5 Solutions of Linear Programs

In the remaining lectures, we will concentrate on linear programs. We begin by studying the special structure of the feasible set and the objective function in this case, and how it affects the set of optimal solutions.

5.1 Basic Solutions

In the LP of Example 1.1, 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,

$$\text{maximize } c^T x \text{ subject to } Ax \leqslant b,\ x \geqslant 0. \tag{5.1}$$

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 $k \leqslant n$. 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 5.1, and we will give a proof momentarily.

Formally, $x \in S$ is an *extreme point* of a convex set S if it cannot be written as a convex combination of two distinct points in S, i.e., if for all $y, z \in S$ and $\delta \in (0, 1)$, $x = \delta y + (1-\delta)z$ implies that $x = y = z$. Since this geometric characterization of extreme points is hard to work with, we consider an alternative, algebraic characterization. To this end, consider the following LP in standard form, which can be obtained from (5.1) by introducing slack variables:

$$\text{maximize } c^T x \text{ subject to } Ax = b,\ x \geqslant 0, \tag{5.2}$$

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 $x \geqslant 0$ is a *basic feasible solution* (BFS) of (5.2).

We will henceforth make the following assumptions:

(i) the rows of A are linearly independent,

(ii) every set of m columns of A are linearly independent, and

(iii) every basic solution is *non-degenerate*, i.e., has exactly m non-zero variables.

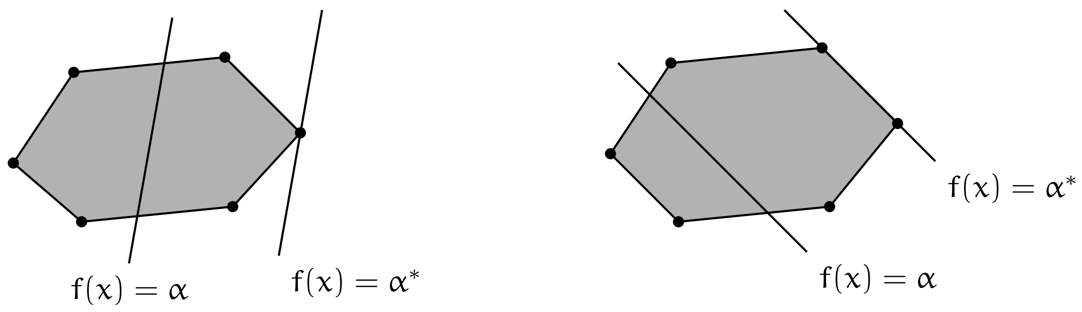

Figure 5.1: Illustration of linear programs with one optimal solution (left) and an infinite number of optimal solutions (right)

Assumptions (i) and (ii) 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. Extra care needs to be taken to handle degeneracies, but this is beyond the scope of this course.

If the above assumptions are satisfied, setting any subset of $n - m$ variables to zero uniquely determines the value of the remaining, basic variables. Computing the set of basic feasible solutions is thus straightforward.

EXAMPLE 5.1. Again consider the LP of Example 1.1. By adding slack variables $x_3 \geqslant 0$ and $x_4 \geqslant 0$, the functional constraint can be written as

$$\begin{pmatrix} 1 & 2 & 1 & 0 \\ 1 & -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 6 \\ 3 \end{pmatrix}.$$

The problem has the following six basic solutions corresponding to the $\binom{4}{2}$ possible ways to choose a basis, which are labeled A through F in Figure 1.1:

	x_1	x_2	x_3	x_4	$f(x)$
A	0	0	6	3	0
B	0	3	0	6	3
C	4	1	0	0	5
D	3	0	3	0	3
E	6	0	0	−3	6
F	0	−3	12	0	−3

5.2 Extreme Points and Optimal Solutions

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

THEOREM 5.2. *A vector is a basic feasible solution of* $Ax = b$ *if and only if it is an extreme point of the set* $X(b) = \{x : Ax = b, x \geqslant 0\}$.

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 $y \geqslant 0$ and $z \geqslant 0$, $x = \delta y + (1-\delta)z$ implies that $y_i = z_i$ whenever $x_i = 0$. By (iii), y and z are basic solutions with the same basis, i.e., both have exactly m non-zero entries, which occur in the same rows. Moreover, $Ay = b = Az$ and thus $A(y-z) = 0$. This yields a linear combination of 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 $\epsilon \in \mathbb{R}$. By choosing $\epsilon > 0$ small enough, $x \pm \epsilon y \geqslant 0$ 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 $X(b)$. □

We are now ready to show that an optimum occurs at an extreme point of the feasible set.

THEOREM 5.3. *If the linear program* (5.2) *is feasible and bounded, then it has an optimal solution that is a basic feasible solution.*

Proof. Let x be an optimal solution of (5.2). 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 $c^Tx \geqslant c^Ty$ and $c^Tx \geqslant c^Tz$ by optimality of x, and since $c^Tx = \delta c^Ty + (1-\delta)c^Tz$, we must have that $c^Tx = c^Ty = c^Tz$, so y and z are optimal as well. As in the proof of Theorem 5.2, $x_i = 0$ 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 $x' \geqslant 0$ and such that x' has strictly fewer than r non-zero entries. □

The result can in fact be extended to show that the maximum of a convex function f over a compact convex set X occurs at an extreme point of X. In this case any point $x \in X$ can be written as a convex combination $x = \sum_{i=1}^{k} \delta_i x^i$ of extreme points $x^1, \dots, x^k \in X$, where $\delta \in \mathbb{R}^k_{\geqslant 0}$ and $\sum_{i=1}^{k} \delta_i = 1$. Convexity of f then implies that

$$f(x) \leqslant \sum_{i=1}^{k} \delta_i f(x^i) \leqslant \max_{1 \leqslant i \leqslant k} f(x^i).$$

5.3 A Naive Approach to Solving Linear Programs

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. By contrast, a large body of work on the theory of computational complexity typically associates efficient computation with methods that for every problem instance can be executed in a number of steps that is at most polynomial in the size of that instance.

In one of the following lectures we will study a well-known method for solving linear programs, the so-called simplex method, which explores the set of basic solutions in a more organized way. It is usually very efficient in practice, but may still require an exponential number of steps for some contrived instances. In fact, no approach is currently know that solves linear programs by inspecting only the boundary of the feasible set and is efficient for every conceivable instance of the problem. There are, however, so-called interior-point method that traverse the interior of the feasible set in search of an optimal solution and are very efficient both in theory and in practice.