Solving Linear Programs

1.0 Basic form

We go back to the form of the basic optimization problem we have considered, which is

$$\begin{aligned} &\min \ f(\underline{x}) \\ &\text{s.t.} \ \ \underline{h}(\underline{x}) = \underline{c} \\ &\qquad \underline{g}(\underline{x}) \leq \underline{b} \end{aligned} \qquad (1)$$

We recall that a linear program (LP) requires all functions *f*, $\underline{h}$, and $\underline{g}$ to be linear in the variables $\underline{x}$.

We make four comments regarding (1).

1. Equality constraints: The equality constraints may be converted into two inequality constraints via the following approach:

$$\underline{h}(\underline{x}) = \underline{c} \Leftrightarrow \begin{matrix} \underline{h}(\underline{x}) \leq \underline{c} \\ \underline{h}(\underline{x}) \geq \underline{c} \end{matrix} \qquad (2)$$

and we may then reverse the sign of the second inequality, resulting in:

$$\underline{h}(\underline{x}) = \underline{c} \Leftrightarrow \begin{matrix} \underline{h}(\underline{x}) \leq \underline{c} \\ -\underline{h}(\underline{x}) \leq -\underline{c} \end{matrix} \qquad (3)$$

This means we may include all of our equality constraints $\underline{h}(\underline{x})=\underline{c}$ from our general form (1) in our inequality constraints $\underline{g}(\underline{x})\leq \underline{b}$, so that the general form of our problem becomes:

$$\begin{aligned} &\min \ f(\underline{x}) \\ &\text{s.t.} \ \ \underline{g}(\underline{x}) \leq \underline{b} \end{aligned} \tag{4}$$

2. Nonnegativity constraints: The form given in (4) places no restriction on the sign of the decision variables in $\underline{x}$. Most problems require all decision variables to be nonnegative. For example, generation offers and demand bids are typically this way. This is convenient, because the algorithm we will present to solve LPs *requires* nonnegativity on the decision variables. It is the case, however, that some problems need to allow negativity for some or all decision variables. For example, we might like to develop a generation dispatch function that computes *changes* in generation rather than generation. For such cases, it is possible to convert one decision variable without the nonnegativity constraint into one, or two decision variables with nonnegativity constraints. We will not go into that here, but suffice it to say there are many references that describe how to do this, including [1, pg. 83-86]. Given this, we pose our general form of the optimization as

$$\begin{aligned} \min\ & f(\underline{x}) \\ \text{s.t.}\ & \underline{g}(\underline{x}) \leq \underline{b} \\ & \underline{x} \geq 0 \end{aligned} \qquad (5)$$

3. No lower bounds on g(x): It is possible that a problem has constraints like $c_l \leq g(\underline{x}) \leq c_h$. In this case, the right hand side $g(\underline{x}) \leq c_h$ is already in the correct form, but the left hand side $c_l \leq g(\underline{x})$ is not. In this case, we can multiply both sides by -1, resulting in

$$g(\underline{x}) \leq -c_l \qquad (6)$$

   The negative right-hand-side is addressed in the next bullet-point.
4. Negative right-hand-sides: We also require that all elements of $\underline{b}$ be nonnegative. This may seem to contradict the step we took in eq. (3) and (6) above. However, we will see at the end of these notes that it is possible to convert inequalities like these, with a negative right-hand-side, to the desirable form.

2.0 A simple solution approach

We concluded our last set of notes (IntroLP) with the statement

> "We will call such intersection points *corner points*. Therefore we see that our solution will always be at a corner point. This provides us with a basis for solution to LPs: Search the corner points!"

This is an effective approach, and if you take it, you will always find the right answer. However, you may also be doing a great deal of work. In our second example (Section 3.0) of our previous notes (IntroLP), we considered following LP:

$$\max f(x, y) = 5x + 8y$$

Subject to

$$40x + 30y \le 480 \quad \text{(person 1)}$$

$$24x + 32y \le 480 \quad \text{(person 2)}$$

$$20x + 24y \le 480 \quad \text{(person 3)}$$

$$x \ge 0, y \ge 0$$

and the constraints are visualized in Fig. 1.

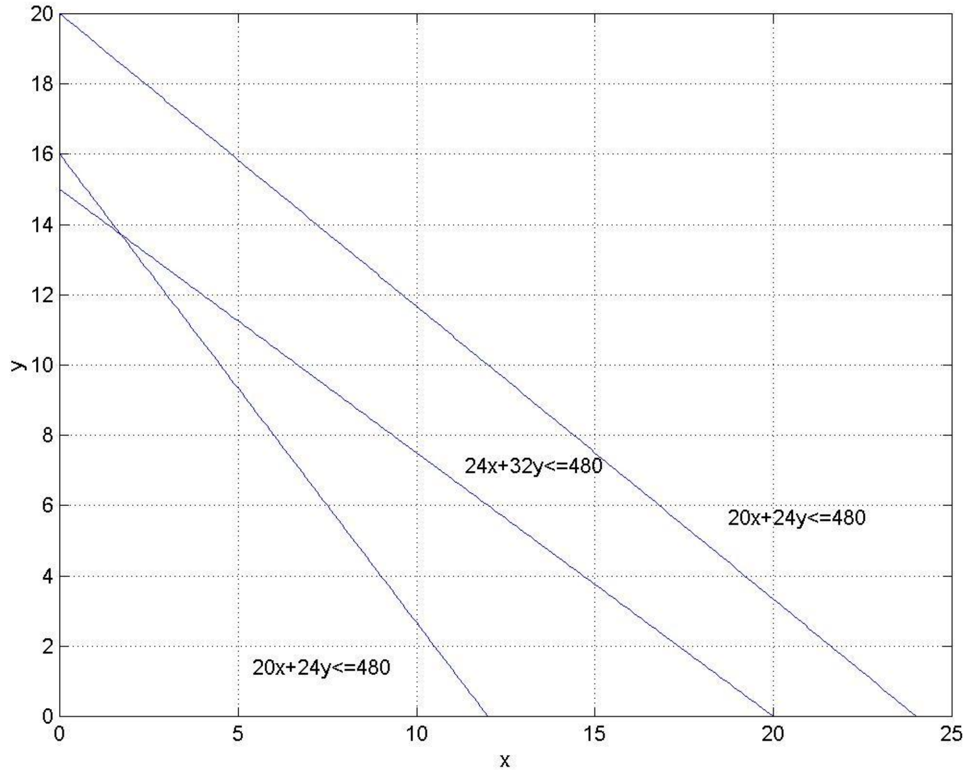

Fig. 1: Constraints for example problem

Here we see there are four corner points to search: (0,0), (12,0), (1.7143, 13.7143), and (0,15). However, we must make an important distinction here. These are *feasible corner points*. Because we know the solution must be feasible, these are the right points to search. Yet there are other infeasible corner points. For example, (0,16), (0,20), (20,0), and (24,0) are four such infeasible corner points. And there are two more that are outside the plane that we have plotted in Fig. 1. Counting them up, we see there are a total of 10 corner points, one for every pair of constraints.

This number of corner points is a combinatorial problem, characterized by 5 distinct things (constraints) taken 2 at a time, i.e.,

$$\begin{bmatrix} 5 \\ 2 \end{bmatrix} = \frac{5!}{2!(5-2)!} = \frac{5*4*3*2*1}{(2*1)(3*2*1)} = 10$$

We can easily distinguish feasible and infeasible corner points from Fig. 1. However, it will not be so easy for larger problems, especially when there are many decision variables and we cannot easily visualize the constraints in 2-D as we are doing here. One could perhaps devise a means to check them all, but it would be highly computational. For example, consider having just 40 constraints, there would be 780 corner points to check. Some problems have millions of constraints.

3.0 A better solution approach

So we need to develop a more effective strategy. To do so, let's consider a graphical portrayal of some LP, as shown in Fig. 2.

In Fig. 2, the dashed lines are the constraint boundaries, and the thick solid blue lines enclose the feasible region. The thin solid lines show the contours of constant objective function.

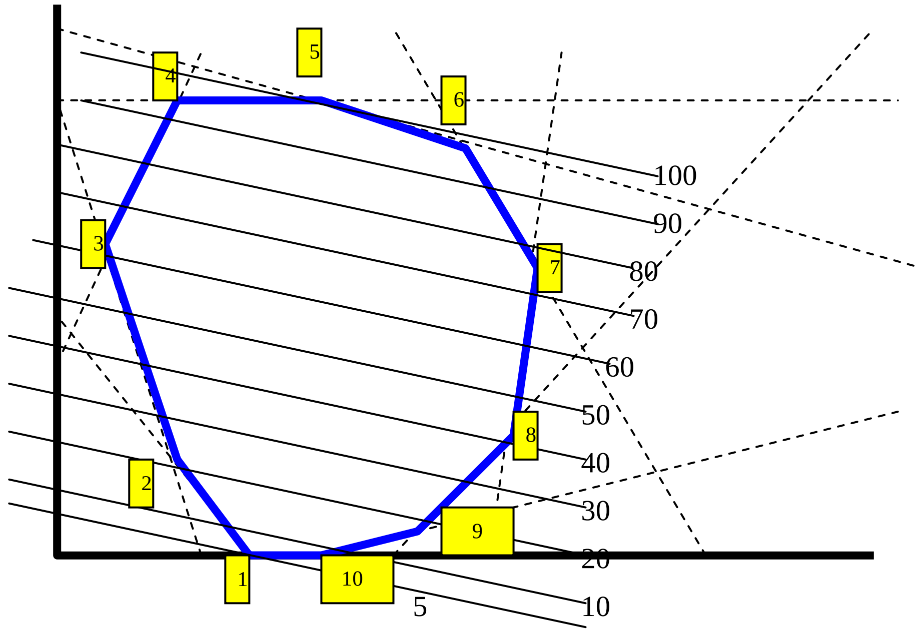

Fig. 2: Feasible region and contours for some LP

It is easy to see in Fig. 2 that if we are minimizing, the solution is corner point 1, and the minimum value is 5.

Likewise, if we are maximizing, the solution is corner point 5, and the maximum value is 100.

We consider a strategy for solving this problem. This strategy depends on the following two definitions:

- *Adjacent* corner points are connected by a single line segment on the boundary of the feasible region.
- One corner point is *better* than another if it has a higher value of the objective function *f*.

Our strategy is as follows:

1. Pick a corner point at random.
2. Move to an adjacent corner point that is better.
   a. If there are two that are better, move to the one that is best.
   b. If there are no better adjacent corner points, the current corner point is the solution to the problem.

Let's apply this strategy to the problem of Fig. 2, assuming we are maximizing the function. We also assume that we initially choose corner point 1.

From corner point 1, we can either move to corner point 2 or 11. But the objective function value at corner point 2, f_2=20, whereas the objective function value at corner point 11 is only f_{11}=10. Although both are better than f_1=5, we choose to move to corner point 2 since it is better.
From corner point 2, f_2=20, we can move to corner point 1, f_1=5 or corner point 3, $f_3=60$. Corner point 1 is not an option since it does not get better. But corner point 3 is better, with f_3=60, so we move there.

In like fashion, we move to corner point 4, f_4=95, and then to corner point 5, f_5=100.

At corner point 5, there are two options: corner point 4, with f_4=95, or corner point 6, with f_6=97. Both of these are worse than corner point 5. So we are done, and corner point 5 is the solution with f_5=100 as the maximum value of the problem.

From this example, we may conjecture a condition for optimality:

> If a corner point feasible solution is equal to or better than all its adjacent corner point feasible solutions, then it is equal to or better than all other corner point feasible solutions, i.e., it is *optimal*.

Formal proofs of this optimality condition are available in some texts; here, we simply state the essence of such proofs, which is contained in the following two points.

1. If the objective function monotonically increases (decreases) in some direction within the decision-vector space, then each adjacent corner point will become progressively better in the direction of objective function increase (decrease) such that the last corner point must have two adjacent corner points that are worse.
2. The monotonicity of objective function increase (decrease) is guaranteed by its linearity.

With the above optimality condition in place, we may outline the algorithm that we are going to study for solving linear programs. It is called the *Simplex Method*, and at a high level, is like this [1]:

1. Initialization: Start at a corner point solution.
2. Iterative step: Move to a better adjacent corner point feasible solution.
3. Optimality test: Determine if the current feasible corner point is optimal using our optimality test (if none of its adjacent feasible corner points are better, then the current feasible corner point is optimal).
    a. If the current feasible corner point is optimal, the solution has been found, and the method terminates.
    b. If the current feasible corner point is not optimal, then go to 2.

4.0 A word about the simplex method

The simplex method was developed in 1947 by George Dantzig (1914-2005) who worked for the US Air Force in the Pentegon to find better ways to plan the Air Force activities. He was trained as a mathematician but had significant experience in developing the plans that the air force required.

It is hard to conceptualize now, but at that time, there was no notion of an *objective function*. Neither was there any understanding that physical constraints on resources could be represented by linear inequalities. Dantzig recognized both of these; in addition, he developed the simplex method we are about to study.

The simplex method almost immediately revolutionized many fields, among which were planning, production, and economics.

It is interesting to note what brought this development to fruition:

- A war (and the needs of the Air Force)
- A person trained in mathematics and with significant practical experience in solving the problems at hand who, it seems, needed a job.

I have posted on our website a short and very readable summary paper written by George Dantzig in 2002 on how the simplex method came to be.

5.0 Setting up the simplex Method

The material of this section is adapted from [1].

The simplex method is comprised of a number of algebraic manipulations. These manipulations are made much easier if we first convert the inequality constraints into equality constraints by introducing *slack variables*.

We use another simple example to explain this idea. The example is as follows:

$$\begin{aligned} \max F &= 3x_1 + 5x_2 \\ \text{s.t.}\quad x_1 \qquad &\le 4 \\ 2x_2 &\le 12 \\ 3x_1 + 2x_2 &\le 18 \\ x_1 \ge 0, x_2 &\ge 0 \end{aligned}$$

Consider the first constraint $x_1 \leq 4$. The slack variable for this constraint is

$$x_3 = 4 - x_1 \qquad (6)$$

which represents the "slack" between the two sides of the inequality $x_1 \leq 4$. If the "slack" is 0, then the inequality is satisfied with equality, and there is really "no slack." This variable cannot be negative, otherwise, $x_1 > 4$.

Therefore, the first inequality, $x_1 \leq 4$, may be replaced with

$$x_1 + x_3 = 4, \qquad x_3 \geq 0 \tag{7}$$

We may similarly introduce slack variables into the other constraints, so that our original LP is converted to the following equivalent LP.

$$\max F = 3x_1 + 5x_2$$

s.t.

$$\begin{aligned} x_1 \qquad\quad + x_3 \qquad\qquad\qquad &= 4 \\ 2x_2 \qquad + x_4 \qquad\quad &= 12 \\ 3x_1 \quad + 2x_2 \qquad\qquad\qquad + x_5 &= 18 \end{aligned} \tag{8}$$

$$x_1 \geq 0, x_2 \geq 0, x_3 \geq 0, x_4 \geq 0, x_5 \geq 0$$

We need a few definitions:

- *Equality form*: In contrast to the original inequality form, the equality form of the problem has all inequality constraints converted to equality constraints via introduction of slack variables.
- *Augmented solution*: A solution to the LP that includes appropriate values of the slack variables (in addition to the values of the decision variables). For example, a solution to the original LP may be stated as $(x_1, x_2) = (3,2)$ whereas an augmented solution would be stated as $(x_1, x_2, x_3, x_4, x_5) = (3,2,1,8,5)$.

- *Basic solution*: An augmented corner-point solution. For example, consider the corner point infeasible solution (4,6) in the example. Augmenting it with the slack variable values x_3=0, x_4=0, and x_5=-6 yields the corresponding basic solution (4,6,0,0,-6). This basic solution is infeasible, as indicated by the presence of the negative slack variable x_5. This point is illustrated by the 'O' in Fig. 3.
- *Basic feasible solution*: A feasible augmented corner-point solution. For example, consider the corner point feasible solution (0,6) in the example. Augmenting it with the slack variables x_3=4, x_4=0, and x_5=6 yields the corresponding basic feasible solution (0,6,4,0,6). This basic solution is feasible. This point is illustrated by the 'X' in Fig. 3.

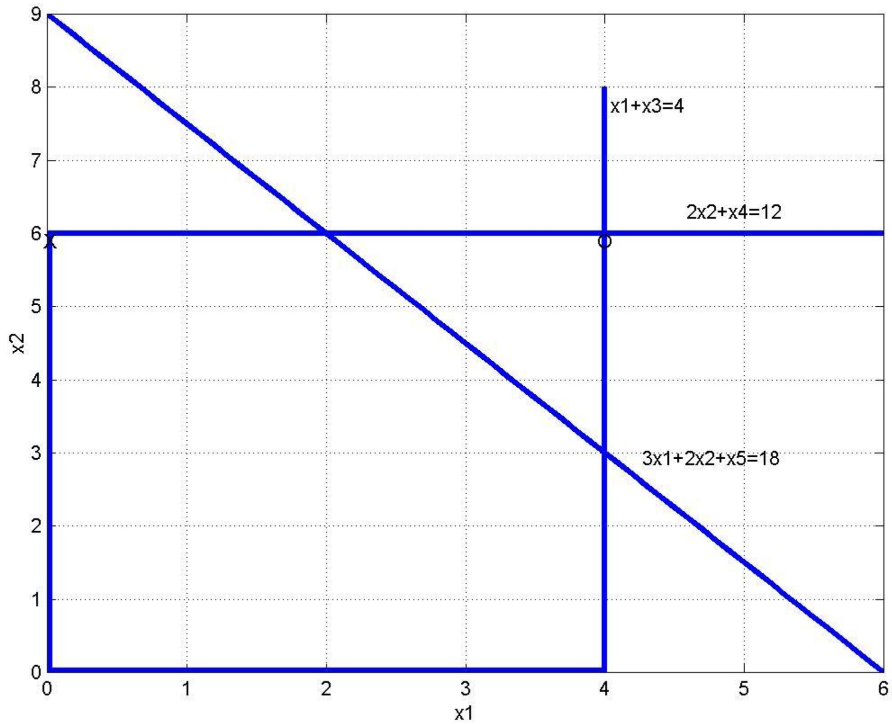

Fig. 3: Illustration

6.0 Final comment on general form of problem

Recall our general LP form:

$$\begin{aligned} &\min\ f(\underline{x}) \\ &\text{s.t. } \underline{g}(\underline{x}) \le \underline{b} \\ &\underline{x} \ge 0 \end{aligned} \qquad (5)$$

We mentioned in Section 1.0 of these notes, under point 3, that we required all elements of $\underline{b}$ to be nonnegative. Yet, in our point #1, we indicated we could handle equality constraints via the following transformation:

$$\underline{h}(\underline{x}) = \underline{c} \Leftrightarrow \begin{array}{c} \underline{h}(\underline{x}) \le \underline{c} \\ -\underline{h}(\underline{x}) \le -\underline{c} \end{array} \qquad (3)$$

This results in a negative right-hand-side; assuming elements of $\underline{c}$ are all positive, then the second equation above would have all negative right-hand sides.

The way to handle this is to first convert the problem into equality form via introduction of slack variables. Then another slack variable can be added for all equations having negative right-hand-sides. An example will clarify.

Let's assume that our last statement of our example problems has one of the constraints with a negative right-hand-side, per below (the last equation, with -18 as the negative right-hand-side).

$$\max F = 3x_1 + 5x_2$$

s.t.

$$\begin{array}{llllllll} x_1 & & + x_3 & & & = & 4 \\ & 2x_2 & & + x_4 & & = & 12 \\ 3x_1 & + 2x_2 & & & + x_5 & = & -18 \end{array}$$

$$x_1 \ge 0, x_2 \ge 0, x_3 \ge 0, x_4 \ge 0, x_5 \ge 0$$

Let's extract the one equation with the negative right-hand-side:

$$3x_1 + 2x_2 + x_5 = -18 \qquad (9)$$

Now we multiply by -1 to get

$$-3x_1 - 2x_2 - x_5 = 18 \qquad (10)$$

Although this creates the positive right-hand-side that we need, we will see later on that our initialization procedure to find a feasible solution will fail here, because it will result in x_5=-18 and therefore violates variable nonnegativity. As a result, we must add another slack variable here, resulting in

$$-3x_1 - 2x_2 - x_5 + x_6 = 18 \qquad (10)$$

Now we have that the right-hand-side is positive and our initialization procedure will result in $x_6=18$, satisfying decision variable nonnegativity and nonnegativity on the element in $\underline{b}$.

---

[1] F. Hillier and G. Lieberman, "Introduction to Operations Research," 4th edition, Holden-Day, Oakland California, 1986.