal{P}$ at $\binom{B}{C} = \binom{45}{75}$. First we check feasibility:

$$0 \le \frac{5}{8}, \qquad 0 \le \frac{375}{2}$$

$$\begin{aligned} 20 \cdot \frac{5}{8} + \frac{1}{15} \cdot \frac{375}{2} &\ge 25 \\ 12 \cdot \frac{5}{8} + \frac{1}{15} \cdot \frac{375}{2} &\ge 20. \end{aligned}$$

Next we check optimality

$$25 \cdot 45 + 20 \cdot 75 = 2625 = 1800 \cdot \frac{5}{8} + 8 \cdot \frac{375}{2}.$$

1.4 LPs in Standard Form and Their Duals

Recall that a linear program is a problem of maximization or minimization of a linear function subject to a finite number of linear inequality and equality constraints. This general definition leads to an enormous variety of possible formulations. In this section we propose one fixed formulation for the purposes of developing an algorithmic solution procedure. We then show that every LP can be recast in this form. We say that an LP is in *standard form* if it has the form

$$\begin{array}{lll} \mathcal{P}: & \text{maximize} & c_1x_1 + c_2x_2 + \cdots + c_nx_n \\ & \text{subject to} & a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \;\;\le\;\; b_i \text{ for } i = 1, 2, \ldots, m \\ & & 0 \le x_j \text{ for } j = 1, 2, \ldots, n\ . \end{array}$$

Using matrix notation, we can rewrite this LP as

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

where the inequalities $Ax \le b$ and $0 \le x$ are to be interpreted componentwise.

Following the results of the previous section on LP duality, we claim that the dual LP to $\mathcal{P}$ is the LP

$$\begin{array}{lll} \mathcal{D}: & \text{minimize} & b_1y_1 + b_2y_2 + \cdots + b_ny_m \\ & \text{subject to} & a_{1j}y_1 + a_{2j}y_2 + \cdots + a_{mj}y_m \;\;\ge\;\; c_j \text{ for } j = 1, 2, \ldots, n \\ & & 0 \le y_i \text{ for } i = 1, 2, \ldots, m\ . \end{array}$$

Again, the statement of this $\mathcal{D}$ can be simplified by the use of matrix notation to give the problem

$$\begin{array}{lll} \mathcal{D}: & \text{minimize} & b^T y \\ & \text{subject to} & A^T y \geq c \\ & & 0 \leq y \ . \end{array}$$

Just as for the cup factory problem, the LPs $\mathcal{P}$ and $\mathcal{D}$ are related via the *Weak Duality Theorem.*

THEOREM: [WEAK DUALITY] *If $x \in \mathbb{R}^n$ is feasible for $\mathcal{P}$ and $y \in \mathbb{R}^m$ is feasible for $\mathcal{D}$, then*

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

Thus, if $\mathcal{P}$ is unbounded, then $\mathcal{D}$ is infeasible, and if $\mathcal{D}$ is unbounded, then $\mathcal{P}$ is infeasible.

PROOF: Let $x \in \mathbb{R}^n$ be feasible for $\mathcal{P}$ and $y \in \mathbb{R}^m$ be feasible for $\mathcal{D}$. Then

$$\begin{array}{rcl} c^T x & = & \sum\limits_{j=1}^{n} c_j x_j \\ & \leq & \sum\limits_{j=1}^{n} (\sum\limits_{i=1}^{m} a_{ij} y_i) x_j \qquad [\text{since } x_j \geq 0 \text{ and } \sum\limits_{i=1}^{m} a_{ij} y_i \geq c_j] \\ & = & y^T A x \\ & = & \sum\limits_{i=1}^{m} (\sum\limits_{j=1}^{n} a_{ij} x_j) y_i \\ & \leq & \sum\limits_{i=1}^{m} b_i y_i \qquad [\text{since } y_i \geq 0 \text{ and } \sum\limits_{j=1}^{n} a_{ij} x_j \leq b_i] \\ & = & b^T y \end{array}$$

■

We caution that the infeasibility of either $\mathcal{P}$ or $\mathcal{D}$ does not imply the unboundedness of the other. Indeed, it is possible for both $\mathcal{P}$ and $\mathcal{D}$ to be infeasible as is illustrated by the following example.

EXAMPLE:

$$\begin{array}{rrcrcr} \text{maximize} & 2x_1 & - & x_2 & & \\ & x_1 & - & x_2 & \leq & 1 \\ & -x_1 & + & x_2 & \leq & -2 \\ & 0 & \leq & x_1, & x_2 & \end{array}$$

The Weak Duality Theorem yields the following elementary corollary.

COROLLARY 1.1 *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,

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

Therefore,

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

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

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

Therefore

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

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

THEOREM 1.1 (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.

2 Solving LPs: The Simplex Algorithm of George Dantzig

2.1 Simplex Pivoting: Dictionary Format

We illustrate a general solution procedure, called the *simplex algorithm*, by implementing it on a very simple example. Consider the LP

$$
\begin{array}{rrcl}
(2.1) \qquad & \max 5x_1 + 4x_2 + 3x_3 & & \\
& \text{s.t. } 2x_1 + 3x_2 + x_3 & \leq & 5 \\
& 4x_1 + x_2 + 2x_3 & \leq & 11 \\
& 3x_1 + 4x_2 + 2x_3 & \leq & 8 \\
& 0 \leq x_1, x_2, x_3 & &
\end{array}
$$

In devising our solution procedure we take a standard mathematical approach; reduce the problem to one that we already know how to solve. Since the structure of this problem is essentially linear, we will try to reduce it to a problem of solving a system of linear equations, or perhaps a series of such systems. By encoding the problem as a system of linear equations we bring into play our knowledge and experience with such systems in the new context of linear programming.

In order to encode the LP (2.1) as a system of linear equations we must first transform linear inequalities into linear equation. This is done by introducing a new non-negative variable, called a *slack variable*, for each inequality:

$$
\begin{array}{rcrcll}
x_4 & = & 5 & - & [2x_1 + 3x_2 + x_3] & \geq 0, \\
x_5 & = & 11 & - & [4x_1 + x_2 + 2x_3] & \geq 0, \\
x_6 & = & 8 & - & [3x_1 + 4x_2 + 2x_3] & \geq 0.
\end{array}
$$

To handle the objective, we introduce a new variable z:

$$z = 5x_1 + 4x_2 + 3x_3.$$

Then all of the information associated with the LP (2.1) can be coded as follows:

$$
\begin{array}{rrcccccl}
& 2x_1 + 3x_2 + x_3 & + & x_4 & & & = & 5 \\
& 4x_1 + x_2 + 2x_3 & & + & x_5 & & = & 11 \\
(2.2) \qquad & 3x_2 + 4x_2 + 2x_3 & & & + & x_6 & = & 8 \\
& -z + 5x_1 + 4x_2 + 3x_3 & & & & & = & 0 \\
& 0 \leq x_1, x_2, x_3, x_4, x_5, x_6. & & & & & &
\end{array}
$$

The new variables x_4, x_5, and x_6 are called slack variables since they take up the *slack* in the linear inequalities. This system can also be written using block structured matrix notation as

$$\begin{bmatrix} 0 & A & I \\ -1 & c^T & 0 \end{bmatrix} \begin{bmatrix} z \\ x \end{bmatrix} = \begin{bmatrix} b \\ 0 \end{bmatrix},$$

where

$$A = \begin{bmatrix} 2 & 3 & 1 \\ 4 & 1 & 2 \\ 3 & 4 & 2 \end{bmatrix}, \; I = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \; b = \begin{bmatrix} 5 \\ 11 \\ 8 \end{bmatrix}, \text{ and } c = \begin{bmatrix} 5 \\ 4 \\ 3 \end{bmatrix}.$$

The augmented matrix associated with the system (2.2) is

$$\left[\begin{array}{ccc|c} 0 & A & I & b \\ -1 & c & 0 & 0 \end{array}\right] \tag{2.3}$$

and is referred to as the *initial simplex tableau* for the LP (2.1).

Again consider the system

$$\begin{array}{rcl} x_4 & = & 5 - 2x_1 - 3x_2 - x_3 \\ x_5 & = & 11 - 4x_1 - x_2 - 2x_3 \\ x_6 & = & 8 - 3x_1 - 4x_2 - 2x_3 \\ z & = & 5x_1 + 4x_2 + 3x_3. \end{array} \tag{2.4}$$

This system defines the variables x_4, x_5, x_6 and z as linear combinations of the variables x_1, x_2, and x_3. We call this system a *dictionary* for the LP (2.1). More specifically, it is the *initial* dictionary for the the LP (2.1). This initial dictionary defines the objective value z and the slack variables as a linear combination of the initial decision variables. The variables that are "defined" in this way are called the *basic variables*, while the remaining variables are called *nonbasic*. The set of all basic variables is called the *basis*. A particular solution to this system is easily obtained by setting the non-basic variables equal to zero. In this case, we get

$$\begin{array}{rcl} x_4 & = & 5 \\ x_5 & = & 11 \\ x_6 & = & 8 \\ z & = & 0. \end{array}$$

Note that the solution

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 5 \\ 11 \\ 8 \end{pmatrix} \tag{2.5}$$

is feasible for the extended system (2.2) since all components are non-negative. For this reason, we call the dictionary (2.4) a *feasible dictionary* for the LP (2.1), and we say that this LP has *feasible origin*.

In general, a dictionary for the LP (2.1) is any system of 4 linear equations that defines three of the variables $x_1, \ldots, x_6$ and z in terms of the remaining 3 variables and has the same solution set as the initial dictionary. The variables other than z that are being defined in the dictionary are called the basis for the dictionary, and the remaining variables are said to be non-basic in the dictionary. Every dictionary identifies a particular solution to the linear system obtained by setting the non-basic variables equal to zero. Such a solution is said to be a *basic feasible solution* (BFS) for the LP (2.1) if it componentwise non-negative, that is, all of the numbers in the vector are non-negative so that the point lies in the feasible region for the LP.

The grand strategy of the simplex algorithm is to move from one feasible dictionary representation of the system (2.2) to another (and hence from one BFS to another) while simultaneously increasing the value of the objective variable z. In the current setting, beginning with the dictionary (2.4), what strategy might one employ in order to determine a new dictionary whose associated BFS gives a greater value for the objective variable z?

Each feasible dictionary is associated with one and only one feasible point. This is the associated BFS obtained by setting all of the non-basic variables equal to zero. This is how we obtain (2.5). To change the feasible point identified in this way, we need to increase the value of one of the non-basic variables from its current value of zero. Note that we cannot decrease the value of a non-basic variable since we wish to remain feasible, that is, we wish to keep all variables non-negative.

Note that the coefficient of each of the non-basic variables in the representation of the objective value z in (2.4) is positive. Hence, if we pick any one of these variables and increase its value from zero while leaving remaining two at zero, we automatically increase the value of the objective variable z. Since the coefficient on x_1 in the representation of z is the greatest, we can increase z the fastest by increasing x_1.

By how much can we increase x_1 and still remain feasible? For example, if we increase x_1 to 3 then (2.4) says that $x_4 = -1$, $x_5 = -1$, $x_6 = -1$ which is not feasible. Let us consider this question by examining the equations in (2.4) one by one. Note that the first equation in the dictionary (2.4),

$$x_4 = 5 - 2x_1 - 3x_2 - x_3,$$

shows that x_4 remains non-negative as long as we do not increase the value of x_1 beyond 5/2 (remember, x_2 and x_3 remain at the value zero). Similarly, using the second equation in the dictionary (2.4),

$$x_5 = 11 - 4x_1 - x_2 - 2x_3,$$

x_5 remains non-negative if $x_1 \leq 11/4$. Finally, the third equation in (2.4),

$$x_6 = 8 - 3x_1 - 4x_2 - 2x_3,$$

implies that x_6 remains non-negative if $x_1 \leq 8/3$. Therefore, we remain feasible, i.e. keep **all** variables non-negative, if our increase to the variable x_1 remains less than

$$\frac{5}{2} = \min\left\{\frac{5}{2}, \frac{11}{4}, \frac{8}{3}\right\}.$$

If we now increase the value of x_1 to $\frac{5}{2}$, then the value of x_4 is driven to zero. One way to think of this is that x_1 *enters the basis while* x_4 *leaves the basis.* Mechanically, we obtain the new dictionary having x_1 basic and x_4 non-basic by using the defining equation for x_4 in the current dictionary:

$$x_4 = 5 - 2x_1 - 3x_2 - x_3.$$

By moving x_1 to the left hand side of this equation and x_4 to the right, we get the new equation

$$2x_1 = 5 - x_4 - 3x_2 - x_3$$

or equivalently

$$x_1 = \frac{5}{2} - \frac{1}{2}x_4 - \frac{3}{2}x_2 - \frac{1}{2}x_3.$$

The variable x_1 can now be *eliminated* from the remaining two equations in the dictionary by substituting in this equation for x_1 where it appears in these equations:

$$\begin{aligned}
x_1 &= \frac{5}{2} - \frac{1}{2}x_4 - \frac{3}{2}x_2 - \frac{1}{2}x_3 \\
x_5 &= 11 - 4\left(\frac{5}{2} - \frac{1}{2}x_4 - \frac{3}{2}x_2 - \frac{1}{2}x_3\right) - x_2 - 2x_3 \\
&= 1 + 2x_4 + 5x_2 \\
x_6 &= 8 - 3\left(\frac{5}{2} - \frac{1}{2}x_4 - \frac{3}{2}x_2 - \frac{1}{2}x_3\right) - 4x_2 - 2x_3 \\
&= \frac{1}{2} + \frac{3}{2}x_4 + \frac{1}{2}x_2 - \frac{1}{2}x_3 \\
z &= 5\left(\frac{5}{2} - \frac{1}{2}x_4 - \frac{3}{2}x_2 - \frac{1}{2}x_3\right) + 4x_2 + 3x_3 \\
&= \frac{25}{2} - \frac{5}{2}x_4 - \frac{7}{2}x_2 + \frac{1}{2}x_3.
\end{aligned}$$

When this substitution is complete, we have the new dictionary and the new BFS:

$$\begin{aligned}
x_1 &= \frac{5}{2} - \frac{1}{2}x_4 - \frac{3}{2}x_2 - \frac{1}{2}x_3 \\
x_5 &= 1 + 2x_4 + 5x_2 \\
x_6 &= \frac{1}{2} + \frac{3}{2}x_4 + \frac{1}{2}x_2 - \frac{1}{2}x_3 \\
z &= \frac{25}{2} - \frac{5}{2}x_4 - \frac{7}{2}x_2 + \frac{1}{2}x_3,
\end{aligned} \tag{2.6}$$

and the associated BFS is

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = \begin{pmatrix} 5/2 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1/2 \end{pmatrix} \qquad \text{with} \qquad z = \frac{25}{2}. \tag{2.7}$$

This process might seem very familiar to you. It is simply Gaussian elimination. As we know from our knowledge of linear systems of equations, Gaussian elimination can be performed in a matrix context with the aid of the augmented matrix (or, simplex tableau) (2.3). We return to this observation later to obtain a more efficient computational technique.

We now have a new dictionary (2.6) which identifies the basic feasible solution (BFS) (2.7) with associated objective value $z = \frac{25}{2}$. Can we improve on this BFS and obtain a higher objective value? Let's try the same trick again, and repeat the process we followed in going from the initial dictionary (2.4) to the new dictionary (2.6). Note that the coefficient of x_3 in the representation of z in the new dictionary (2.6) is positive. Hence if we increase the value of x_3 from zero, we will increase the value of z. By how much can we increase the value of x_3 and yet keep all the remaining variables non-negative? As before, we see that the first equation in the dictionary (2.6) combined with the need to keep x_1 non-negative implies that we cannot increase x_3 by more than $(5/2)/(1/2) = 5$. However, the second equation in (2.6) places no restriction on increasing x_3 since x_3 does not appear in this equation. Finally, the third equation in (2.6) combined with the need to keep x_6 non-negative implies that we cannot increase x_3 by more than $(1/2)/(1/2) = 1$. Therefore, in order to preserve the non-negativity of all variables, we can increase x_3 by at most

$$1 = \min\{5, 1\}.$$

When we do this x_6 is driven to zero, so x_3 enters the basis and x_6 leaves. More precisely, first move x_3 to the left hand side of the defining equation for x_6 in (2.6),

$$\frac{1}{2}x_3 = \frac{1}{2} + \frac{3}{2}x_4 + \frac{1}{2}x_2 - x_6,$$

or, equivalently,

$$x_3 = 1 + 3x_4 + x_2 - 2x_6,$$

then substitute this expression for x_3 into the remaining equations,

$$\begin{aligned} x_1 &= \frac{5}{2} - \frac{1}{2}x_4 - \frac{3}{2}x_2 - \frac{1}{2}[1 + 3x_4 + x_2 - 2x_6] \\ &= 2 - 2x_4 - 2x_2 + x_6 \\ x_5 &= 1 + 2x_4 + 5x_2 \\ z &= \frac{25}{2} - \frac{5}{2}x_4 - \frac{7}{2}x_2 + \frac{1}{2}[1 + 3x_4 + x_2 - 2x_6] \\ &= 13 - x_4 - 3x_2 - x_6\ , \end{aligned}$$

yielding the dictionary

$$\begin{aligned} x_3 &= 1 + 3x_4 + x_2 - 2x_6 \\ x_1 &= 2 - 2x_4 + 2x_2 + x_6 \\ x_5 &= 1 + 2x_4 + 2x_2 \\ z &= 13 - x_4 - 3x_2 - x_6 \end{aligned}$$

which identifies the feasible solution

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

having objective value $z = 13$.

Can we do better? NO! This solution is optimal! The coefficient on the variables in the cost row of the dictionary,

$$z = 13 - x_4 - 3x_2 - x_6,$$

are all non-positive, so increasing any one of their values will not increase the value of the objective. (Why does this prove optimality?) The process of moving from one feasible dictionary to the next is called a *simplex pivot*. The overall process of stringing a sequence of simplex pivots together in order to locate an optimal solution is called the *Simplex Algorithm*. The simplex algorithm is consistently ranked as one of the ten most important algorithmic discoveries of the 20th century (http://www.uta.edu/faculty/rcli/TopTen/topten.pdf). The algorithm is generally attributed to George Dantzig (1914-2005)who is known as the father of linear programming. In 1984 Narendra Karmarkar published a paper describing a new approach to solving linear programs that was both numerically efficient and had *polynomial complexity*. This new class of methods are called *interior point* methods. These new methods have revolutionized the optimization field over the last 30 years, and they have led to efficient numerical methods for a wide variety of optimization problems well beyond the confines of linear programming. However, the simplex algorithm continues as an important numerical method for solving LPs, and for many specially structured LPs it is still the most efficient algorithm.

2.2 Simplex Pivoting: Tableau Format (Augmented Matrix Format)

We now review the implementation of the simplex algorithm by applying Gaussian elimination to the augmented matrix (2.3), also known as the simplex tableau. For this problem, the initial simplex tableau is given by

$$\left[\begin{array}{ccc|c} 0 & A & I & b \\ -1 & c & 0 & 0 \end{array}\right] = \left[\begin{array}{ccccccc|c} 0 & 2 & 3 & 1 & 1 & 0 & 0 & 5 \\ 0 & 4 & 1 & 2 & 0 & 1 & 0 & 11 \\ 0 & 3 & 4 & 2 & 0 & 0 & 1 & 8 \\ -1 & 5 & 4 & 3 & 0 & 0 & 0 & 0 \end{array}\right]. \tag{2.8}$$

Each simplex pivot on a dictionary corresponds to one step of Gaussian elimination on the augmented matrix associated with the dictionary. For example, in the first simplex pivot, x_1

enters the basis and x_4 leaves the basis. That is, we use the first equation of the dictionary to rewrite x_1 as a function of the remaining variables, and then use this representation to eliminate x_1 form the remaining equations. In terms of the augmented matrix (2.8), this corresponds to first making the coefficient for x_1 in the first equation the number 1 by dividing this first equation through by 2. Then use this entry to eliminate the column under x_1, that is, make all other entries in this column zero (Gaussian elimination):

$$\begin{array}{ccccccc|c|cl}
 & \begin{array}{c}\text{Pivot}\\ \text{column}\\ \downarrow\end{array} & & & & & & & \begin{array}{c}\text{ratios}\\ \downarrow\end{array} & \\
0 & \enclose{circle}{2} & 3 & 1 & 1 & 0 & 0 & 5 & \enclose{circle}{5/2} & \leftarrow\ \text{Pivot row}\\
0 & 4 & 1 & 2 & 0 & 1 & 0 & 11 & 11/4 & \\
0 & 3 & 4 & 2 & 0 & 0 & 1 & 8 & 8/3 & \\
\hline
-1 & \enclose{circle}{5} & 4 & 3 & 0 & 0 & 0 & 0 & & \\
\hline
 & & & & & & & & & \\
\hline
0 & 1 & 3/2 & 1/2 & 1/2 & 0 & 0 & 5/2 & & \\
0 & 0 & -5 & 0 & -2 & 1 & 0 & 1 & & \\
0 & 0 & -1/2 & 1/2 & -3/2 & 0 & 1 & 1/2 & & \\
\hline
 & & & & & & & & & \\
-1 & 0 & -7/2 & 1/2 & -5/2 & 0 & 0 & -25/2 & & \\
\hline
\end{array}$$

In this illustration, we have placed a line above the cost row to delineate its special roll in the pivoting process. In addition, we have also added a column on the right hand side which contains the ratios that we computed in order to determine the pivot row. Recall that we must use the smallest ratio in order to keep all variables in the associated BFS non-negative. Note that we performed the exact same arithmetic operations but in the more efficient matrix format. The new augmented matrix,

$$\left[\begin{array}{ccccccc|c}
0 & 1 & 3/2 & 1/2 & 1/2 & 0 & 0 & 5/2\\
0 & 0 & -5 & 0 & -2 & 1 & 0 & 1\\
0 & 0 & -1/2 & 1/2 & -3/2 & 0 & 1 & 1/2\\
-1 & 0 & -7/2 & 1/2 & -5/2 & 0 & 0 & -25/2
\end{array}\right], \tag{2.9}$$

is the augmented matrix for the dictionary (2.6).

The initial augmented matrix (2.8) has basis x_4, x_5, and x_6. The columns associated with these variables in the initial tableau (2.8) are distinct columns of the identity matrix. Correspondingly, the basis for the second tableau is x_1, x_5, and x_6, and again this implies that the columns for these variables in the tableau (2.9) are the corresponding distinct columns of the identity matrix. In tableau format, this will always be true of the basic variables, i.e., their associated columns are distinct columns of the identity matrix. To recover the BFS (basic feasible solution) associated with this tableau we first set the non-basic variables equal to zero (i.e. the variables not associated with columns of the identity matrix (except in very unusual circumstances)): $x_2 = 0$, $x_3 = 0$, and $x_4 = 0$. To find the value of the basic variables go to the column associated with that variable (for example, x_1 is in the second column), in that column find the row with the number 1 in it, then in that row go to the number to the right of the vertical bar (for x_1 this is the first row with the number to the right of the bar being 5/2). Then set this basic variable equal to that number ($x_1 = 5/2$). Repeating this for x_5 and x_6 we get $x_5 = 1$ and $x_6 = 1/2$. To get the corresponding value for z, look at the z row and observe that the corresponding linear equation is

$$-z - \frac{7}{2}x_2 + \frac{1}{2}x_3 - \frac{5}{2}x_4 = -\frac{25}{2},$$

but $x_2,\ x_3$, and x_4 are non-basic and so take the value zero giving $-z = -25/2$, or $z = 25/2$.

Of course this is all exactly the same information we obtained from the dictionary approach. The simplex, or augmented matrix approach is simply a more efficient computational procedure. We will use the simplex procedure in class to solve LPs. In addition, on quizzes and exams you will be required to understand how to go back and forth between these two representations, i.e the dictionary representation and its corresponding simplex tableau (or, augmented matrix). Let us now continue with the second simplex pivot.

In every tableau we always reserve the bottom row for encoding the linear relationship between the objective variable z and the currently non-basic variables. For this reason we call this row the *cost row*, and to distiguish its special role, we place a line above it in the tableau (this is reminiscent of the way we place a vertical bar in an augmented matrix to distinguish the right hand side of a linear equation). In the cost row of the tableau (2.9),

$$[-1,\ 0,\ -7/2,\ 1/2,\ -5/2,\ 0,\ 0,\ |\ -25/2],$$

we see a positive coefficient, 1/2, in the 4th column. Hence the cost row coefficient for the non-basic variable x_3 in this tableau is 1/2. This indicates that if we increase the the value of x_3, we also increase the value of the objective z. This is not true for any of the other currently non-basic variables since their cost row coefficients are all non-positive. Thus, the only way we can increase the value of z is to bring x_3 into the basis, or equivalently, pivot on the x_3 column which is the 4th column of the tableau. For this reason, we call the x_3 column the *pivot column*. Now if x_3 is to enter the basis, then which variable leaves? Just as with the dictionary representation, the variable that leaves the basis is that currently basic variable whose non-negativity places the greatest restriction on increasing the value of x_3. This restriction is computed as the smallest ratio of the right hand sides and the positive coefficients in the x_3 column:

$$1 = \min\{(5/2)/(1/2),\ (1/2)/(1/2)\}.$$

The ratios are only computed with the positive coefficients since a non-positive coefficient means that by increasing this variable we do not decrease the valued of the corresponding basic variable and so it is not a restricting equation. Since the minimum ratio in this instance is 1 and it comes from the third row, we find that the *pivot row* is the third row. Looking across the third row, we see that this row identifies x_6 as a basic variable since the x_6 column is a column of the identity with a 1 in the third row. Hence x_6 is the variable leaving the basis when x_3 enters. The intersection of the pivot column and the pivot row is called the *pivot*. In this instance it is the number 1/2 which is the $(3,4)$ entry of the simplex tableau. Pivoting on this entry requires us to first make it 1 by multiplying this row through by 2, and then to apply Gaussian elimination to force all other entries in this column to zero:

$$\begin{array}{rrrrrrr|rccl}
 & & & \text{Pivot} & & & & & & & \\
 & & & \text{column} & & & & & \text{r} & & \\
 & & & \downarrow & & & & & & & \\
0 & 1 & 3/2 & 1/2 & 1/2 & 0 & 0 & 5/2 & 5 & & \\
0 & 0 & -5 & 0 & -2 & 1 & 0 & 1 & & & \\
0 & 0 & -1/2 & \textcircled{1/2} & -3/2 & 0 & 1 & 1/2 & \textcircled{1} & \leftarrow & \text{pivot row} \\
\cline{1-8}
-1 & 0 & -7/2 & \textcircled{1/2} & -5/2 & 0 & 0 & -25/2 & & & \\
\cline{1-8}
\multicolumn{8}{c}{} \\
\cline{1-8}
0 & 1 & 2 & 0 & 2 & 0 & -1 & 2 & & & \\
0 & 0 & -5 & 0 & -2 & 1 & 0 & 1 & & & \\
0 & 0 & -1 & 1 & -3 & 0 & 2 & 1 & & & \\
\cline{1-8}
-1 & 0 & -3 & 0 & -1 & 0 & -1 & -13 & & &
\end{array}$$

This simplex tableau is said to be optimal since it is feasible (the associated BFS is non-negative) and the cost row coefficients for the variables are all non-positive. A BFS that is optimal is called and *optimal basic feasible solution*. The optimal BFS is obtained by setting the non-basic variables equal to zero and setting the basic variables equal to the value on the right hand side corresponding to the one in its column: $x_1 = 2,\ x_2 = 0,\ x_3 = 1,\ x_4 = 0,\ x_5 = 1,\ x_6 = 0$. The optimal objective value is obtained by taking the negative of the number in the lower right hand corner of the optimal tableau: $z = 13$.

We now recap the complete sequence of pivots in order to make a final observation that will help streamline the pivoting process: pivots are circled,

$$\begin{array}{rrrrrrr|r}
0 & \textcircled{2} & 3 & 1 & 1 & 0 & 0 & 5 \\
0 & 4 & 1 & 2 & 0 & 1 & 0 & 11 \\
0 & 3 & 4 & 2 & 0 & 0 & 1 & 8 \\
\hline
-1 & 5 & 4 & 3 & 0 & 0 & 0 & 0 \\
\hline
\multicolumn{8}{c}{} \\
\hline
0 & 1 & 3/2 & 1/2 & 1/2 & 0 & 0 & 5/2 \\
0 & 0 & -5 & 0 & -2 & 1 & 0 & 1 \\
0 & 0 & -1/2 & \textcircled{1/2} & -3/2 & 0 & 1 & 1/2 \\
\hline
-1 & 0 & -7/2 & 1/2 & -5/2 & 0 & 0 & -25/2 \\
\hline
\multicolumn{8}{c}{} \\
\hline
0 & 1 & 2 & 0 & 2 & 0 & -1 & 2 \\
0 & 0 & -5 & 0 & -2 & 1 & 0 & 1 \\
0 & 0 & -1 & 1 & -3 & 0 & 2 & 1 \\
\hline
-1 & 0 & -3 & 0 & -1 & 0 & -1 & -13
\end{array}$$

Observe from this sequence of pivots that the z column is never touched, that is, it remains the same in all tableaus. Essentially, it just serves as a place holder reminding us that in the linear equation for the cost row the coefficient of z is -1. Therefore, for the sake of expediency we will drop this column from our simplex computations in most settings, and simply re-insert it whenever instructive or convenient. However, *it is very important to always remember that it is there*! Indeed, we will make explicit and essential use of this column in order to arrive at a full understanding of the duality theory for linear programming. After removing this column, the above pivots take the following form:

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

$$
\begin{array}{cccccc|c}
\hline
1 & 3/2 & 1/2 & 1/2 & 0 & 0 & 5/2 \\
0 & -5 & 0 & -2 & 1 & 0 & 1 \\
0 & -1/2 & \textcircled{1/2} & -3/2 & 0 & 1 & 1/2 \\
\hline
0 & -7/2 & 1/2 & -5/2 & 0 & 0 & -25/2 \\
\hline
\end{array}
$$

$$
\begin{array}{cccccc|c}
\hline
1 & 2 & 0 & 2 & 0 & -1 & 2 \\
0 & -5 & 0 & -2 & 1 & 0 & 1 \\
0 & -1 & 1 & -3 & 0 & 2 & 1 \\
\hline
0 & -3 & 0 & -1 & 0 & -1 & -13
\end{array}
$$

We close this section with a final example of simplex pivoting on a tableau giving only the essential details.

The LP

$$
\begin{array}{lrcrcrcl}
\text{maximize} & 3x & + & 2y & - & 4z & & \\
 & & & & & & & \\
\text{subject to} & x & + & 4y & & & \leq & 5 \\
 & 2x & + & 4y & - & 2z & \leq & 6 \\
 & x & + & y & - & 2z & \leq & 2 \\
 & 0 & \leq & x, & y, & z & &
\end{array}
$$

Simplex Iterations

							ratios
1	4	0	1	0	0	5	5
2	4	-2	0	1	0	6	3
①	1	-2	0	0	1	2	2
3	2	-4	0	0	0	0	

0	3	2	1	0	-1	3	3/2
0	2	②	0	1	-2	2	1
1	1	−2	0	0	1	2	
0	−1	2	0	0	−3	−6	

0	1	0	1	−1	1	1
0	1	1	0	1/2	−1	1
1	3	0	0	1	−1	4
0	−3	0	0	−1	−1	−8

Optimal Solution

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 4 \\ 0 \\ 1 \end{bmatrix} \qquad \text{optimal value} \ = \ 8$$

A final word of advise, when doing simplex pivoting by hand, it is helpful to keep the tableaus vertically aligned in order to keep track of the arithmetic operations. Lined paper helps to keep the rows straight. But the columns need to be straight as well. Many students find that it is easy to keep both the rows and columns straight if they do pivoting on graph paper having large boxes for the numbers.

2.3 Dictionaries: The General Case for LPs in Standard Form

Recall the following standard form for LPs:

$$\begin{array}{rll} \mathcal{P}: & \text{maximize} & c^T x \\ & \text{subject to} & Ax \leq b \\ & & 0 \leq x \ , \end{array}$$

where $c \in \mathbb{R}^n$, $b \in \mathbb{R}^m$, $A \in \mathbb{R}^{m\times n}$ and the inequalities $Ax \leq b$ and $0 \leq x$ are to be interpreted componentwise. We now provide a formal definition for a dictionary associated with an LP in standard form. Let

$$(D_I) \qquad \begin{array}{rl} x_{n+i} = & b_i - \displaystyle\sum_{j=1}^{n} a_{ij} x_j \\ z = & \displaystyle\sum_{j=1}^{n} c_j x_j \end{array}$$

be the defining system for the slack variables x_{n+i}, $i = 1, \cdots, n$ and the objective variable z. A dictionary for $\mathcal{P}$ is any system of the form

$$(D_B) \qquad \begin{aligned} x_i &= \widehat{b}_i - \sum_{j \in N} \widehat{a}_{ij} x_j \qquad & i \in B \\ z &= \widehat{z} + \sum_{j \in N} \widehat{c}_j x_j \end{aligned}$$

where B and N are index sets contained in the set of integers $\{1, \ldots, n+m\}$ satisfying

(1) B contains m elements,

(2) $B \cap N = \emptyset$

(3) $B \cup N = \{1, 2, \ldots, n+m\}$,

and such that the systems (D_I) and (D_B) have identical solution sets. The set $\{x_j : j \in B\}$ is said to be the basis associated with the dictionary (D_B) (sometimes we will refer to the index set B as the basis for the sake of simplicity), and the variables x_i, $i \in N$ are said to be the non-basic variables associated with this dictionary. The point identified by this dictionary is

$$(2.10) \qquad \begin{aligned} x_i &= \widehat{b}_i \qquad & i \in B \\ x_j &= 0 \qquad & j \in N. \end{aligned}$$

The dictionary is said to be feasible if $0 \leq \widehat{b}_i$ for $i \in N$. If the dictionary D_B is feasible, then the point identified by the dictionary (2.10) is said to be a basic feasible solution (BFS) for the LP. A feasible dictionary and its associated BFS are said to be optimal if $\widehat{c}_j \leq 0$ $j \in N$.

Simplex Pivoting by Matrix Multiplication

As we have seen simplex pivoting can either be performed on dictionaries or on the augmented matrices that encode the linear equations of a dictionary in matrix form. In matrix form, simplex pivoting reduces to our old friend, Gaussian elimination. In this section, we show that Gaussian elimination can be represented as a consequence of left multiplication by a specially designed matrix called a *Gaussian pivot matrix*.

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,

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

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 following initial feasible 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. In this case,

$$s = 1, \quad t = 2, \quad a = 4, \quad \alpha = 2, \quad \text{and} \quad b = \begin{bmatrix} 2 \\ 5 \end{bmatrix},$$

and so the corresponding Gaussian pivot matrix is

$$G_1 = \begin{bmatrix} I_{1\times 1} & -\alpha^{-1}a & 0 \\ 0 & \alpha^{-1} & 0 \\ 0 & -\alpha^{-1}b & I_{2\times 2} \end{bmatrix} = \begin{bmatrix} 1 & -2 & 0 & 0 \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & \frac{-5}{2} & 0 & 1 \end{bmatrix}.$$

Multiplying the simplex on the left by G_1 gives

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

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

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

and left multiplication by G_2 gives

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

yielding the optimal tableau.

If

$$\begin{bmatrix} A & I & b \\ c^T & 0 & 0 \end{bmatrix} \tag{2.4}$$

is the initial tableau, then

$$G_2G_1 \begin{bmatrix} A & I & b \\ c^T & 0 & 0 \end{bmatrix} = \left[\begin{array}{cccccc|c} -5 & 0 & 0 & 1 & -2 & 0 & 1 \\ 1 & 1 & 0 & 0 & 1 & \frac{-1}{2} & 1 \\ 1 & 0 & 1 & 0 & -1 & 1 & 3 \\ \hline -4 & 0 & 0 & 0 & -2 & \frac{-1}{2} & -14 \end{array}\right]$$

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

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

beforehand. Moreover, the matrix G is invertible since both G_1 and G_2 are invertible:

$$G^{-1} = G_1^{-1}G_2^{-1} = \begin{bmatrix} 1 & 4 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 5 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & \frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{1}{2} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 4 & 2 & 0 \\ 0 & 2 & 1 & 0 \\ 0 & 2 & 3 & 0 \\ 0 & 5 & 3 & 1 \end{bmatrix}$$

(you should check that $GG^{-1} = I$ by doing the multiplication by hand). In general, every sequence of simplex pivots has a representation as left multiplication by some invertible matrix since pivoting corresponds to left multiplication of the tableau by a Gaussian pivot matrix, and Gaussian pivot matrices are always invertible. We now examine the consequence of this observation more closely in the general case. In this discussion, it is essential that we include the column associated with the objective variable z which we have largely ignored up to this point.

Recall the initial simplex tableau, or augmented matrix associated with the system (D_I):

$$T_0 = \begin{bmatrix} 0 & A & I & b \\ -1 & c^T & 0 & 0 \end{bmatrix} .$$

In this discussion, it is essential that we include the first column, i.e. the column associated with the objective variable z in the augmented matrix. Let the matrix

$$T_k = \begin{bmatrix} 0 & \widehat{A} & R & \widehat{b} \\ -1 & \widehat{c}^T & -y^T & \widehat{z} \end{bmatrix}$$

be another simplex tableau obtained from the initial tableau after a series of k simplex pivots. Note that the first column remains unchanged. Indeed, the fact that simplex pivots do not alter the first column is the reason why we drop it in our hand computations. But in the discussion that follows its presence and the fact that it remains unchanged by simplex pivoting is very important. Since T_k is another simplex tableau the $m \times (n + m)$ matrix $[\widehat{A} \quad R]$ must posses among its columns the m columns of the $m \times m$ identity matrix. These columns of the identity matrix correspond precisely to the basic variables associated with this tableau.

Our prior discussion on Gaussian pivot matrices tells us that T_k can be obtained from T_0 by multiplying T_0 on the left by some nonsingular $(m+1) \times (m+1)$ matrix G where G is the product of a sequence of Gaussian pivot matrices. In order to better understand the action of G on T_0 we need to decompose G into a block structure that is conformal with that of T_0:

$$G = \begin{bmatrix} M & u \\ v^T & \beta \end{bmatrix} ,$$

where $M \in \mathbb{R}^{m\times m}$, $u, v \in \mathbb{R}^m$, and $\beta \in \mathbb{R}$. Then

$$\begin{aligned}
\begin{bmatrix} 0 & \widehat{A} & R & \widehat{b} \\ -1 & \widehat{c}^T & -y^T & \widehat{z} \end{bmatrix} &= T_k \\
&= GT_0 \\
&= \begin{bmatrix} M & u \\ v^T & \beta \end{bmatrix} \begin{bmatrix} 0 & A & I & b \\ -1 & c^T & 0 & 0 \end{bmatrix} \\
&= \begin{bmatrix} -u & MA + uc^T & M & Mb \\ -\beta & v^TA + \beta c^T & v^T & v^Tb \end{bmatrix} .
\end{aligned}$$

By equating the blocks in the matrices on the far left and far right hand sides of this equation, we find from the first column that

$$u = 0 \qquad \text{and} \qquad \beta = 1 .$$

Here we see the key role played by our knowledge of the structure of the objective variable column (the first column). From the $(1, 3)$ and the $(2, 3)$ terms on the far left and right hand sides of (2.5), we also find that

$$M = R, \qquad \text{and} \qquad v = -y .$$

Putting all of this together gives the following representation of the k^{th} tableau T_k:

$$T_k = \begin{bmatrix} R & 0 \\ -y^T & 1 \end{bmatrix} \begin{bmatrix} 0 & A & I & b \\ -1 & c^T & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & RA & R & Rb \\ -1 & c^T - y^TA & -y^T & -y^Tb \end{bmatrix} , \tag{2.5}$$

where the matrix R is necessarily invertible since the matrix

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

is invertible (prove this!):

$$G^{-1} = \begin{bmatrix} R^{-1} & 0 \\ y^TR^{-1} & 1 \end{bmatrix} . \qquad \text{(check by multiplying out } GG^{-1}\text{)}$$

The matrix R is called the *record* matrix for the tableau as it keeps track of all of the transformations required to obtain the new tableau. Again, the variables associated with the columns of the identity correspond to the basic variables. The tableau T_k is said to be *primal feasible*, or just feasible, if $\widehat{b} = Rb \geq 0$.

3 Initialization and the Two Phase Simplex Algorithm

We now turn to the problem of finding an initial basic feasible solution. Again consider an LP in standard form,

$$\mathcal{P} \qquad \begin{array}{ll} \text{maximize} & c^T x \\ \text{subject tp} & Ax \leq b, \quad 0 \leq x. \end{array}$$

We associate with this LP an auxiliary LP of the form

$$\mathcal{Q} \qquad \begin{array}{ll} \text{minimize} & x_0 \\ \text{subject to} & Ax - x_0 \mathsf{e} \leq b, \quad 0 \leq x_0, x\ . \end{array}$$

where $\mathsf{e} \in \mathsf{R}^m$ is the vector of all ones. The ith row of the system of inequalities $Ax - x_0\mathsf{e} \leq b$ takes the form

$$a_{i1}x_1 \,+\, a_{i2}x_2 \,+\, \ldots \,+\, a_{in}x_n \,-\, x_0 \,\leq\, b_i\ .$$

The system of inequalities can also be written in block matrix form as

$$\left[\begin{array}{cc} \mathsf{e} & A \end{array}\right] \begin{pmatrix} x_0 \\ x \end{pmatrix} \leq b\ .$$

Note that if the optimal value in the auxiliary problem is zero, then at the optimal solution $(\tilde{x}_0, \tilde{x})$ we have $\tilde{x}_0 = 0$. If we plug this into the inequality $Ax - x_0\mathsf{e} \leq b$, we get $A\tilde{x} \leq b$. That is, $\tilde{x}$ is feasible for the original LP $\mathcal{P}$. Corresponding, if $\hat{x}$ is feasible for $\mathcal{P}$, then $(\hat{x}_0, \hat{x})$ with $\hat{x}_0 = 0$ is feasible for $\mathcal{A}$, in which case $(\hat{x}_0, \hat{x})$ must be optimal for $\mathcal{A}$. Therefore, $\mathcal{P}$ is feasible if and only if the optimal value in $\mathcal{A}$ is zero. In particular the feasibility of $\mathcal{P}$ can be determined by solving the LP $\mathcal{A}$.

The auxiliary problem $\mathcal{A}$ is also called the *Phase I* problem since solving it is the first phase of a two phase process of solving general LPs. In Phase I we solve the auxiliary problem to obtain an initial feasible tableau for the original problem, and in Phase II we solve the original LP starting with the feasible tableau provided in Phase I.

Solving $\mathcal{Q}$ by the simplex algorithm yields an initial feasible dictionary for $\mathcal{P}$. However, to solve $\mathcal{Q}$ we need an initial feasible dictionary for $\mathcal{Q}$. But if $\mathcal{P}$ does not have feasible origin neither does $\mathcal{Q}$! Fortunately, an initial feasible dictionary for $\mathcal{Q}$ is easily constructed. Observe that if we set $x_0 = -\min\{b_i : i = 0, \ldots, n\}$ with $b_0 = 0$, then $b + x_0\mathsf{e} \geq 0$ since

$$\begin{array}{rl} & \min\{b_i + x_0 : i = 1, \ldots, m\} \\ = & \min\{b_i : i = 1, \ldots, m\} - \min\{b_i : i = 0, \ldots, m\} \geq 0. \end{array}$$

Hence, by setting $x_0 = -\min\{b_i : i = 0, \ldots, m\}$ and $x = 0$ we obtain a feasible solution for $\mathcal{Q}$. It is also a basic feasible solution. To see this, consider the initial dictionary for $\mathcal{Q}$

$$\begin{array}{rcl} x_{n+i} & = & b_i + x_0 - \displaystyle\sum_{j=1}^{m} a_{ij}x_j \\ z & = & -x_0. \end{array}$$

Let i_0 be any index such that

$$b_{i_0} = \min\{b_i : i = 0, 1, \ldots, m).$$

If $i_0 = 0$, then this LP has feasible origin and so the initial dictionary is optimal. If $i_0 > 0$, then pivot on this row bringing x_0 into the basis. This yields the dictionary

$$\begin{aligned}
x_0 &= -b_{i_0} + x_{n+i_0} + \sum_{j=1}^{m} a_{i_0 j} x_j \\
x_{n+i} &= b_i - b_{i_0} + x_{n+i_0} - \sum_{j=1}^{m} (a_{ij} - a_{i_0 j}) x_j, \quad i \neq i_0 \\
z &= b_{i_0} - x_{n+i_0} - \sum_{j=1}^{m} a_{i_0 j} x_j.
\end{aligned}$$

But $b_{i_0} \leq b_i$ for all $i = 1, \ldots, m$, so

$$0 \leq b_i - b_{i_0} \text{ for all } i = 1, \ldots, m.$$

Therefore this dictionary is feasible. We illustrate this initialization procedure by example.

Consider the LP

$$\begin{array}{lrcrcrcr}
\max & x_1 & - & x_2 & + & x_3 & & \\
\text{s.t.} & 2x_1 & - & x_2 & + & 2x_3 & \leq & 4 \\
 & 2x_1 & - & 3x_2 & + & x_3 & \leq & -5 \\
 & -x_1 & + & x_2 & - & 2x_3 & \leq & -1 \\
 & & & \multicolumn{3}{c}{0 =\leq x_1, x_2, x_3\ .} & &
\end{array}$$

This LP does not have feasible origin since the right hand side vector $b = (4, -5, -1)^T$ is not componentwise non-negative. Hence the initial dictionary D_I is not feasible. Therefore, we must first solve the auxiliary problem $\mathcal{Q}$ to obtain a feasible dictionary. The auxiliary problem has the form

$$\begin{array}{lrcrcrcrcr}
\max & -x_0 & & & & & & & & \\
\text{s.t.} & -x_0 & + & 2x_1 & - & x_2 & + & 2x_3 & \leq & 4 \\
 & -x_0 & + & 2x_1 & - & 3x_2 & + & x_3 & \leq & -5 \\
 & -x_0 & - & x_1 & + & x_2 & - & 2x_3 & \leq & -1 \\
 & & & \multicolumn{5}{c}{0 \leq x_0, x_1, x_2, x_3\ .} & &
\end{array}$$

The initial tableau for this LP has the following form:

$$\begin{array}{rrrrrrr|r}
-1 & 2 & -1 & 2 & 1 & 0 & 0 & 4 \\
-1 & 2 & -3 & 1 & 0 & 1 & 0 & -5 \\
-1 & -1 & 1 & -2 & 0 & 0 & 1 & -1 \\
\hline
-1 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{array}$$

However, instead of pivoting on this tableau we will pivot on a somewhat different tableau which differs in one respect. In the Phase I tableau we also include the objective row for the original LP. This is done to save the effort of having to compute the proper coefficients for this row after solving the auxiliary problem. Having these coefficients in hand at the end of Phase I allows one to immediately begin Phase II. This is illustrated in the following example. Here we denote the objective variable for the Phase I problem $\mathcal{Q}$ as w.

$$
\begin{array}{c ccccccc|c l}
 & x_0 & & & & & & & & \\
 & \downarrow & & & & & & & & \\
 & -1 & 2 & -1 & 2 & 1 & 0 & 0 & 4 & \\
 & \enclose{circle}{-1} & 2 & -3 & 1 & 0 & 1 & 0 & \text{-5} & \leftarrow \text{most negative} \\
 & -1 & -1 & 1 & -2 & 0 & 0 & 1 & -1 & \\
\cline{2-9}
z & 0 & 1 & -1 & 1 & 0 & 0 & 0 & 0 & \\
w & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \\
\cline{2-9}
\cline{2-9}
 & 0 & 0 & 2 & 1 & 1 & -1 & 0 & 9 & \\
 & 1 & -2 & 3 & -1 & 0 & -1 & 0 & 5 & \\
 & 0 & -3 & \enclose{circle}{4} & -3 & 0 & -1 & 1 & 4 & \\
\cline{2-9}
z & 0 & 1 & -1 & 1 & 0 & 0 & 0 & 0 & \\
w & 0 & -2 & 3 & -1 & 0 & -1 & 0 & 5 & \\
\cline{2-9}
\cline{2-9}
 & 0 & \frac{3}{2} & 0 & \frac{5}{2} & 1 & -\frac{1}{2} & -\frac{1}{2} & 7 & \\
 & 1 & \frac{1}{4} & 0 & \enclose{circle}{\frac{5}{4}} & 0 & -\frac{1}{4} & -\frac{3}{4} & 2 & \\
 & 0 & -\frac{3}{4} & 1 & -\frac{3}{4} & 0 & -\frac{1}{4} & \frac{1}{4} & 1 & \\
\cline{2-9}
z & 0 & \frac{1}{4} & 0 & \frac{1}{4} & 0 & -\frac{1}{4} & \frac{1}{4} & 1 & \\
w & 0 & \frac{1}{4} & 0 & \frac{5}{4} & 0 & -\frac{1}{4} & -\frac{3}{4} & 2 & \\
\cline{2-9}
\end{array}
$$

$$
\begin{array}{c c|cccccc|c l}
\hline
 & -2 & 1 & 0 & 0 & 1 & 0 & 1 & 3 & \text{Auxiliary problem} \\
 & \frac{4}{5} & \frac{1}{5} & 0 & 1 & 0 & -\frac{1}{5} & -\frac{3}{5} & \frac{8}{5} & \text{solved.} \\
 & \frac{3}{5} & -\frac{3}{5} & 1 & 0 & 0 & -\frac{2}{5} & -\frac{1}{5} & \frac{11}{5} & \text{Extract an initial} \\
\cline{3-9}
z & -\frac{1}{5} & \frac{4}{20} & 0 & 0 & 0 & -\frac{4}{20} & \frac{8}{20} & \frac{3}{5} & \text{feasible tableau.} \\
\cline{3-9}
w & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \\
\end{array}
$$

We have solved the Phase I problem. The optimal value in the Phase I problem is zero. Hence the original problem is feasible. In addition, we can extract from the tableau above an initial feasible tableau for the the original LP.

$$
\begin{array}{cccccc|c}
1 & 0 & 0 & 1 & 0 & \textcircled{1} & 3 \\
\frac{1}{5} & 0 & 1 & 0 & -\frac{1}{5} & -\frac{3}{5} & \frac{8}{5} \\
-\frac{3}{5} & 1 & 0 & 0 & -\frac{2}{5} & -\frac{1}{5} & \frac{11}{5} \\
\hline
\frac{1}{5} & 0 & 0 & 0 & -\frac{1}{5} & \frac{2}{5} & \frac{3}{5} \\
\hline\hline
1 & 0 & 0 & 1 & 0 & 1 & 3 \\
\frac{4}{5} & 0 & 1 & \frac{3}{5} & -\frac{1}{5} & 0 & \frac{17}{5} \\
-\frac{2}{5} & 1 & 0 & \frac{1}{5} & 0 & 0 & \frac{14}{5} \\
\hline
-\frac{1}{5} & 0 & 0 & -\frac{2}{5} & -\frac{1}{5} & 0 & -\frac{3}{5}
\end{array}
$$

Hence the optimal solution is

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 2.8 \\ 3.4 \end{pmatrix}$$

with optimal value $z = .6$.

Recapping, we have

The Two Phase Simplex Algorithm

Phase I Formulate and solve the auxiliary problem. Two outcomes are possible:

- (i) The optimal value in the auxiliary problem is positive. In this case the original problem is infeasible.
- (ii) The optimal value is zero and an initial feasible tableau for the original problem is obtained.

Phase II If the original problem is feasible, apply the simplex algorithm to the initial feasible tableau obtained from Phase I above. Again, two outcomes are possible:

- (i) The LP is determined to be unbounded.
- (ii) An optimal basic feasible solution is obtained.

Clearly the two phase simplex algorithms can be applied to solve any LP. This yields the following theorem.

THEOREM: [THE FUNDAMENTAL THEOREM OF LINEAR PROGRAMMING] *Every LP has the following three properties:*

(i) If it has no optimal solution, then it is either infeasible or unbounded.

(ii) If it has a feasible solution, then it has a basic feasible solution.

(iii) If it is bounded, then it has an optimal basic feasible solution.

Proof: The first phase of the two-phase simplex algorithm either discovers that the problem is infeasible or produces a basic feasible solution. The second phase of the two-phase simplex algorithm either discovers that the problem is unbounded or produces an optimal basic feasible solution. ■

4 Duality Theory

Recall from Section 1 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

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

Since the problem $\mathcal{D}$ is a linear program, it too has a dual. The *duality* terminology suggests that the problems $\mathcal{P}$ and $\mathcal{D}$ come as a pair implying that the dual to $\mathcal{D}$ should be $\mathcal{P}$. This is indeed the case as we now show:

$$\begin{array}{ll} \text{minimize} & b^T y \\ \text{subject to} & A^T y \ge c, \\ & 0 \le y \end{array} \quad = \quad \begin{array}{ll} -\text{maximize} & (-b)^T y \\ \text{subject to} & (-A^T) y \le (-c), \\ & 0 \le y. \end{array}$$

The problem on the right is in standard form so we can take its dual to get the LP

$$\begin{array}{ll} \text{minimize} & (-c)^T x \\ \text{subject to} & (-A^T)^T x \ge (-b),\ 0 \le x \end{array} \quad = \quad \begin{array}{ll} \text{maximize} & c^T x \\ \text{subject to} & Ax \le b,\ 0 \le x\ . \end{array}$$

The primal-dual pair of LPs $\mathcal{P} - \mathcal{D}$ are related via the Weak Duality Theorem.

THEOREM 4.1 (WEAK DUALITY THEOREM) *If $x \in \mathsf{R}^n$ is feasible for $\mathcal{P}$ and $y \in \mathsf{R}^m$ is feasible for $\mathcal{D}$, then*

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

Thus, if $\mathcal{P}$ is unbounded, then $\mathcal{D}$ is necessarily infeasible, and if $\mathcal{D}$ is unbounded, then $\mathcal{P}$ is necessarily infeasible. Moreover, if $c^T \bar{x} = b^T \bar{y}$ with $\bar{x}$ feasible for $\mathcal{P}$ and $\bar{y}$ feasible for $\mathcal{D}$, then $\bar{x}$ must solve $\mathcal{P}$ and $\bar{y}$ must solve $\mathcal{D}$.

We now use The Weak Duality Theorem in conjunction with The Fundamental Theorem of Linear Programming to prove the *Strong Duality Theorem.* The key ingredient in this proof is the general form for simplex tableaus derived at the end of Section 2 in (2.5).

THEOREM 4.2 (THE STRONG DUALITY THEOREM) *If either $\mathcal{P}$ or $\mathcal{D}$ has a finite optimal value, then so does the other, the optimal values coincide, and optimal solutions to both $\mathcal{P}$ and $\mathcal{D}$ exist.*

REMARK: 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

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

This problem has a finite optimal value, namely zero; however, this value is not attained by any point $x \in \mathsf{R}$. 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.

Proof: 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. By our formula for the general form of simplex tableaus (2.5), we know that there exists a nonsingular record matrix $R \in \mathsf{R}^{n\times n}$ and a vector $y \in \mathsf{R}^m$ such that the optimal tableau has the form

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

Since this is an optimal tableau we know that

$$c - A^Ty \le 0, \qquad\qquad -y^T \le 0$$

with y^Tb equal to optimal value in the primal problem. But then $A^Ty \ge c$ and $0 \le y$ so that y is feasible for the dual problem $\mathcal{D}$. In addition, the Weak Duality Theorem implies that

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

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.

4.1 General Duality Theory

Thus far we have discussed duality theory as it pertains to LPs in standard form. Of course, one can always transform any LP into one in standard form and then apply the duality theory. However, from the perspective of applications, this is cumbersome since it obscures the meaning of the dual variables. It is very useful to be able to compute the dual of an LP without first converting to standard form. In this section we show how this can easily be done. For this, we still make use of a standard form, but now we choose one that is much more flexible:

$$\begin{array}{lll} \mathcal{P} & \max & \sum_{j=1}^n c_jx_j \\ & \text{subject to} & \sum_{j=1}^n a_{ij}x_j \le b_i \qquad i \in I \\ & & \sum_{j=1}^n a_{ij}x_j = b_i \qquad i \in E \\ & & 0 \le x_j \qquad\qquad\quad j \in R \quad . \end{array}$$

Here the index sets I, E, and R are such that

$$I \cap E = \emptyset,\ I \cup E = \{1, 2, \dots, m\},\ \text{and } R \subset \{1, 2, \dots, n\}.$$

We use the following primal-dual correspondences to compute the dual of an LP.

In the Dual	In the Primal
Restricted Variables	Inequality Constraints
Free Variables	Equality Constraints
Inequality Constraints	Restricted Variables
Equality Constraints	Free Variables

Using these rules we obtain the dual to $\mathcal{P}$.

$$\begin{array}{lll} \mathcal{D} & \text{min} & \sum_{i=1}^m b_i y_i \\ & \text{subject to} & \sum_{i=1}^m a_{ij} y_i \geq c_j \quad j \in R \\ & & \sum_{i=1}^m a_{ij} y_i = c_j \quad j \in F \\ & & 0 \leq y_i \qquad\qquad i \in I \quad , \end{array}$$

where $F = \{1, 2, \dots, n\} \setminus R$.

For example, the LP

$$\begin{array}{lll} \text{maximize} & x_1 - 2x_2 + 3x_3 & \\ \text{subject to} & 5x_1 + x_2 - 2x_3 & \leq 8 \\ & -x_1 + 5x_2 + 8x_3 & = 10 \\ & x_1 \leq 10,\ 0 \leq x_3 & \end{array}$$

has dual

$$\begin{array}{lll} \text{minimize} & 8y_1 + 10y_2 + 10y_3 & \\ \text{subject to} & 5y_1 - y_2 + y_3 & = 1 \\ & y_1 + 5y_2 & = -2 \\ & -2y_1 + 8y_2 & \geq 3 \\ & 0 \leq y_1,\ 0 \leq y_3 & . \end{array}$$

The primal-dual pair $\mathcal{P}$ and $\mathcal{D}$ above are related by the following weak duality theorem.

THEOREM 4.3 *[General Weak Duality Theorem]*
Let $A \in \mathsf{R}^{m \times n}$, $b \in \mathsf{R}^m$, and $c \in \mathsf{R}^n$. If $x \in \mathsf{R}^n$ is feasible for $\mathcal{P}$ and $y \in \mathsf{R}^m$ is feasible for $\mathcal{D}$, then

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

Moreover, the following statements hold.

(i) If $\mathcal{P}$ is unbounded, then $\mathcal{D}$ is infeasible.

(ii) If $\mathcal{D}$ is unbounded, then $\mathcal{P}$ is infeasible.

(iii) If $\bar{x}$ is feasible for $\mathcal{P}$ and $\bar{y}$ is feasibe for $\mathcal{D}$ with $c^T \bar{x} = b^T \bar{y}$, then $\bar{x}$ is and optimal solution to $\mathcal{P}$ and $\bar{y}$ is an optimal solution to $\mathcal{D}$.

PROOF: Suppose $x \in \mathsf{R}^n$ is feasible for $\mathcal{P}$ and $y \in \mathsf{R}^m$ is feasible for $\mathcal{D}$. Then

$$\begin{aligned}
c^T x &= \sum_{j\in R} c_j x_j + \sum_{j \in F} c_j x_j \\
&\le \sum_{j\in R}(\sum_{i=1}^{m} a_{ij}y_i)x_j + \sum_{j\in F}(\sum_{i=1}^{m} a_{ij}y_i)x_j \\
&\qquad\qquad \text{(Since } c_j \le \textstyle\sum_{i=1}^{n} a_{ij}y_i \text{ and } x_j \ge 0 \text{ for } j \in R \\
&\qquad\qquad\quad \text{and } c_j = \textstyle\sum_{i=1}^{n} a_{ij}y_i \text{ for } j \in F.) \\
&= \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}y_i x_j \\
&= y^T A x \\
&= \sum_{i\in I}(\sum_{j=1}^{n} a_{ij}x_j)y_i + \sum_{i\in E}(\sum_{j=1}^{n} a_{ij}x_j)y_i \\
&\le \sum_{i\in I} b_i y_i + \sum_{i\in E} b_i y_i \\
&\qquad\qquad \text{(Since } \textstyle\sum_{j=1}^{n} a_{ij}x_j \le b_i \text{ and } 0 \le y_i \text{ for } i \in I \\
&\qquad\qquad\quad \text{and } \textstyle\sum_{j=1}^{n} a_{ij}x_j = b_i \text{ for } i \in E. \\
&= \sum_{i=1}^{m} b_i y_i \\
&= b^T y \ .
\end{aligned}$$

■

4.2 The Dual Simplex Algorithm

Consider the linear program

$$\begin{array}{lll}
\mathcal{P} & \text{maximize} & -4x_1 - 2x_2 - x_3 \\
& \text{subject to} & -x_1 - x_2 + 2x_3 \le -3 \\
& & -4x_1 - 2x_2 + x_3 \le -4 \\
& & x_1 + x_2 - 4x_3 \le 2 \\
& & 0 \le x_1, x_2, x_3 \ .
\end{array}$$

The dual to this LP is

$$\begin{array}{llr}
\mathcal{D} & \text{minimize} & -3y_1 - 4y_2 + 2y_3 \\
& \text{subject to} & -y_1 - 4y_2 + y_3 \ge -4 \\
& & -y_1 - 2y_2 + y_3 \ge -2 \\
& & 2y_1 + y_2 - 4y_3 \ge -1 \\
& & 0 \le y_1, y_2, y_3 \ .
\end{array}$$

The problem $\mathcal{P}$ does not have feasible origin, and so it appears that one must apply phase I of the two phase simplex algorithm to obtain an initial basic feasible solution. On the other hand, the dual problem $\mathcal{D}$ does have feasible origin, so why not just apply the simplex algorithm to $\mathcal{D}$ and avoid phase I altogether? This is exactly what we will do. However, we do it in a way that may at first seem odd. We *reverse* the usual simplex procedure by choosing a pivot row first, and then choosing the pivot column. The initial tableau for the problem $\mathcal{P}$ is

$$\begin{array}{|rrrrrr|r|}
-1 & -1 & 2 & 1 & 0 & 0 & -3 \\
-4 & -2 & 1 & 0 & 1 & 0 & -4 \\
1 & 1 & -4 & 0 & 0 & 1 & 2 \\
\hline
-4 & -2 & -1 & 0 & 0 & 0 & 0
\end{array}$$

A striking and important feature of this tableau is that every entry in the cost row is nonpositive! This is exactly what we are trying to achieve by our pivots in the simplex algorithm. This is a consequence of the fact that the dual problem $\mathcal{D}$ has feasible origin. Any tableau having this property we will call *dual feasible.* Unfortunately, the tableau is not feasible since some of the right hand sides are negative. Henceforth, we will say that such a tableau is not *primal feasible.* That is, instead of saying that a tableau (or dictionary) is feasible or infeasible in the usual sense, we wil now say that the tableau is *primal feasible*, respectively, *primal infeasible.*

Observe that if a tableau is *both* primal and dual feasible, then it must be optimal, i.e. the basic feasible solution that it identifies is an optimal solution. We now describe an implementation of the simplex algorithm, called the *dual simplex algorithm*, that can be applied to tableaus that are dual feasible but not primal feasible. Essentially it is the simplex algorithm applied to the dual problem but using the tableau structure associated with the primal problem. The goal is to use simplex pivots to attain primal feasibility while maintaining dual feasibility.

Consider the tableau above. The right hand side coefficients are -3, -4, and 2. These correspond to the cost coefficients of the dual objective. Not that this tableau also identifies a basic feasible solution for the dual problem by setting the dual variable equal to the negative of the cost row coefficients associated with the slack variables:

$$\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} .$$

The dual variables are currently nonbasic and so their values are zero. Next note that by increasing the value of either y_1 or y_2 we decrease the value of the dual objective since the coefficients of these variables are -3 and -4. In the simplex algorithm terminology, we can pivot on either the first or second row to decrease the value of the dual objective. Let's choose the first row as our pivot row. How do we choose the pivot column? Similar to the primal simplex algorithm, we choose the pivot column to maintain dual feasibility. For this we again must compute ratios, but this time it is the ratios of the negative entries in the pivot row with the corresponding cost row entries:

ratios for the first two columns are 4 and 2

$$\begin{array}{|cccccc|c|l}
-1 & \boxed{-1} & 2 & 1 & 0 & 0 & -3 & \leftarrow \text{pivot row} \\
-4 & -2 & 1 & 0 & 1 & 0 & -4 & \\
1 & 1 & -4 & 0 & 0 & 1 & 2 & \\
\hline
-4 & -2 & -1 & 0 & 0 & 0 & 0 &
\end{array}$$

The smallest ratio is 2 so the pivot column is column 2 in the tableau, and the pivot element is therefore the (1,2) entry of the tableau. Note that this process of choosing the pivot is the reverse of how the pivot is chosen in the primal simplex algorithm. In the dual simplex algorithm we fist choose a pivot row, then compute ratios to determine the pivot column which identifies the pivot. We now successive apply this process to the above tableau until optimality is achieved.

$$\begin{array}{|cccccc|c|l}
-1 & \boxed{-1} & 2 & 1 & 0 & 0 & -3 & \leftarrow \text{pivot row} \\
-4 & -2 & 1 & 0 & 1 & 0 & -4 & \\
1 & 1 & -4 & 0 & 0 & 1 & 2 & \\
\hline
-4 & -2 & -1 & 0 & 0 & 0 & 0 & \\
\hline
 & & & & & & & \\
1 & 1 & -2 & -1 & 0 & 0 & 3 & \\
-2 & 0 & -3 & -2 & 1 & 0 & 2 & \\
0 & 0 & \boxed{-2} & 1 & 0 & 1 & -1 & \leftarrow \text{pivot row} \\
\hline
-2 & 0 & -5 & -2 & 0 & 0 & 6 & \\
\hline
1 & 1 & 0 & -2 & 0 & -1 & 4 & \\
-2 & 0 & 0 & -7/2 & 1 & -3/2 & 7/2 & \\
0 & 0 & 1 & -1/2 & 0 & -1/2 & 1/2 & \\
\hline
-2 & 0 & 0 & -9/2 & 0 & -5/2 & 17/2 & \text{optimal}
\end{array}$$

Therefore, the optimal solutions to $\mathcal{P}$ and $\mathcal{D}$ are

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \\ 1/2 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 9/2 \\ 0 \\ 5/2 \end{pmatrix} ,$$

respectively, with optimal value $-17/2$.

Next consider the LP

$$\begin{array}{lll}
\mathcal{P} & \text{maximize} & -4x_1 - 2x_2 - x_3 \\
 & \text{subject to} & -x_1 - x_2 + 2x_3 \leq -3 \\
 & & -4x_1 - 2x_2 + x_3 \leq -4 \\
 & & x_1 + x_2 - x_3 \leq 2 \\
 & & 0 \leq x_1, x_2, x_3 \ .
\end{array}$$

This LP differs from the previous LP only in the x_3 coefficient of the third linear inequality. Let's apply the dual simplex algorithm to this LP.

$$\begin{array}{|cccccc|c|l}
-1 & \boxed{-1} & 2 & 1 & 0 & 0 & -3 & \leftarrow \text{pivot row}\\
-4 & -2 & 1 & 0 & 1 & 0 & -4 & \\
1 & 1 & -1 & 0 & 0 & 1 & 2 & \\
\hline
-4 & -2 & -1 & 0 & 0 & 0 & 0 & \\
\hline
\end{array}$$

$$\begin{array}{|cccccc|c|l}
1 & 1 & -2 & -1 & 0 & 0 & 3 & \\
-2 & 0 & -3 & -2 & 1 & 0 & 2 & \\
0 & 0 & 1 & 1 & 0 & 1 & -1 & \leftarrow \text{pivot row}\\
\hline
-2 & 0 & -5 & -2 & 0 & 0 & 6 & \\
\end{array}$$

The first dual simplex pivot is given above. Repeating this process again, we see that there is only one candidate for the pivot row in our dual simplex pivoting strategy. What do we do now? It seems as though we are stuck since there are no negative entries in the third row with which to compute ratios to determine the pivot column. What does this mean? Recall that we chose the pivot row because the negative entry in the right hand side implies that we can decrease the value of the dual objective by bring the dual variable y_3 into the dual basis. The ratios are computed to preserve dual feasibility. In this problem, the fact that there are no negative entries in the pivot row implies that we can increase the value of y_3 as much as we want without violating dual feasibility. That is, the dual problem is unbounded below, and so by the weak duality theorem the primal problem must be infeasible!

We will make extensive use of the dual simplex algorithm in our discussion of sensitivity analysis in linear programming.