ISyE 3133 Lecture Notes
©Shabbir Ahmed

Linear Programming: Geometry, Algebra and the Simplex Method

A linear programming problem (LP) is an optimization problem where all variables are continuous, the objective is a linear (with respect to the decision variables) function , and the feasible region is defined by a finite number of linear inequalities or equations. LP[1] is possibly the best known and most frequently used branch of optimization. Beginning with the seminal work of Dantzig[2], LP has witnessed great success in real world applications from such diverse areas as sociology, finance, transportation, manufacturing and medicine. The importance and extensive use of LPs also come from the fact that the solution of certain optimization problems that are not LPs, e.g., integer, stochastic, and nonlinear programming problems, is often carried out by solving a sequence of related linear programs.

In this note, we discuss the geometry and algebra of LPs and present the Simplex method.

1.1 Geometry of LP

Recall that an LP involves optimizing a linear objective subject to linear constraints, and so can be written in the form

$$\min\ \{\mathbf{c}^\top \mathbf{x} :\ \mathbf{a}_i^\top \mathbf{x} \le b_i \ \ i = 1, \ldots, m\}.$$

An LP involving equality constraints can be written in the above form by replacing each equality constraint by two inequality constraints. The feasible region of an LP is of the form

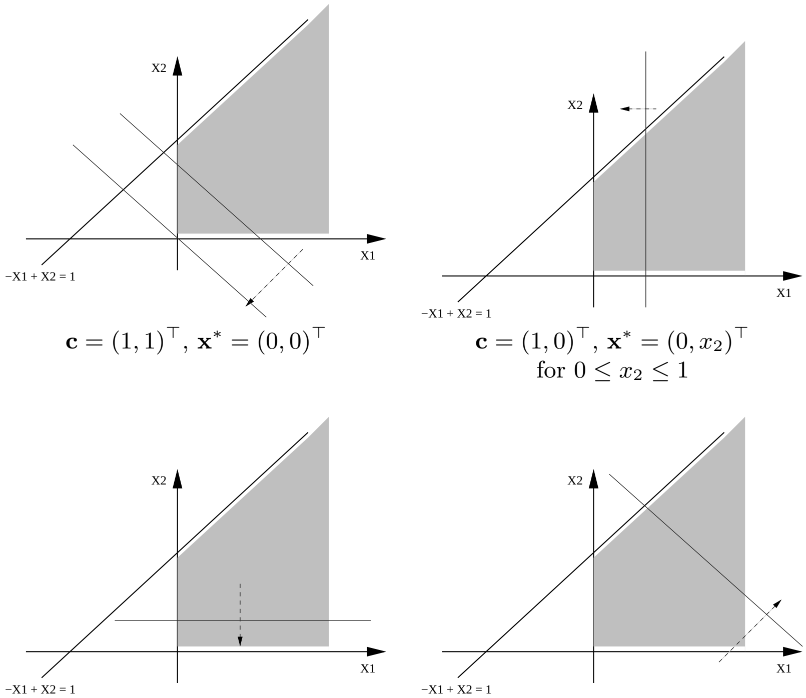
$$X = \{\mathbf{x} \in \mathbb{R}^n :\ \mathbf{a}_i^\top \mathbf{x} \le b_i \ \ i = 1, \ldots, m\}$$

which is a polyhedron (recall the definitions/properties of hyperplanes, halfspaces, and polyhedral sets). Thus an LP involves minimizing a linear function over a polyhedral set. Since both the objective function and constraint set are convex, an LP is a convex optimization problem. An LP could be either infeasible, unbounded or have an optimal solution.

Example 1.1 Figure 1.1 shows the set of optimal solutions of the feasible LP

$$\begin{array}{ll} \min & c_1x_1 + c_2x_2 \\ \text{s.t.} & -x_1 + x_2 \le 1,\ x_1 \ge 0,\ x_2 \ge 0. \end{array}$$

[1]We use LP for both linear programming and a linear programming problem.
[2]G.B. Dantzig. *Linear Programming and Extensions*, Princeton University Press, 1963



$\mathbf{c} = (1,1)^\top$, $\mathbf{x}^* = (0,0)^\top$

$\mathbf{c} = (1,0)^\top$, $\mathbf{x}^* = (0,x_2)^\top$ for $0 \leq x_2 \leq 1$

$\mathbf{c} = (0,1)^\top$, $\mathbf{x}^* = (x_1,0)^\top$ for $0 \leq x_1$

$\mathbf{c} = (-1,-1)^\top$ Unbounded.

Figure 1.1: Optimal solutions of the LP in Example 1.1

corresponding to different objective function vectors. This example illustrates LP could have a unique optimal solution or could have an infinite number of optimal solutions. Moreover the set of optimal solutions may also be unbounded. This example also illustrates that if an LP (whose feasible region does not contain a line) has an optimal solution then there is an extreme point (recall the definition of extreme points of convex sets) of the feasible region that is optimal.

Theorem 1.1 *Consider the linear program*

$$\min\{\mathbf{c}^\top \mathbf{x} : \ \mathbf{x} \in X\},$$

and suppose that the feasible region X does not contain a line. If the above LP has an optimal solution then there exists an extreme point of X which is optimal.

1.2 Algebra of LP

By Theorem 1.1 the search for an optimal solution to an LP could be restricted to just the finite set of extreme points of the feasible region (a polyhedron has a finite set of extreme points). To do this we need to have an algebraic characterization of extreme points. First we consider a fixed format for the LP known as the *standard form*. Next we consider an algebraic construction to obtain a specific type of feasible solutions, and show that such solutions correspond to extreme points of the feasible region.

Standard form

An LP in standard form has only equality constraints and non-negative variables. The objective function and constraints are simplified so that each variable appears only once. Any constant term in the objective function is not considered and the constraint system is a system of linear equations where all variables appear on the left and constants appear on the right hand side. An LP in any form can always be converted to standard form by including additional non-negative variables. For every $\leq$ constraint, add a "slack" variable to the left-hand-side, and transform the constraint to $=$ form. The slack variable is non-negative and has a zero cost coefficient. For every $\geq$ constraint, subtract a "slack" variable from the left-hand-side, and transform the constraint to $=$ form. Again the slack variable is non-negative and has a zero cost coefficient. Any variable restricted to be non-positive, i.e, a variable $x_i \leq 0$, should be replaced by a new variable x'_i such that $x'_i = -x_i$. Note that $x'_i \geq 0$. Any variable x_i that is unrestricted in sign, should be replaced by the difference of two new variables u_i and v_i such that $x_i = u_i - v_i$ and $u_i,\ v_i \geq 0$.

Example 1.2 Consider the LP

$$\begin{array}{rrcrcrcl}
\min & 2x_1 & + & 3x_2 & - & 2x_3 & & \\
\text{s.t.} & x_1 & - & x_2 & + & 2x_3 & \leq & 5 \\
 & 2x_1 & + & 3x_2 & - & 2x_3 & \geq & 7 \\
 & x_1 & + & 2x_2 & - & 3x_3 & = & 12 \\
 & x_1 & & & & & \geq & 0 \\
 & & & x_2 & & & \leq & 0 \\
 & & & & & x_3 & & \text{unrestricted},
\end{array}$$

and its standard form representation

$$\begin{array}{rrcrcrcrcrcrcr}
\min & 2x_1 & - & 3x'_2 & - & 2u_3 & + & 2v_3 & & & & & & \\
\text{s.t.} & x_1 & + & x'_2 & + & 2u_3 & - & 2v_3 & + & s_1 & & & = & 5 \\
 & 2x_1 & - & 3x'_2 & - & 2u_3 & + & 2v_3 & & & - & s_2 & = & 7 \\
 & x_1 & - & 2x'_2 & - & 3u_3 & + & 3v_3 & & & & & = & 12 \\
 & x_1 & , & x'_2 & , & u_3 & , & v_3 & , & s_1 & , & s_2 & \geq & 0.
\end{array}$$

Using indices $i = 1, \ldots, m$ for constraints and $j = 1, \ldots, n$ for variables, let x_j be the j-th variable, c_j be the objective coefficient of x_j, a_{ij} be the coefficient of x_j in the i-th constraint and b_i be the right-hand-side of the i-th constraint. A standard form LP is then

$$\begin{array}{lll}
\min & \sum_{j=1}^{n} c_j x_j & \\
\text{s.t.} & \sum_{j=1}^{n} a_{ij} x_j = b_i & \text{for all } i = 1, \ldots, m \\
 & x_j \geq 0 & \text{for all } j = 1, \ldots, n.
\end{array}$$

Using vector matrix notation an LP in standard form is

$$(LP) \quad \begin{array}{ll}
\min & \mathbf{c}^\top \mathbf{x} \\
\text{s.t.} & \mathbf{A}\mathbf{x} = \mathbf{b} \\
 & \mathbf{x} \geq \mathbf{0},
\end{array} \tag{1.1}$$

where $\mathbf{c} = \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$, and $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$. The j-th column of $\mathbf{A}$ is denoted by $\mathbf{A}_j$ and consists of the coefficients of the variable x_j in all equations.

Basic solutions and Basic feasible solutions

Any solution to the LP (1.1) has to satisfy $\mathbf{Ax} = \mathbf{b}$ and $\mathbf{x} \geq \mathbf{0}$. Let us ignore the non-negativity constraints $\mathbf{x} \geq \mathbf{0}$ for a moment. Note that $\mathbf{Ax} = \mathbf{b}$ is a system of linear equations with m equations and n unknowns. Let us assume that all rows of $\mathbf{A}$ are linearly independent, i.e., we ignore constraints that are obtained by taking linear combinations of other constraints. In other words we assume that $\mathbf{A}$ has *full row rank* (we will always make this assumption from now on). Then we must have $m \leq n$ otherwise the system does not have a solution. If $m = n$, then $\mathbf{Ax} = \mathbf{b}$ has a unique solution $\mathbf{x}^* = \mathbf{A}^{-1}\mathbf{b}$. (Note that $\mathbf{A}^{-1}$ exists since $\mathbf{A}$ has linearly independent rows and columns and hence is non-singular.) If $m < n$, then there are an infinite number of solutions. Consider the following special type of solutions. Fix $(n - m)$ of the variables to 0 and solve for the remaining m variables.

Example 1.3 Consider the following system (with $m = 2$ and $n = 4$):

$$\begin{array}{ccccccccc} x_1 & + & x_2 & + & x_3 & & & = & 6 \\ & & x_2 & & & + & x_4 & = & 3 \end{array}$$

Fixing $x_3 = x_4 = 0$ and then solving for the remaining variables: $x_1 = 3, x_2 = 3$, we have the solution $(3, 3, 0, 0)^\top$. Note that if we fix $x_2 = x_4 = 0$ then we cannot find a solution.

The above example illustrates that we must take care in deciding which variables to fix to zero. We need to fix $(n - m)$ of the variables to 0 in such a way that we can solve for the remaining m variables *uniquely*. The $(n - m)$ variables that are fixed to 0 are called *non-basic* variables, the m variables that are solved for are called *basic* variables, and the corresponding solution is called a *basic solution* to the system $\mathbf{Ax} = \mathbf{b}$. We now provide a precise algebraic description of basic solutions.

Consider the system $\mathbf{Ax} = \mathbf{b}$. Let us partition $\mathbf{x}$ as $\mathbf{x} = \begin{bmatrix} \mathbf{x}_B \\ \mathbf{x}_N \end{bmatrix}$, where $\mathbf{x}_B$ denote the vector of m basic variables and $\mathbf{x}_N$ be the vector of $(n - m)$ non-basic variables. Similarly, partition $\mathbf{A}$ as

$$\mathbf{A} = [\mathbf{A}_B, \mathbf{A}_N]$$

where $\mathbf{A}_B$ is an $m \times m$ matrix formed by the columns of $\mathbf{A}$ corresponding to m basic variables, and $\mathbf{A}_N$ is an $m \times (n - m)$ matrix formed by the columns of $\mathbf{A}$ corresponding to the $n - m$ non-basic variables. Then $\mathbf{Ax} = \mathbf{b}$ is equivalent to

$$\mathbf{A}_B\mathbf{x}_B + \mathbf{A}_N\mathbf{x}_N = \mathbf{b}.$$

Example 1.4 For the system in Example 1.3:

$$\underbrace{\begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}}_{\mathbf{x}} = \underbrace{\begin{bmatrix} 6 \\ 3 \end{bmatrix}}_{\mathbf{b}}.$$

Suppose $\mathbf{x}_B = (x_2, x_4)^\top$, i.e., $\mathbf{x}_N = (x_1, x_3)^\top$. Then we have

$$\underbrace{\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}}_{\mathbf{A}_B} \underbrace{\begin{bmatrix} x_2 \\ x_4 \end{bmatrix}}_{\mathbf{x}_B} + \underbrace{\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}}_{\mathbf{A}_N} \underbrace{\begin{bmatrix} x_1 \\ x_3 \end{bmatrix}}_{\mathbf{x}_N} = \underbrace{\begin{bmatrix} 6 \\ 3 \end{bmatrix}}_{\mathbf{b}}.$$

Setting $\mathbf{x}_N = (x_1, x_3)^\top = (0, 0)^\top$, the above system reduces to

$$\underbrace{\begin{bmatrix} x_2 \\ x_4 \end{bmatrix}}_{\mathbf{x}_B} = \underbrace{\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}^{-1}}_{\mathbf{A}_B^{-1}} \underbrace{\begin{bmatrix} 6 \\ 3 \end{bmatrix}}_{\mathbf{b}} = \begin{bmatrix} 6 \\ -3 \end{bmatrix}$$

Notice that we need $\mathbf{A}_B$ to be invertible to obtain a solution.

Let $\mathbf{A}_B$ be a $m \times m$ matrix formed by any m linearly independent columns of $\mathbf{A}$. Such a matrix is called a *basis* of $\mathbf{A}$. (The variables corresponding to these columns are the *basic* variables $\mathbf{x}_B$). The remaining columns form the matrix $\mathbf{A}_N$ and the associated *non-basic* variables are $\mathbf{x}_N$. Then a solution to $\mathbf{A}\mathbf{x} = \mathbf{b}$ is

$$\mathbf{x} = \begin{bmatrix} \mathbf{x}_B \\ \mathbf{x}_N \end{bmatrix} = \begin{bmatrix} \mathbf{A}_B^{-1}\mathbf{b} \\ \mathbf{0} \end{bmatrix}$$

Such a solution is called a *basic solution.*

Recall that until now we have ignored the non-negativity restriction in the feasible region $\mathbf{A}\mathbf{x} = \mathbf{b}$, $\mathbf{x} \geq \mathbf{0}$. By construction a basic solution $\mathbf{x}$ satisfies $\mathbf{A}\mathbf{x} = \mathbf{b}$ but not necessarily $\mathbf{x} \geq \mathbf{0}$. A basic solution that has all components non-negative is a *basic feasible solution* (bfs).

Example 1.5 For the system in Example 1.3 let us enumerate all choices of $m = 2$ out of $n = 4$ columns of $\mathbf{A}$ that lead to a basis and check if the corresponding solution is a bfs.

$\mathbf{A}_B$	$\mathbf{x}$	bfs?
$[\mathbf{A}_1, \mathbf{A}_2]$	$(3, 3, 0, 0)^\top$	Yes
$[\mathbf{A}_1, \mathbf{A}_4]$	$(6, 0, 0, 3)^\top$	Yes
$[\mathbf{A}_2, \mathbf{A}_3]$	$(0, 3, 3, 0)^\top$	Yes
$[\mathbf{A}_2, \mathbf{A}_4]$	$(0, 6, 0, -3)^\top$	No
$[\mathbf{A}_3, \mathbf{A}_4]$	$(0, 0, 6, 3)^\top$	Yes

where $\mathbf{A}_j$ is the column of $\mathbf{A}$ corresponding to variable x_j. Note that the columns $[\mathbf{A}_1, \mathbf{A}_3]$ do not form a basis since they are linearly dependent.

There are only $\binom{n}{m}$ ways to choose a set of m columns from n columns of $\mathbf{A}$. So the number of basis(es) of $\mathbf{A}$ and the number of basic solutions and hence basic feasible solutions is finite and at most $\binom{n}{m}$. For Example 1.3, $n = 4, m = 2$. The number of basic solutions $= \binom{4}{2} - 1 = 5$ and the number of basic feasible solutions is 4.

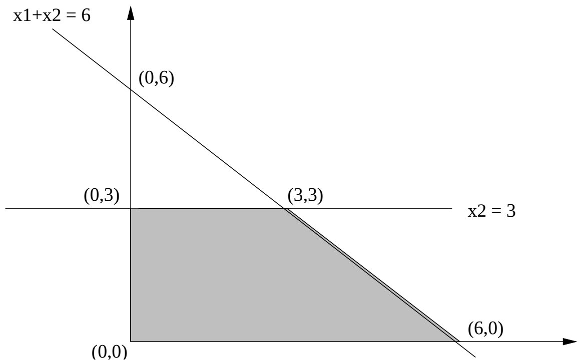


Figure 1.2: Feasible region of Example 1.6

Equivalence of bfs and extreme points

Now we state (without proof) that basic feasible solutions are precisely the extreme point solutions.

Theorem 1.2 *A solution* $\mathbf{x}$ *is an extreme point of the polyhedron* $X = \{\mathbf{x} \in \mathbb{R}^n : \ \mathbf{Ax} = \mathbf{b}, \ \mathbf{x} \geq \mathbf{0}\}$ *if and only it is a basic feasible solution of the system* $\mathbf{Ax} = \mathbf{b}, \ \mathbf{x} \geq \mathbf{0}$.

Example 1.6 Consider the LP feasible region defined by the constraints

$$x_1 + x_2 \leq 6, \ x_2 \leq 3, \ x_1 \geq 0 \ x_2 \geq 0.$$

Figure 1.2 shows that the above feasible region along with its extreme points. In standard form the above system is

$$\begin{array}{ccccccccc} x_1 & + & x_2 & + & x_3 & & & = & 6 \\ & & x_2 & & & + & x_4 & = & 3 \\ x_1, & & x_2, & & x_3, & & x_4, & \geq & 0 \end{array}$$

This is the system from Example 1.3. The basic solution and the basic feasible solutions/extreme points are

$\mathbf{A}_B$	$\mathbf{x}$	bfs/extreme point
$[\mathbf{A}_1, \mathbf{A}_2]$	$(3,3,0,0)^\top$	Yes
$[\mathbf{A}_1, \mathbf{A}_4]$	$(6,0,0,3)^\top$	Yes
$[\mathbf{A}_2, \mathbf{A}_3]$	$(0,3,3,0)^\top$	Yes
$[\mathbf{A}_2, \mathbf{A}_4]$	$(0,6,0,-3)^\top$	No
$[\mathbf{A}_3, \mathbf{A}_4]$	$(0,0,6,3)^\top$	Yes

Theorem 1.3 *If an LP* $\min\{\mathbf{c}^\top\mathbf{x} : \ \mathbf{Ax} = \mathbf{b}, \ \mathbf{x} \geq \mathbf{0}\}$ *has an optimal solution then there is a basic feasible solution that is optimal.*

Theorem 1.3 suggests a straight-forward scheme for solving LPs: enumerate all bfs (there are only finitely many of these) checking their objective function values, and choose one with the smallest objective value. Unfortunately for real-world LPs involving hundreds or thousands of variables and constraints enumerating all bfs is hardly practical. In the next section we provide a systematic way to search over bfs and solve LPs.

1.3 The Simplex Method

The basic idea of the Simplex method for LPs is to move from a bfs to an "adjacent" to improve the objective function value. Two bfs $\mathbf{x}^1$ and $\mathbf{x}^2$ are *adjacent* if one of the non-basic variables in $\mathbf{x}^1$ is basic in $\mathbf{x}^2$, and vice versa. The main steps of the Simplex algorithm are as follows:

0. Start from a bfs, if none exists STOP the problem is infeasible
1. Try to move to "adjacent" bfs to improve the objective function. There are three possibilities
   (a) There is no adjacent bfs which will improve the objective function, in this case STOP the current bfs is optimal
   (b) We discover the problem is unbounded, in this case STOP
   (c) There is adjacent bfs which improves the objective function, in this case REPEAT step 1 from this bfs

We postpone the discussion of how to find a starting bfs for later. Recall that moving from a bfs involves making one of the non-basic variables basic (i.e. increasing its value from zero). We first investigate the effect of this.

Effect of increasing non-basic variables

Suppose we are at a bfs, given by the basis matrix $\mathbf{A}_B$. (Recall we let B denote indices of the basic variables, and N denote the indices of the non-basic variables.) Since moving to an adjacent bfs entails increasing one of the current non-basic variables (i.e. one of the variables in $\mathbf{x}_N$) we need to track how such a change effects the basic variables ($\mathbf{x}_B$) and the objective function. To understand the effect on the basic variables, note that

$$\begin{aligned} & \mathbf{A}_B\mathbf{x}_B + \mathbf{A}_N\mathbf{x}_N && = \mathbf{b} \\ \Leftrightarrow \quad & \mathbf{x}_B && = \mathbf{A}_B^{-1}\mathbf{b} - \mathbf{A}_B^{-1}\mathbf{A}_N\mathbf{x}_N \\ \Leftrightarrow \quad & \mathbf{x}_B && = \mathbf{A}_B^{-1}\mathbf{b} - \sum_{j\in N}\mathbf{A}_B^{-1}\mathbf{A}_j x_j. \end{aligned} \tag{1.2}$$

Currently all non-basic variables are at zero, and so the basic variables have the value

$$\mathbf{x}_B \;=\; \mathbf{A}_B^{-1}\mathbf{b}.$$

According to (1.2), a unit increase in a non-basic variable x_j will decrease the vector of basic variables $\mathbf{x}_B$ by the vector $\mathbf{A}_B^{-1}\mathbf{A}_j$.

To understand the effect on the objective function value, observe that

$$\begin{aligned} z = \quad \mathbf{c}^\top\mathbf{x} & = \mathbf{c}_B^\top\mathbf{x}_B + \mathbf{c}_N^\top\mathbf{x}_N \\ & = \mathbf{c}_B^\top(\mathbf{A}_B^{-1}\mathbf{b} - \mathbf{A}_B^{-1}\mathbf{A}_N\mathbf{x}_N) + \mathbf{c}_N^\top\mathbf{x}_N \\ & = \mathbf{c}_B^\top\mathbf{A}_B^{-1}\mathbf{b} + (\mathbf{c}_N^\top - \mathbf{c}_B^\top\mathbf{A}_B^{-1}\mathbf{A}_N)\mathbf{x}_N \\ & = \mathbf{c}_B^\top\mathbf{A}_B^{-1}\mathbf{b} + \sum_{j\in N}(c_j - \mathbf{c}_B^\top\mathbf{A}_B^{-1}\mathbf{A}_j)x_j, \end{aligned} \tag{1.3}$$

where in line 2 above, we have plugged in the expression for $\mathbf{x}_B$ in terms of $\mathbf{x}_N$ from (1.2). Currently all non-basic variables are at zero, and so the objective value is

$$z \;=\; \mathbf{c}_B^\top\mathbf{A}_B^{-1}\mathbf{b}.$$

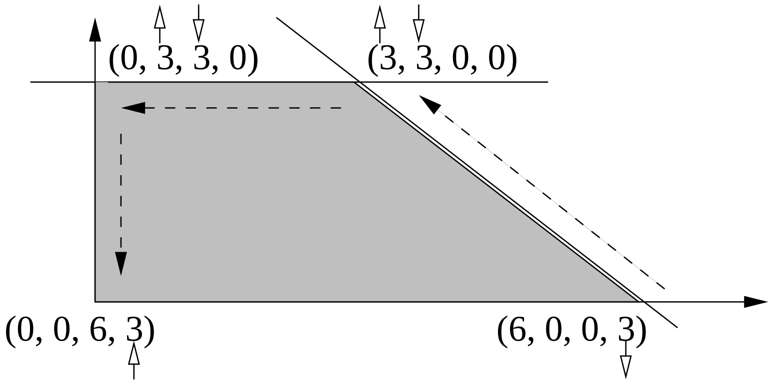


Figure 1.3: Consider the feasible region of Example 1.6. Moving from one bfs to an adjacent one corresponds to making one of the non-basic variables basic (i.e. increasing its value and making the corresponding tight constraint loose) and moving as far as possible until we hit another constraint (the slack of this constraint which was basic now becomes non-basic). E.g. Moving from (6,0,0,3) to the adjacent bfs (3,3,0,0) the basic variable x_4 becomes non-basic, and the non-basic variable x_2 becomes basic.

A unit change in a non-basic variable x_j will increase the objective by $(c_j - \mathbf{c}_B^\top \mathbf{A}_B^{-1}\mathbf{A}_j)$. This quantity is known as the reduced cost. Given a bfs defined by basic variables B, the *reduced cost* of a variable x_j is defined as

$$r_j = c_j - \mathbf{c}_B^\top \mathbf{A}_B^{-1}\mathbf{A}_j. \tag{1.4}$$

Note that reduced costs of basic variables is zero (why?).

Moving to an adjacent bfs

To move from the current bfs we have to increase the value of one of the non-basic variables from zero. Two questions arise:

1. Which non-basic variable in N should we consider increasing?

2. How much should the chosen non-basic variable be increased to?

1. Which non-basic variable in N should we consider increasing? Note that we are trying to improve the objective value from its current value of $\mathbf{c}_B^\top \mathbf{A}_B^{-1}\mathbf{b}$. According to (1.3) increasing the non-basic variable x_j increases the objective value at the rate of r_j (its reduced cost). Since we are *minimizing* we would like to *increase only one of the non-basic variables which has a negative reduced cost.*

Rule: Pick a non-basic variable which has a negative reduced cost. (Anyone will do if there are multiple). The chosen non-basic variable is said to *enter* the basis. If there are no non-basic variables with negative reduced cost, STOP the current bfs is optimal.

2.How much should the chosen non-basic variable be increased to? Suppose that we have picked a non-basic variable x_j with $j \in N$ with $r_j < 0$ to enter the basis. All other non-basic variables will remain at zero. From (1.2) the basic variables will change according to

$$\mathbf{x}_B = \mathbf{A}_B^{-1}\mathbf{b} - \mathbf{A}_B^{-1}\mathbf{A}_j x_j.$$

Let us introduce another bit of notation:

$$\mathbf{d}_B^j = \mathbf{A}_B^{-1}\mathbf{A}_j.$$

Note that $\mathbf{d}_B^j$ is an m-dimensional vector and its components d_i^j are associated with the basic variables $i \in B$.

Example 1.7 Consider the bfs $\mathbf{x} = (6,0,0,3)^\top$ for the system in Example 1.6 (see also Figure 1.3). Here $B = \{1,4\}$. Consider increasing the value of the non-basic variable x_2, i.e., $j = 2$. Here

$$\mathbf{d}_B^2 = \begin{bmatrix} d_1^2 \\ d_4^2 \end{bmatrix} = \mathbf{A}_B^{-1}\mathbf{A}_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Writing out (1.2) component-wise (i.e. for each basic variable) we get

$$x_i = (\mathbf{A}_B^{-1}\mathbf{b})_i - d_i^j x_j \quad \forall\, i \in B.$$

As we increase x_j, the i-th basic variable x_i will decrease at the rate of d_i^j. We need to make sure that all basic variables remain non-negative. If the quantity d_i^j is non-positive then the basic variable x_i will remain non-negative. If d_i^j is positive then the basic variable x_i will remain non-negative as long as the non-basic variable x_j satisfies

$$x_j \le \frac{(\mathbf{A}_B^{-1}\mathbf{b})_i}{d_i^j}.$$

The current value of the basic variable $x_i = (\mathbf{A}_B^{-1}\mathbf{b})_i$. This suggests the following rule:

Rule: Set the value of the entering (nonbasic) variable as follows:

$$x_j = \min\left\{ \frac{x_i}{d_i^j} : \quad i \in B \text{ s.t. } d_i^j > 0 \right\}. \tag{1.5}$$

The above formula is known as the minimum ratio test.

What if $d_j \le 0$ for all basic variables? Then we can increase the nonbasic variable x_j as much as we want and improve the objective. This then means that the problem is *unbounded.*

Note that the basic variable(s) which determines the minimum in the right-hand-side of the expression (1.5) becomes zero when we increase the value of the entering nonbasic variable. Thus one of these basic variables which are now zero becomes nonbasic. Now we repeat the process with the new basis matrix. This is the Simplex algorithm. A summary of the method is presented in Algorithm 1.1.

Algorithm 1.1 The Simplex Method

0. Initialize with a starting bfs $\mathbf{x}^0$. Let $\mathbf{A}_B$ be the associated basis. If no such bfs exists, STOP, the problem is infeasible.

1. For each non-basic variable $j \in N$ compute the reduced cost r_j. If all reduced costs are non-negative, STOP, the current bfs is optimal. Otherwise, compute $\mathbf{d}_B^j$ corresponding to a nonbasic variable x_j with $r_j < 0$.

2. If $\mathbf{d}_B^j \leq \mathbf{0}$, STOP, the problem is unbounded. Otherwise, compute the new value of x_j according to the minimum ratio test formula (1.5). Let x_i be (one of) the basic variable which become nonbasic.

3. The basic variable x_i leaves the basis and the non-basic variable x_j enters the basis. Update B and $\mathbf{A}_B$ and return to step 1.

Example 1.8 Consider the LP

$$\begin{array}{lccccccccc}
\min & 3x_1 & - & 2x_2 & & & & & & \\
\text{s.t.} & x_1 & + & x_2 & + & x_3 & & & = & 6 \\
& & & x_2 & & & + & x_4 & = & 3 \\
& x_1, & & x_2, & & x_3, & & x_4, & \geq & 0.
\end{array}$$

Following are iterations of the Simplex method starting from the bfs $(6, 0, 0, 3)^\top$.

#	$\mathbf{x}_B$	$\mathbf{x}_N$	$\mathbf{A}_B$	r_N	$\mathbf{x}_B$	$\mathbf{d}_B$	$\left[\frac{x_i}{-\mathbf{d}_i}\right]$
1.	$\begin{bmatrix} x_1 \\ x_4 \end{bmatrix}$	$\begin{bmatrix} x_2 \\ x_3 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} -5 \\ -5 \end{bmatrix}$	$\begin{bmatrix} 6 \\ 3 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 6 \\ 3 \end{bmatrix}$
2.	$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$	$\begin{bmatrix} x_3 \\ x_4 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} -3 \\ 5 \end{bmatrix}$	$\begin{bmatrix} 3 \\ 3 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 3 \\ - \end{bmatrix}$
3.	$\begin{bmatrix} x_2 \\ x_3 \end{bmatrix}$	$\begin{bmatrix} x_1 \\ x_4 \end{bmatrix}$	$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}$	$\begin{bmatrix} 3 \\ 2 \end{bmatrix}$	$\begin{bmatrix} 3 \\ 3 \end{bmatrix}$	-	-

---

Finding an initial bfs

The Simplex method requires an initial basic feasible solution for the system $\mathbf{A}\mathbf{x} = \mathbf{b}$, $\mathbf{x} \geq \mathbf{0}$. If the constraint matrix $\mathbf{A}$ has an $m \times m$ identity submatrix $\mathbf{I}$, we can use $\mathbf{I}$ as the initial basis. Otherwise, we can suitably multiply the constraints by ± 1 so that $\mathbf{b} \geq \mathbf{0}$ and then introduce auxiliary variables $\mathbf{y} = (y_1, \ldots, y_m)^\top$, to construct the following *Phase-I* LP:

$$\begin{array}{ll}
\min & \mathbf{e}^\top \mathbf{y} \\
\text{s.t.} & \mathbf{A}\mathbf{x} + \mathbf{I}\mathbf{y} = \mathbf{b}, \\
& \mathbf{x},\ \mathbf{y} \geq \mathbf{0},
\end{array}$$

where $\mathbf{I}$ is an $m \times m$ identity matrix and $\mathbf{e}$ is a vector of ones. We can now solve the above LP using $\mathbf{I}$ as the starting basis. If the optimal objective value is zero, implying $\mathbf{y} = \mathbf{0}$, then we have a basis consisting of only the columns of $\mathbf{A}$, which we can use as an initial BFS for the original problem. However if the optimal objective value greater than 0, then the original problem is infeasible.

Degeneracy, stalling, cycling and termination

If the current basic feasible solution is such that one (or some) of the basic variables has a value of zero, then the minimum ration test (1.5) gives a value of 0 for the entering variable. Such a bfs is known as a *degenerate* bfs. In this case, the basic variable with a value of zero, leaves the basis. So the basis changes but the actual solution does not change (since none of variable values has changed). So we have more than one basis corresponding to the same bfs (extreme point). We have gone through an iteration without making any improvement.

Example 1.9 Consider the system

$$\begin{array}{rcccccccccccl} - & x_1 & + & x_2 & + & s_1 & & & & & = & 0 \\ & x_1 & & & & & - & s_2 & & & = & 1 \\ & & & x_2 & & & & & - & s_3 & = & 1 \\ & x_1, & & x_2, & & s_1, & & s_2, & & s_3 & \geq & 0 \end{array}$$

and the three basis defined by choosing the basic variables as $\mathbf{x}_B = (x_1, x_2, s_1)^\top$, $\mathbf{x}_B = (x_1, x_2, s_2)^\top$, and $\mathbf{x}_B = (x_1, x_2, s_3)^\top$, respectively. All three basis define the same bfs $\mathbf{x} = (1, 1, 0, 0, 0)^\top$.

In the presence of degeneracy there might be several simplex iterations (change of basis) without any progress (improvement in the objective value). This is known as *stalling*. It may also happen that starting from a basis, the algorithm may go through several iterations (change of basis) and return to the starting basis. This is known as *cycling*. Figure 1.4 provides an example of cycling of the Simplex method. In case of degeneracy, there are ties in choosing which non-basic variable (with negative reduced cost) to enter the basis, and which basic variable (that determines λ) to leave the basis. Cycling can be avoided by careful choice when breaking ties.

With proper care in breaking ties) the Simplex method is guaranteed to terminate finitely with an optimal solution to the LP. In the worst case, it might require $\begin{pmatrix} n \\ m \end{pmatrix}$ steps to terminate (since it may require exploring all bfs). In most practical problems the performance is much better. The practical performance is sensitive to the number of rows m. The optimal solution produced by the simplex algorithm is a bfs, therefore, (at most) m of the n variables are positive.

An example of cycling

$$\begin{aligned} \min\ & -2x_1 - 3x_2 + x_3 + 12x_4 \\ s.t.\ & -2x_1 - 9x_2 + x_3 + 9x_4 + x_5 = 0 \\ & \frac{1}{3}x_1 + x_2 - \frac{1}{3}x_3 - 2x_4 + x_6 = 0 \\ & x_1, x_2, x_3, x_4, x_5, x_6 \geq 0 \end{aligned}$$

$$c^T = (-2, -3, 1, 12, 0, 0)$$

$$A = \begin{bmatrix} -2 & -9 & 1 & 9 & 1 & 0 \\ 1/3 & 1 & -1/3 & -2 & 0 & 1 \end{bmatrix}$$

$$b = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Iteration #	x_B	x_N	A_B	$r_N = c_N - (c_B^T A_B^{-1} A_N)^T$	$x_B = A_B^{-1} b$	$d_B = A_B^{-1} A_j$	$\frac{(x_B)_i}{(d_B)_i}$	
1.	$\begin{bmatrix} x_5 \\ x_6 \end{bmatrix}$	$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} -2 \\ -3 \leftarrow \\ 1 \\ 12 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} -9 \\ 1 \end{bmatrix}$	$\begin{bmatrix} - \\ 0 \rightarrow \end{bmatrix}$	x_2 enters x_6 leaves
2.	$\begin{bmatrix} x_2 \\ x_5 \end{bmatrix}$	$\begin{bmatrix} x_1 \\ x_3 \\ x_4 \\ x_6 \end{bmatrix}$	$\begin{bmatrix} -9 & 1 \\ 1 & 0 \end{bmatrix}$	$\begin{bmatrix} -1 \leftarrow \\ 0 \\ 6 \\ 3 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1/3 \\ 1 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \rightarrow \end{bmatrix}$	x_1 enters x_5 leaves
3.	$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$	$\begin{bmatrix} x_3 \\ x_4 \\ x_5 \\ x_6 \end{bmatrix}$	$\begin{bmatrix} -2 & -9 \\ 1/3 & 1 \end{bmatrix}$	$\begin{bmatrix} -2 \\ -3 \leftarrow \\ 1 \\ 12 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} -9 \\ 1 \end{bmatrix}$	$\begin{bmatrix} - \\ 0 \rightarrow \end{bmatrix}$	x_4 enters x_2 leaves
4.	$\begin{bmatrix} x_1 \\ x_4 \end{bmatrix}$	$\begin{bmatrix} x_2 \\ x_3 \\ x_5 \\ x_6 \end{bmatrix}$	$\begin{bmatrix} -2 & 9 \\ 1/3 & 2 \end{bmatrix}$	$\begin{bmatrix} 3 \\ -1 \leftarrow \\ 0 \\ 6 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 1/3 \end{bmatrix}$	$\begin{bmatrix} 0 \rightarrow \\ 0 \end{bmatrix}$	x_3 enters x_1 leaves
5.	$\begin{bmatrix} x_3 \\ x_4 \end{bmatrix}$	$\begin{bmatrix} x_1 \\ x_2 \\ x_5 \\ x_6 \end{bmatrix}$	$\begin{bmatrix} 1 & 9 \\ -1/3 & -2 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 12 \\ -2 \\ -3 \leftarrow \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} -9 \\ 1 \end{bmatrix}$	$\begin{bmatrix} - \\ 0 \rightarrow \end{bmatrix}$	x_6 enters x_4 leaves
6.	$\begin{bmatrix} x_3 \\ x_6 \end{bmatrix}$	$\begin{bmatrix} x_1 \\ x_2 \\ x_4 \\ x_5 \end{bmatrix}$	$\begin{bmatrix} 1 & 0 \\ -1/3 & 1 \end{bmatrix}$	$\begin{bmatrix} 0 \\ 6 \\ 3 \\ -1 \leftarrow \end{bmatrix}$	$\begin{bmatrix} 0 \\ 0 \end{bmatrix}$	$\begin{bmatrix} 1 \\ 1/3 \end{bmatrix}$	$\begin{bmatrix} 0 \rightarrow \\ 0 \end{bmatrix}$	x_5 enters x_3 leaves
7.	$\begin{bmatrix} x_5 \\ x_6 \end{bmatrix}$	We are back to the same basis we started with.						

Figure 1.4: Cycling of the Simplex method