

3 The Simplex Method

3.1 Basic Solutions

In the LP of Example 2.3, the optimal solution happened to lie at an extreme point of the feasible set. This was not a coincidence. Consider an LP in general form,

$$\text{maximize } c^T x \text{ subject to } Ax \leq b, x \geq 0. \tag{3.1}$$

The feasible set of this LP is a convex polytope in \mathbb{R}^n , i.e., an intersection of half-spaces. Each level set of the objective function $c^T x$, i.e., each set $L_\alpha = \{x \in \mathbb{R}^n : c^T x = \alpha\}$ of points for which the value of the objective function is equal to some constant $\alpha \in \mathbb{R}$, is a k -dimensional flat for some $k \leq n$. The goal is to find the largest value of α for which $L_\alpha(f)$ intersects with the feasible set. If such a value exists, the intersection contains either a single point or an infinite number of points, and it is guaranteed to contain an extreme point of the feasible set. This fact is illustrated in Figure 3.1, and we will give a proof momentarily.

The geometric characterization of extreme points, as points that cannot be written as a convex combination of two different points, is somewhat hard to work with. We therefore use an alternative, algebraic characterization. To this end, consider the following LP in standard form, which can be obtained from (3.1) by introducing slack variables:

$$\text{maximize } c^T x \text{ subject to } Ax = b, x \geq 0, \tag{3.2}$$

where $A \in \mathbb{R}^{m \times n}$ and $b \in \mathbb{R}^m$. Call a solution $x \in \mathbb{R}^n$ of the equation $Ax = b$ *basic* if at most m of its entries are non-zero, i.e., if there exists a set $B \subseteq \{1, \dots, n\}$ with $|B| = m$ such that $x_i = 0$ if $i \notin B$. The set B is then called *basis*, and variable x_i is called *basic* if $i \in B$ and *non-basic* if $i \notin B$. A basic solution x that also satisfies $x \geq 0$ is a *basic feasible solution* (BFS) of (3.2). We finally distinguish basic solutions that

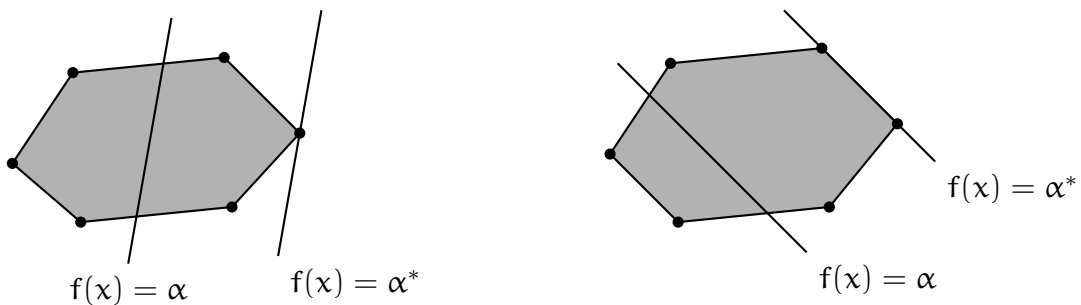


Figure 3.1: Illustration of linear programs with one optimal solution (left) and an infinite number of optimal solutions (right)

have exactly m non-zero entries from those that have strictly fewer, and refer to the latter as degenerate.

In what follows we will assume that (i) the rows of A are linearly independent and that (ii) every set of m columns of A are linearly independent. These assumptions are without loss of generality: if a set of rows are linearly dependent, one of the corresponding constraints can be removed without changing the feasible set; similarly, if a set of columns are linearly dependent, one of the corresponding variables can be removed.

3.2 Extreme Points and Optimal Solutions

It turns out that the extreme points of the feasible set are precisely the basic feasible solutions.

THEOREM 3.1. *A vector is a basic feasible solution of $Ax = b$ if and only if it is an extreme point of the set $X(b) = \{x : Ax = b, x \geq 0\}$.*

Proof. Consider a BFS x and suppose that $x = \delta y + (1 - \delta)z$ for $y, z \in X(b)$ and $\delta \in (0, 1)$. Since $y \geq 0$ and $z \geq 0$, $x = \delta y + (1 - \delta)z$ implies that $y_i = z_i = 0$ whenever $x_i = 0$. This means in particular that $y - z$ has at most m non-zero entries. At the same time, $y, z \in X(b)$ implies that $Ay = b = Az$ and thus $A(y - z) = 0$. This yields a linear combination of at most m columns of A that is equal to zero, which by (ii) implies that $y = z$. Thus x is an extreme point of $X(b)$.

Now consider a feasible solution $x \in X(b)$ that is not a BFS. Let i_1, \dots, i_r be the rows of x that are non-zero, and observe that $r > m$. This means that the columns a^{i_1}, \dots, a^{i_r} , where $a^i = (a_{1i}, \dots, a_{mi})^T$, have to be linearly dependent, i.e., there has to exist a collection of r non-zero numbers y_{i_1}, \dots, y_{i_r} such that $y_{i_1}a^{i_1} + \dots + y_{i_r}a^{i_r} = 0$. Extending y to a vector in \mathbb{R}^n by setting $y_i = 0$ if $i \notin \{i_1, \dots, i_r\}$, we have $Ay = y_{i_1}a^{i_1} + \dots + y_{i_r}a^{i_r}$ and thus $A(x \pm \epsilon y) = b$ for every $\epsilon \in \mathbb{R}$. Since x_i is non-zero whenever y_i is non-zero, we can choose $\epsilon > 0$ small enough such that $x \pm \epsilon y \geq 0$ and thus $x \pm \epsilon y \in X(b)$. Moreover $x = 1/2(x - \epsilon y) + 1/2(x + \epsilon y)$, so x is not an extreme point of $X(b)$. \square

Moreover, when looking for an optimum, we can restrict our attention to the set of basic feasible solutions.

THEOREM 3.2. *If the linear program (3.2) is feasible and bounded, then it has an optimal solution that is a basic feasible solution.*

Proof. Let x be an optimal solution of (3.2). If x has exactly m non-zero entries, then it is a BFS and we are done. So suppose that x has r non-zero entries for $r > m$, and that it is not an extreme point of $X(b)$, i.e., that $x = \delta y + (1 - \delta)z$ for $y, z \in X(b)$ with $y \neq z$ and $\delta \in (0, 1)$. We will show that there must exist an optimal solution with strictly fewer than r non-zero entries; the claim then follows by induction.

Since $c^T x \geq c^T y$ and $c^T x \geq c^T z$ by optimality of x , and since $c^T x = \delta c^T y + (1 - \delta)c^T z$, we must have that $c^T x = c^T y = c^T z$, so y and z are optimal as well. As in the proof of Theorem 3.1, $x_i = 0$ implies that $y_i = z_i = 0$, so y and z have at most r non-zero entries, which must occur in the same rows as in x . If y or z has strictly fewer than r non-zero entries, we are done. Otherwise let $x' = \delta' y + (1 - \delta')z = z + \delta'(y - z)$, and observe that x' is optimal for every $\delta' \in \mathbb{R}$. Moreover, $y - z \neq 0$, and all non-zero entries of $y - z$ occur in rows where x is non-zero as well. We can thus choose $\delta' \in \mathbb{R}$ such that $x' \geq 0$ and such that x' has strictly fewer than r non-zero entries. \square

Since there are only finitely many basic solutions, a naive approach to solving an LP would be to go over all basic solutions and pick one that optimizes the objective. The problem with this approach is that it would not in general be efficient, as the number of basic solutions may grow exponentially in the number of variables. We will now study a well-known method for solving linear programs, the simplex method, which explores the set of basic solutions in a more organized way.

3.3 The Simplex Tableau

One way to understand the simplex method is in terms of the so-called simplex tableau, which stores all the information required to explore the set of basic solutions.

Let $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, and $x \in \mathbb{R}^n$ such that $Ax = b$. Let B be a *basis*, i.e., a set $B \subseteq \{1, \dots, n\}$ with $|B| = m$, corresponding to a choice of m non-zero variables. Then we have

$$A_B x_B + A_N x_N = b,$$

where $A_B \in \mathbb{R}^{m \times m}$ and $A_N \in \mathbb{R}^{m \times (n-m)}$ respectively consist of the columns of A indexed by B and those not indexed by B , and x_B and x_N respectively consist of the rows of x indexed by B and those not indexed by B . Moreover, if x is a basic solution, then there is a basis B such that $x_N = 0$ and $A_B x_B = b$, and if x is a basic feasible solution, there is a basis B such that $x_N = 0$, $A_B x_B = b$, and $x_B \geq 0$.

For every x with $Ax = b$ and every basis B , we have that $x_B = A_B^{-1}(b - A_N x_N)$, and thus

$$\begin{aligned} f(x) &= c^T x = c_B^T x_B + c_N^T x_N \\ &= c_B^T A_B^{-1}(b - A_N x_N) + c_N^T x_N \\ &= c_B^T A_B^{-1} b + (c_N^T - c_B^T A_B^{-1} A_N) x_N \end{aligned}$$

Suppose that we want to maximize $c^T x$ and find that

$$c_N^T - c_B^T A_B^{-1} A_N \leq 0 \quad \text{and} \quad A_B^{-1} b \geq 0. \quad (3.3)$$

Then, for any feasible $x \in \mathbb{R}^n$, it holds that $x_N \geq 0$ and therefore $f(x) \leq c_B^T A_B^{-1} b$. The basic solution x^* with $x_B^* = A_B^{-1} b$ and $x_N^* = 0$, on the other hand, is feasible and satisfies $f(x^*) = c_B^T A_B^{-1} b$. It must therefore be optimal.

If alternatively $(c_N^T - c_B^T A_B^{-1} A_N)_i > 0$ for some i , then we can increase the value of the objective by increasing $(x_N)_i$. Either this can be done indefinitely, which means

that the maximum is unbounded, or the constraints force some of the variables in the basis to become smaller and we have to stop when the first such variable reaches zero. In that case we have found a new BFS and can repeat the process.

Assuming that the LP is feasible and has a bounded optimal solution, there exists a basis B^* for which (3.3) is satisfied. The basic idea behind the simplex method is to start from an initial BFS and then move from basis to basis until B^* is found. The information required for this procedure can conveniently be represented by the so-called *simplex tableau*. For a given basis B , it takes the following form:¹

$$\begin{array}{c}
 \begin{array}{ccc}
 \overbrace{\hspace{1.5cm}}^m & \overbrace{\hspace{2.5cm}}^{n-m} & \overbrace{\hspace{1cm}}^1 \\
 B & N & \\
 \end{array} \\
 \left. \begin{array}{l} m \\ 1 \end{array} \right\} \begin{array}{|ccc|}
 \hline
 A_B^{-1}A_B = I & A_B^{-1}A_N & A_B^{-1}b \\
 \hline
 c_B^T - c_B^T A_B^{-1}A_B = 0 & c_N^T - c_B^T A_B^{-1}A_N & -c_B^T A_B^{-1}b \\
 \hline
 \end{array}
 \end{array}$$

The first m rows consist of the matrix A and the column vector b , multiplied by the inverse of A_B . It is worth pointing out that for any basis B , the LP with constraints $A_B^{-1}Ax = A_B^{-1}b$ is equivalent to the one with constraints $Ax = b$. The first n columns of the last row are equal to $c^T - \lambda^T A$ for $\lambda^T = c_B^T A_B^{-1}$. The vector λ can be interpreted as a solution, not necessarily feasible, to the dual problem. In the last column of the last row we finally have the value $-f(x)$, where x is the BFS with $x_B = A_B^{-1}b$ and $x_N = 0$.

We will see later that the simplex method always maintains feasibility of this solution x . As a consequence it also maintains complementary slackness for x and $\lambda^T = c_B^T A_B^{-1}$: since we work with an LP in standard form, $\lambda^T(Ax - b) = 0$ follows automatically from the feasibility condition, $Ax = b$; the condition $(c^T - \lambda^T A)x = 0$ holds because $x_N = 0$ and $c_B^T - \lambda^T A_B = c_B^T - c_B^T A_B^{-1}A_B = 0$. What it then means for (3.3) to become satisfied is that $c^T - \lambda^T A \leq 0$, i.e., that λ is a feasible solution for the dual. Optimality of x is thus actually a consequence of Theorem 2.4.

3.4 The Simplex Method in Tableau Form

Consider a tableau of the following form, where the basis can be identified by the identity matrix embedded in (a_{ij}) :

(a_{ij})	a_{i0}
a_{0j}	a_{00}

¹The columns of the tableau have been permuted such that those corresponding to the basis appear on the left. This has been done just for convenience: in practice we will always be able to identify the columns corresponding to the basis by the embedded identity matrix.

The simplex method then proceeds as follows:

1. Find an initial BFS with basis B .
2. Check whether $a_{0j} \leq 0$ for every j . If yes, the current solution is optimal, so stop.
3. Choose j such that $a_{0j} > 0$, and choose $i \in \{i' : a_{i'j} > 0\}$ to minimize a_{i0}/a_{ij} . If $a_{ij} \leq 0$ for all i , then the problem is unbounded, so stop. If multiple rows minimize a_{i0}/a_{ij} , the problem has a degenerate BFS.
4. Update the tableau by multiplying row i by $1/a_{ij}$ and adding a $-(a_{kj}/a_{ij})$ multiple of row i to each row $k \neq i$. Then return to Step 2.

We will now describe the different steps of the simplex method in more detail and illustrate them using the LP of Example 2.3.

Finding an initial BFS

Finding an initial BFS is very easy when the constraints are of the form $Ax \leq b$ for $b \geq 0$. We can then write the constraints as $Ax + z = b$ for a vector z of slack variables with regional constraint $z \geq 0$, and obtain a BFS by setting $x = 0$ and $z = b$. This can alternatively be thought of as extending x to (x, z) and setting $(x_B, x_N) = (z, x) = (b, 0)$. We then have $A_B^{-1} = I$ and $c_B = 0$, and the entries in the tableau become A_N and c_N^T for the variables x_1 and x_2 that are not in the basis, and b and 0 in the last column. For the LP of Example 2.3 we obtain the following tableau, where rows and columns have been labeled with the names of the corresponding variables:

	x_1	x_2	z_1	z_2	a_{i0}
z_1	1	2	1	0	6
z_2	1	-1	0	1	3
a_{0j}	1	1	0	0	0

If the constraints do not have this convenient form, finding an initial BFS requires more work. We will discuss this case in the next lecture.

Choosing a pivot column

If $a_{0j} \leq 0$ for all $j \geq 1$, the current solution is optimal. Otherwise we can choose a column j such that $a_{0j} > 0$ as the pivot column and let the corresponding variable enter the basis. If multiple candidate columns exist, choosing any one of them will lead to a new basis, but we could for example break ties toward the column that maximizes a_{0j} or the one with the smallest index. The candidate variables in our example are x_1 and x_2 , so let us choose x_1 . The pivot operation will cause this variable to enter the basis.

Choosing the pivot row

If $a_{ij} \leq 0$ for all i , then the problem is unbounded and the objective can be increased by an arbitrary amount. Otherwise we choose a row $i \in \{i' : a_{i'j} > 0\}$ that minimizes a_{i0}/a_{ij} . This row is called the pivot row, and a_{ij} is called the pivot. If multiple rows

minimize a_{i0}/a_{ij} , the problem has a degenerate BFS. In our example there is a unique choice, corresponding to variable z_2 . The pivot operation will cause this variable to leave the basis.

Pivoting

The purpose of the pivoting step is to get the tableau into the appropriate form for the new BFS. For this, we multiply row i by $1/a_{ij}$ and add a $-(a_{kj}/a_{ij})$ multiple of row i to each row $k \neq i$, including the last one. Our choice of the pivot row as a row that minimizes a_{i0}/a_{ij} turns out to be crucial, as it guarantees that the solution remains feasible after pivoting. In our example, we need to subtract the second row from both the first and the last row, after which the tableau looks as follows:

	x_1	x_2	z_1	z_2	a_{i0}
z_1	0	3	1	-1	3
x_1	1	-1	0	1	3
a_{0j}	0	2	0	-1	-3

Note that the second row now corresponds to variable x_1 , which has replaced z_2 in the basis.

We are now ready to choose a new pivot column. In our example, one further iteration yields the following tableau:

	x_1	x_2	z_1	z_2	a_{i0}
x_2	0	1	$\frac{1}{3}$	$-\frac{1}{3}$	1
x_1	1	0	$\frac{1}{3}$	$\frac{2}{3}$	4
a_{0j}	0	0	$-\frac{2}{3}$	$-\frac{1}{3}$	-5

This corresponds to the BFS where $x_1 = 4$, $x_2 = 1$, and $z_1 = z_2 = 0$, with an objective of -5 . All entries in the last row are non-positive, so this solution is optimal.

3.5 Degeneracies and Cycling

In the absence of degeneracies, the value of the objective function increases in every iteration of the simplex method, and an optimal solution or a certificate for unboundness is found after a finite number of steps. When the simplex method encounters a degenerate BFS, however, it may remain at the same BFS despite changing basis. This would obviously cause the value of the objective function to remain the same as well, and the simplex method may in fact cycle indefinitely through a number of bases that all represent the same BFS.

Such cycling can be avoided by a more careful choice of pivot rows and columns, and thus of the variables entering and leaving the basis. Bland's rule achieves this by fixing some ordering of the variables and then choosing, among all variables that could enter and leave in a given iteration, those that are minimal according to the ordering.