Chapter 2

THEORY OF SIMPLEX METHOD

2.1 Mathematical Programming Problems

A *mathematical programming problem* is an optimization problem of finding the values of the unknown variables $x_1, x_2, \cdots, x_n$ that

$$\begin{array}{lll} \text{maximize (or minimize)} & f(x_1, x_2, \cdots, x_n) & \\ \text{subject to} & g_i(x_1, x_2, \cdots, x_n)(\leq, =, \geq)b_i, & i = 1, 2, \cdots, m \end{array} \tag{2.1}$$

where the b_i are real constants and the functions f and g_i are real-valued. The function $f(x_1, x_2, \cdots, x_n)$ is called the *objective function* of the problem (2.1) while the functions $g_i(x_1, x_2, \cdots, x_n)$ are called the *constraints* of (2.1). In vector notations, (2.1) can be written as

$$\begin{array}{lll} \text{max or min} & f(\mathbf{x}^T) & \\ \text{subject to} & g_i(\mathbf{x}^T)(\leq, =, \geq)b_i, & i = 1, 2, \cdots, m \end{array}$$

where $\mathbf{x}^T = (x_1, x_2, \cdots, x_n)$ is the solution vector.

Example 2.1. Consider the following problem.

$$\begin{array}{ll} \text{max} & f(x, y) = xy \\ \text{subject to} & x^2 + y^2 = 1 \end{array}$$

A classical method for solving this problem is the Lagrange multiplier method. Let

$$L(x, y, \lambda) = xy - \lambda(x^2 + y^2 - 1).$$

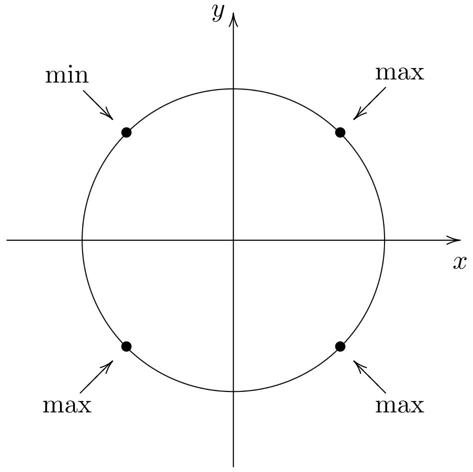
Then differentiate L with respect to x, y, λ and set the partial derivative to 0 we get

$$\frac{\partial L}{\partial x} = y - 2\lambda x = 0,$$
$$\frac{\partial L}{\partial y} = x - 2\lambda y = 0,$$
$$\frac{\partial L}{\partial \lambda} = x^2 + y^2 - 1 = 0.$$

The third equation is redundant here. The first two equations give

$$\frac{y}{2x} = \lambda = \frac{x}{2y}$$

which gives $x^2 = y^2$ or $x = \pm y$. We find that the extrema of xy are obtained at $x = \pm y$. Since $x^2 + y^2 = 1$ we then have $x = \pm\frac{1}{\sqrt{2}}$ and $y = \pm\frac{1}{\sqrt{2}}$. It is then easy the verify that the maximum occurs at $x = y = \frac{1}{\sqrt{2}}$ and $x = y == \frac{1}{\sqrt{2}}$ giving $f(x, y) = \frac{1}{2}$.



A *linear programming problem* (LPP) is a mathematical programming problem having a linear objective function and linear constraints. Thus the general form of an LP problem is

$$\begin{array}{ll} \text{maxormin} & z = c_1x_1 + c_2x_2 + \cdots + c_nx_n \\ \text{subject to} & \begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n (\leq,\ =,\ \geq) b_1, \\ \quad \vdots \qquad\qquad\quad \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n (\leq,\ =,\ \geq) b_m, \end{cases} \end{array} \tag{2.2}$$

Here the constants a_{ij}, b_i and c_j are assumed to be real. The constants c_j are called the *cost* or *price coefficients* of the unknowns x_j and the vector $(c_1, \cdots, c_n)^T$ is called the *cost* or *price vector*.

If in problem (2.2), all the constraints are inequality with sign $\leq$ and the unknowns x_i are restricted to nonnegative values, then the form is called *canonical*. Thus the *canonical form* of an LP problem can be written as

$$\begin{array}{ll} \text{maxormin} & z = c_1x_1 + c_2x_2 + \cdots + c_nx_n \\ \text{subject to} & \begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n \leq b_1, \\ \quad \vdots \qquad\qquad\quad \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \leq b_m, \end{cases} \\ \text{where} & x_i \geq 0, \quad i = 1, 2, \cdots, n. \end{array} \tag{2.3}$$

If all $b_i \geq 0$, then the form is called a *feasible canonical form*.

Before the simplex method can be applied to an LPP, we must first convert it into what is known as the *standard form*:

$$\begin{array}{ll} \max & z = c_1x_1 + \cdots + c_nx_n \\ \text{subject to} & \begin{cases} a_{i1}x_1 + \cdots + a_{in}x_n = b_i, \quad i = 1, 2, \cdots, m \\ \qquad\qquad\qquad\quad x_j \geq 0, \quad j = 1, 2, \cdots, n\ . \end{cases} \end{array} \tag{2.4}$$

Here the b_i are assumed to be nonnegative. We note that the number of variables may or may not be the same as before.

One can always change an LPP problem into the canonical form or into the standard form by the following procedures.

(i) If the LP as originally formulated calls for the minimization of the functional

$$z = c_1x_1 + c_2x_2 + \cdots + c_nx_n,$$

we can instead substitute the equivalent objective function

$$\text{maximize} \quad z' = (-c_1)x_1 + (-c_2)x_2 + \cdots (-c_n)x_n = -z \ .$$

(ii) If any variable x_j is *free* i.e., not restricted to non-negative values, then it can be replaced by

$$x_j = x_j^+ - x_j^-,$$

where $x_j^+ = \max(0, x_j)$ and $x_j^- = \max(0, -x_j)$ are now non-negative. We substitute $x_j^+ - x_j^-$ for x_j in the constraints and objective function in (2.2). The problem then has $(n+1)$ non-negative variables $x_1, \cdots, x_j^+, x_j^-, \cdots, x_n$.

(iii) If $b_i \le 0$, we can multiply the i-th constraint by -1.

(iv) An equality of the form $\sum_{j=1}^n a_{ij}x_j = b_i$ can be replaced by

$$\sum_{j=1}^n a_{ij}x_j \le b_i \quad \text{and} \quad \sum_{j=1}^n (-a_{ij})x_j \le (-b_i).$$

(v) Finally, any inequality constraint in the original formulation can be converted to equations by the addition of non-negative variables called the *slack* and the *surplus* variables. For example, the constraint

$$a_{i1}x_1 + \cdots + a_{ip}x_p \le b_i$$

can be written as

$$a_{i1}x_1 + \cdots + a_{ip}x_p + x_{p+1} = b_i$$

where $x_{p+1} \ge 0$ is a slack variable. Similarly, the constraint

$$a_{j1}x_1 + \cdots + a_{jp}x_p \ge b_j$$

can be written as

$$a_{j1}x_1 + \cdots a_{jp}x_p - x_{p+2} = b_j$$

where $x_{p+2} \ge 0$ is a surplus variable. The new variables would be assigned zero cost coefficients in the objective function, i.e. $c_{p+i} = 0$.

In matrix notations, the standard form of an LPP is

$$\begin{aligned} &\text{Max} && z = \mathbf{c}^T\mathbf{x} \\ &\text{subject to} && A\mathbf{x} = \mathbf{b} && (2.5)\\ &\text{and} && \mathbf{x} \ge \mathbf{0} && (2.6) \end{aligned}$$

where A is $m \times n$, $\mathbf{b}$ is $m \times 1$, $\mathbf{x}$ is $n \times 1$ and rank $(A) = m$.

Definition 2.1. A *feasible solution* (FS) to an LPP is a vector $\mathbf{x}$ which satisfies constraints (2.5) and (2.6). The set of all feasible solutions is called the *feasible region.* A feasible solution to an LPP is said to be an *optimal solution* if it maximizes the objective function of the LPP. A feasible solution to an LPP is said to be a *basic feasible solution* (BFS) if it is a basic solution with respect to the linear system (2.5). If a basic feasible solution is non-degenerate, then we call it a *non-degenerate basic feasible solution.*

We note that the optimal solution may not be unique, but the optimum value of the problem should be unique. For LPP in feasible canonical form, the zero vector is always a feasible solution. Hence the feasible region is always non-empty.

2.2 Basic Feasible Solutions and Extreme Points

In this section, we discuss the relationship between basic feasible solutions to a LPP and extreme points of the corresponding feasible region. We will show that they are indeed the same. We assume a LLP is given in its feasible canonical form.

Theorem 2.1. *The feasible region to an LPP is convex, closed and bounded from below.*

Proof. That the feasible region is convex follows from Lemma 1.1. The closeness follows from the fact that the set that satisfies the equality or inequalities of the type $\leq$ and $\geq$ are closed and that the intersection of closed sets are closed. Finally, by (2.6), we see that the feasible region is a subset of $\mathbb{R}^n_+$, where $\mathbb{R}^n_+$ is given in (1.3) and is a set bounded from below. Hence the feasible region itself must also be bounded from below. $\square$

Theorem 2.2. *If there is a feasible solution then there is a basic feasible solution.*

Proof. Assume that there is a feasible solution $\mathbf{x}$ with p positive variables where $p \leq n$. Let us reorder the variables so that the first p variables are positive. Then the feasible solution can be written as $\mathbf{x}^T = (x_1, x_2, x_3, \cdots, x_p, 0, \cdots, 0)$, and we have

$$\sum_{j=1}^{p} x_j \mathbf{a}_j = \mathbf{b}. \tag{2.7}$$

Let us first consider the case where $\{\mathbf{a}_j\}_{j=1}^p$ is linearly independent. Then we have $p \leq m$ for m is the largest number of linearly independent vectors in A. If $p = m$, $\mathbf{x}$ is basic according to the definition, and in fact it is also non-degenerate. If $p < m$, then there exist $\mathbf{a}_{p+1}, \cdots, \mathbf{a}_m$ such that the set $\{\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_m\}$ is linearly independent. Since $x_{p+1}, x_{p+2}, \cdots, x_m$ are all zero, it follows that

$$\sum_{j=1}^{m} x_j \mathbf{a}_j = \sum_{j=1}^{p} x_j \mathbf{a}_j = \mathbf{b}$$

and $\mathbf{x}$ is a degenerate basic feasible solution.

Suppose that $\{\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_p\}$ is linearly dependent and without loss of generality, assume that none of the $\mathbf{a}_j$ are zero vectors. For if $\mathbf{a}_j$ is zero, we can set x_j to zeros and hence reduce p by 1. With this assumption, then there exists $\{\alpha_j\}_{j=1}^p$ not all zero such that

$$\sum_{j=1}^{p} \alpha_j \mathbf{a}_j = \mathbf{0}.$$

Let $\alpha_r \neq 0$, we have

$$\mathbf{a}_r = \sum_{\substack{j=1 \\ j \neq r}}^{p} \left(-\frac{\alpha_j}{\alpha_r} \right) \mathbf{a}_j.$$

Substitute this into (2.7), we have

$$\sum_{\substack{j=1 \\ j \neq r}}^{p} \left(x_j - x_r \frac{\alpha_j}{\alpha_r} \right) \mathbf{a}_j = \mathbf{b}\ .$$

Hence we have a new solution for (2.5), namely,

$$\left[x_1 - x_r \frac{\alpha_1}{\alpha_r},\ \cdots,\ x_{r-1} - x_r \frac{\alpha_{r-1}}{\alpha_r},\ 0,\ x_{r+1} - x_r \frac{\alpha_{r+1}}{\alpha_r},\ \cdots, x_p - x_r \frac{\alpha_p}{\alpha_r},\ 0, \cdots, 0 \right]^T$$

which has no more than $p-1$ non-zero variables.

We now claim that, by choosing α_r suitably, the new solution above is still a feasible solution. In order that this is true, we must choose our α_r such that

$$x_j - x_r\frac{\alpha_j}{\alpha_r} \geq 0, \qquad j = 1, 2, \cdots, p. \tag{2.8}$$

For those $\alpha_j = 0$, the inequality obviously holds as $x_j \geq 0$ for all $j = 1, 2, \cdots, p$. For those $\alpha_j \neq 0$, the inequality becomes

$$\frac{x_j}{\alpha_j} - \frac{x_r}{\alpha_r} \geq 0, \text{ for } \alpha_j > 0, \tag{2.9}$$

$$\frac{x_j}{\alpha_j} - \frac{x_r}{\alpha_r} \leq 0, \text{ for } \alpha_j < 0. \tag{2.10}$$

If we choose our $\alpha_r > 0$, then (2.10) automatically holds. Moreover if α_r is chosen as

$$\frac{x_r}{\alpha_r} = \min_j \left\{\frac{x_j}{\alpha_j} : \alpha_j > 0\right\}, \tag{2.11}$$

then (2.9) is also satisfied. Thus by choosing α_r as in (2.11), then (2.8) holds and $\mathbf{x}$ is a feasible solution of $p-1$ non-zero variables. We note that we can also choose $\alpha_r < 0$ and such that

$$\frac{x_r}{\alpha_r} = \max_j \left\{\frac{x_j}{\alpha_j} :\ \alpha_j < 0\right\}, \tag{2.12}$$

then (2.8) is also satisfied, though clearly the two $\alpha_r's$ in general will not give the same new solution with $(p-1)$ non-zero entries.

Assume that a suitable α_r has been chosen and the new feasible solution is given by

$$\hat{\mathbf{x}} = [\hat{x}_1, \hat{x}_2, \cdots, \hat{x}_{r-1}, 0, \hat{x}_{r+1}, \cdots, \hat{x}_p, 0, 0, \cdots, 0]^T$$

which has no more than $p-1$ non-zero variables. We can now check whether the corresponding column vectors of A are linearly independent or not. If it is, then we have a basic feasible solution. If it is not, we can repeat the above process to reduce the number of non-zero variables to $p-2$. Since p is finite, the process must stop after at most $p-1$ operations, at which we only have one nonzero variable. The corresponding column of A is clearly linearly independent. (If that column is a zero column, then the corresponding α can be arbitrary chosen, and hence it can always be eliminated first). The $\mathbf{x}$ obtained then is a basic feasible solution. $\square$

Example 2.2. Consider the linear system

$$\begin{bmatrix} 2 & 1 & 4 \\ 3 & 1 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 11 \\ 14 \end{bmatrix}$$

where $\mathbf{x} = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}$ is a feasible solution. Since $\mathbf{a}_1 + 2\mathbf{a}_2 - \mathbf{a}_3 = \mathbf{0}$, we have

$$\alpha_1 = 1\ , \quad \alpha_2 = 2\ , \quad \alpha_3 = -1.$$

By (2.11),

$$\frac{x_2}{\alpha_2} = \frac{3}{2} = \min_{j=1,2,3} \left\{\frac{x_j}{\alpha_j} :\ \alpha_j > 0\right\}.$$

Hence $r = 2$ and

$$\begin{aligned}
\hat{x}_1 &= x_1 - x_2\frac{\alpha_1}{\alpha_2} = 2 - 3\cdot\frac{1}{2} = \frac{1}{2},\\
\hat{x}_2 &= 0,\\
\hat{x}_3 &= x_3 - x_2\frac{\alpha_3}{\alpha_2} = 1 - 3\cdot\frac{-1}{2} = \frac{5}{2},
\end{aligned}$$

Thus the new feasible solution is $\mathbf{x}_1{}^T = (\frac{1}{2}, 0, \frac{5}{2})$. Since

$$\mathbf{a}_1 = \begin{bmatrix}2\\3\end{bmatrix}, \quad \mathbf{a}_3 = \begin{bmatrix}4\\5\end{bmatrix}$$

are linearly independent, the solution is also basic.

Similarly, if we use (2.12), we get

$$\frac{x_3}{\alpha_3} = \frac{1}{-1} = \max_{j=1,2,3}\left\{\frac{x_j}{\alpha_j} : \alpha_j < 0\right\}.$$

Hence $r = 3$ and we have

$$\begin{aligned}
\hat{x}_1 &= x_1 - x_3\frac{\alpha_1}{\alpha_3} = 2 - (-1)1 = 3,\\
\hat{x}_2 &= x_2 - x_3\frac{\alpha_2}{\alpha_3} = 3 - (-1)2 = 5.
\end{aligned}$$

Thus $\mathbf{x}_2{}^T = (3, 5, 0)$, which is also a basic feasible solution.

Suppose that we eliminate $\mathbf{a}_1$ instead, then we have

$$\begin{aligned}
\hat{x}_1 &= 0\\
\hat{x}_2 &= x_1\frac{x_2}{\alpha_1} = 3 - 2\cdot\frac{2}{1} = -1\\
\hat{x}_3 &= x_3 - x_1\frac{\alpha_3}{\alpha_1} = 1 - 2\cdot\frac{(-1)}{1} = 3
\end{aligned}$$

Thus the new solution is $\mathbf{x}_1{}^T = (0, -1, 3)$. We note that it is a basic solution but is no longer feasible.

Theorem 2.3. *The basic feasible solutions of an LPP are extreme points of the corresponding feasible region.*

Proof. Suppose that $\mathbf{x}$ is a basic feasible solution and without loss of generality, assume that $\mathbf{x}$ has the form

$$\mathbf{x} = \begin{bmatrix}\mathbf{x}_B\\ \mathbf{0}\end{bmatrix}$$

where

$$\mathbf{x}_B = B^{-1}\mathbf{b}$$

is an $m \times 1$ vector. Suppose on the contrary that there exist two feasible solutions $\mathbf{x}_1, \mathbf{x}_2$, different from $\mathbf{x}$, such that

$$\mathbf{x} = \lambda\mathbf{x}_1 + (1-\lambda)\mathbf{x}_2$$

for some $\lambda \in (0, 1)$. We write

$$\mathbf{x}_1 = \begin{bmatrix}\mathbf{u}_1\\ \mathbf{v}_1\end{bmatrix}, \quad \mathbf{x}_2 = \begin{bmatrix}\mathbf{u}_2\\ \mathbf{v}_2\end{bmatrix}$$

where $\mathbf{u}_1$, $\mathbf{u}_2$ are m-vectors and $\mathbf{v}_1$, $\mathbf{v}_2$ are $(n-m)$-vectors. Then we have

$$\mathbf{0} = \lambda \mathbf{v}_1 + (1-\lambda)\mathbf{v}_2.$$

As $\mathbf{x}_1, \mathbf{x}_2$ are feasible, $\mathbf{v}_1, \mathbf{v}_2 \geq \mathbf{0}$. Since $\lambda, (1-\lambda) > 0$, we have $\mathbf{v}_1 = \mathbf{0}$ and $\mathbf{v}_2 = \mathbf{0}$. Thus

$$\mathbf{b} = A\mathbf{x}_1 = [B, R]\begin{bmatrix}\mathbf{u}_1 \\ \mathbf{v}_1\end{bmatrix} = B\mathbf{u}_1 \ ,$$

and similarly,

$$\mathbf{b} = A\mathbf{x}_2 = B\mathbf{u}_2 \ .$$

Thus we have

$$B\mathbf{u}_1 = B\mathbf{u}_2 = \mathbf{b} = B\mathbf{x}_B.$$

Since B is non-singular, this implies that $\mathbf{u}_1 = \mathbf{u}_2 = \mathbf{x}_B$. Hence

$$\mathbf{x} = \begin{bmatrix}\mathbf{x}_B \\ \mathbf{0}\end{bmatrix} = \begin{bmatrix}\mathbf{u}_1 \\ \mathbf{0}\end{bmatrix} = \mathbf{x}_1 = \begin{bmatrix}\mathbf{u}_2 \\ \mathbf{0}\end{bmatrix} = \mathbf{x}_2,$$

and that is a contradiction. Hence $\mathbf{x}$ must be an extreme point of the feasible region. □

The next theorem is the converse of Theorem 2.3.

Theorem 2.4. *The extreme points of a feasible region are basic feasible solutions of the corresponding LPP.*

Proof. Suppose $\mathbf{x}_0 = (x_1, \cdots, x_n)^T$ is an extreme point of the feasible region. Assume that there are r components of $\mathbf{x}_0$ which are non-zero. Without loss of generality, let $x_i > 0$ for $i = 1, 2, \cdots, r$ and $x_i = 0$ for $i = r+1, r+2, \cdots, n$. Then we have

$$\sum_{i=1}^{r} x_i \mathbf{a}_i = \mathbf{b}.$$

We claim that $\{\mathbf{a}_1, \cdots, \mathbf{a}_r\}$ is linearly independent. Suppose on contrary that there exist $\alpha_i, i = 1, 2, \cdots, r$, not all zero, such that

$$\sum_{i=1}^{r} \alpha_i \mathbf{a}_i = \mathbf{0}. \tag{2.13}$$

Let ϵ be such that

$$0 < \epsilon < \min_{\alpha_i \neq 0} \frac{x_i}{|\alpha_i|},$$

then

$$x_i \pm \epsilon \cdot \alpha_i > 0, \quad \forall i = 1, 2, \cdots, r. \tag{2.14}$$

Put $\mathbf{x}_1 = \mathbf{x}_0 + \epsilon \cdot \boldsymbol{\alpha}$ and $\mathbf{x}_2 = \mathbf{x}_0 - \epsilon \cdot \boldsymbol{\alpha}$ where $\boldsymbol{\alpha}^T = (\alpha_1, \alpha_2, \cdots, \alpha_r, 0, \cdots, 0)$. We claim that $\mathbf{x}_1, \mathbf{x}_2$ are feasible solutions. Clearly by (2.14), $\mathbf{x}_1 \geq \mathbf{0}$ and $\mathbf{x}_2 \geq \mathbf{0}$. Moreover by (2.13),

$$A\mathbf{x}_1 = A\mathbf{x}_0 + \epsilon A\boldsymbol{\alpha} = A\mathbf{x}_0 + \mathbf{0} = \mathbf{b}.$$

Therefore, $\mathbf{x}_1$ is a feasible solution. Similarly,

$$A\mathbf{x}_2 = A\mathbf{x}_0 - \epsilon A\boldsymbol{\alpha} = A\mathbf{x}_0 + \mathbf{0} = \mathbf{b},$$

and $\mathbf{x}_2$ is also a feasible solution. Since clearly $\mathbf{x}_0 = \dfrac{1}{2}\mathbf{x}_1 + \dfrac{1}{2}\mathbf{x}_2$, $\mathbf{x}_0$ is not an extreme point, hence we have a contradiction. Therefore the set $\{\mathbf{a}_1, \cdots, \mathbf{a}_r\}$ must be linearly independent, and that $\mathbf{x}_0$ is a basic feasible solution. □

The significance of the extreme points of feasible regions is given by the following theorem.

Theorem 2.5. *The optimal solution to an LPP occurs at an extreme point of the feasible region.*

Proof. We first claim that no point on the hyperplane that corresponds to the optimal value of the objective function can be an interior point of the feasible region. Suppose on contrary that $z = \mathbf{c}^T\mathbf{x}$ is an optimal hyperplane and that the optimal value is attained by a point $\mathbf{x}_0$ in the interior of the feasible region. Then there exists an $\epsilon > 0$ such that the sphere

$$X = \{\mathbf{x} : |\mathbf{x} - \mathbf{x}_0| < \epsilon\}$$

is in the feasible region. Then the point

$$\mathbf{x}_1 = \mathbf{x}_0 + \frac{\epsilon}{2}\frac{\mathbf{c}}{|\mathbf{c}|} \in X$$

is a feasible solution. But

$$\mathbf{c}^T\mathbf{x}_1 = \mathbf{c}^T\mathbf{x}_0 + \mathbf{c}^T\frac{\epsilon}{2}\frac{\mathbf{c}}{|\mathbf{c}|} = z + \frac{\epsilon}{2}|\mathbf{c}| > z,$$

a contradiction to the optimality of z. Thus $\mathbf{x}_0$ has to be a boundary point.

Now since $\mathbf{c}^T\mathbf{x} \le z$ for all feasible solutions $\mathbf{x}$, we see that the optimal hyperplane is a supporting hyperplane of the feasible region at the point $\mathbf{x}_0$. By Theorem 1, the feasible region is bounded from below. Therefore by Theorem 1.5, the supporting hyperplane $z = \mathbf{c}^T\mathbf{x}$ contains at least one extreme point of the feasible region. Clearly that extreme point must also be an optimal solution to the LPP. □

Summarizing what we have proved so far, we see that the optimal value of an LPP can be obtained at the basic feasible solutions, or equivalently, at the extreme points. Simplex method is a method that systematically searches through the basic feasible solutions for the optimal one.

Example 2.3. Consider the constraint set in $\mathbb{R}^2$ defined by

$$x_1 + \frac{8}{3}x_2 \le 4 \tag{2.15}$$

$$x_1 + \quad x_2 \le 2 \tag{2.16}$$

$$2x_1 \qquad \le 3 \tag{2.17}$$

$$x_1, x_2 \ge 0 \tag{2.18}$$

Adding slack variables x_3, x_4 and x_5 to convert it into standard form gives,

$$x_1 + \frac{8}{3}x_2 + x_3 \qquad = 4 \tag{2.19}$$

$$x_1 + \quad x_2 \quad + x_4 \quad = 2 \tag{2.20}$$

$$2x_1 \qquad\qquad + x_5 = 3 \tag{2.21}$$

$$x_1, x_2, x_3, x_4, x_5 \ge 0 \tag{2.22}$$

A basic solution is obtained by setting any two variables of x_1, x_2, x_3, x_4, x_5 to zero and solving for the remaining three. For example, let us set $x_1 = 0$ and $x_3 = 0$ and solve for x_2, x_4 and x_5 in

$$\begin{cases} \frac{8}{3}x_2 & & = 4 \\ x_2 + x_4 & & = 2 \\ & x_5 & = 3 \end{cases}$$

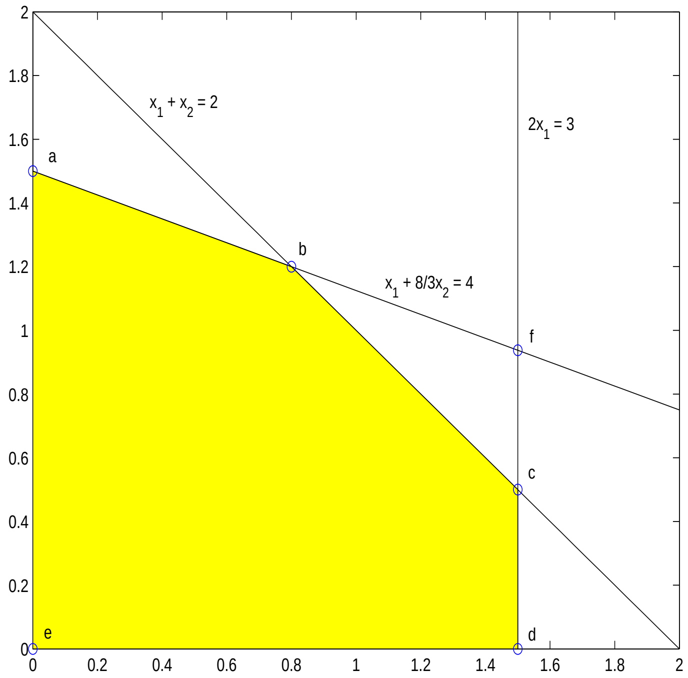


Figure 2.1. *There are 5 extreme points by inspection* $\{a, b, c, d, e\}$.

We get the point $[0, \frac{3}{2}, 0, \frac{1}{2}, 3]$. From Figure 2.3, we see that this point corresponds to the extreme point a of the convex polyhedron K defined by (2.19), (2.20), (2.21), (2.22). The equations $x_1 = 0$ and $x_3 = 0$ are called the *binding equations* of the extreme point a.

We note that not all basic solutions are feasible. In fact there is a maximum total of $\binom{5}{3} = \binom{5}{2} = 10$ basic solutions and here we have only five extreme points. The extreme points of the region K and their corresponding binding equations are given in the following table.

extreme point	a	b	c	d	e
set to zero	x_1	x_3	x_4	x_2	x_1
	x_3	x_4	x_5	x_5	x_2

If we set x_3 and x_5 to zero, the basic solution we get is $[\frac{3}{2}, \frac{15}{16}, 0, -\frac{7}{16}, 0]$. Hence it is not a basic *feasible* solution and therefore does not correspond to any one of the extreme point in K. In fact, it is given by the point f in the above figure.

2.3 Improving a Basic Feasible Solution

Let the LPP be

$$\max \qquad z = \mathbf{c}^T\mathbf{x}$$

$$\text{subject to} \begin{cases} A\mathbf{x} = \mathbf{b} \\ \mathbf{x} \geq \mathbf{0}\ . \end{cases}$$

Here we assume that $\mathbf{b} \geq \mathbf{0}$ and $\text{rank}(A) = m$. Let the columns of A be given by $\mathbf{a}_j$, i.e. $A = [\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_n]$. Let B be an $m \times m$ non-singular matrix whose columns are linearly independent columns of A and we denote B by $[\mathbf{b}_1, \mathbf{b}_2, \cdots, \mathbf{b}_m]$. We learn from §1.5 that for any choice of basic matrix B, there corresponds a basic solution to $A\mathbf{x} = \mathbf{b}$. The basic solution is given by the m-vector $\mathbf{x}_B = [x_{B_1}, \cdots, x_{B_m}]$ where

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

Corresponding to any such $\mathbf{x}_B$, we define an m-vector $\mathbf{c}_B$, called the *reduced cost vector*, containing the prices of the basic variables, i.e.

$$\mathbf{c}_B = \begin{bmatrix} c_{B_1} \\ \vdots \\ c_{B_m} \end{bmatrix}.$$

Note that for any BFS $\mathbf{x}_B$, the value of the objective function is given by

$$z = \mathbf{c}_B^T \mathbf{x}_B.$$

To improve z, we must be able to locate other BFS easily, or equivalently, we must be able to replace a basic matrix B by others easily. In order to do this, we need to express the columns of A in terms of the columns of B. Since A is an $m \times n$ matrix and $\text{rank}(A) = m$, the column space of A is m dimensional. Thus the columns of B form a basis for the column space of A. Let

$$\mathbf{a}_j = \sum_{i=1}^{m} y_{ij}\mathbf{b}_i, \quad \text{for all } j = 1, 2, \cdots, n. \tag{2.23}$$

Put

$$\mathbf{y}_j = \begin{bmatrix} y_{1j} \\ y_{2j} \\ \vdots \\ y_{mj} \end{bmatrix}, \quad \text{for all } j = 1, 2, \cdots, n,$$

then we have $\mathbf{a}_j = B\mathbf{y}_j$. Hence

$$\mathbf{y}_j = B^{-1}\mathbf{a}_j, \quad \text{for all } j = 1, 2, \cdots, n. \tag{2.24}$$

The matrix B can be considered as the *change of coordinate* matrix from $\{\mathbf{y}_1, \cdots, \mathbf{y}_n\}$ to $\{\mathbf{a}_1, \cdots, \mathbf{a}_n\}$. We remark that if $\mathbf{a}_j$ appears as $\mathbf{b}_i$, i.e., x_j is a basic variable and $x_j = x_{B_i}$, then $\mathbf{y}_j = \mathbf{e}_i$, the i-th unit vector.

Given a BFS $\mathbf{x}_B = B^{-1}\mathbf{b}$ corresponding to the basic matrix B, we would like to improve the value of the objective function which is currently given by $z = \mathbf{c}_B^T \mathbf{x}_B$. For simplicity, we will confine ourselves to those BFS in which only one column of B is changed. We claim that if $y_{rj} \neq 0$ for some r, then $\mathbf{b}_r$ can be replaced by $\mathbf{a}_j$ and the new set of vectors still form a basis. We can then form the corresponding new basic solution. In fact, we note that if $y_{rj} \neq 0$, then by (2.23)

$$\mathbf{b}_r = \frac{1}{y_{rj}}\mathbf{a}_j - \sum_{\substack{i=1 \\ i \neq r}}^{m} \frac{y_{ij}}{y_{rj}}\mathbf{b}_i.$$

Using this, we can replace $\mathbf{b}_r$ in

$$B\mathbf{x}_B = \sum_{i=1}^{m} x_{B_i}\mathbf{b}_i = \mathbf{b},$$

by $\mathbf{a}_j$ and we get

$$\sum_{\substack{i=1 \\ i \neq r}}^{m} \left(x_{B_i} - x_{B_r}\frac{y_{ij}}{y_{rj}} \right)\mathbf{b}_i + \frac{x_{B_r}}{y_{rj}}\mathbf{a}_j = \mathbf{b}.$$

Let

$$\hat{x}_{B_i} = \begin{cases} x_{B_i} - x_{B_r}\dfrac{y_{ij}}{y_{rj}}, & i = 1, 2, \cdots, m \\ \dfrac{x_{B_r}}{y_{rj}}, & i = j, \end{cases}$$

we see that the vector

$$\hat{\mathbf{x}}_B = [\hat{x}_{B_1}, \cdots, \hat{x}_{B_{r-1}}, 0, \hat{x}_{B_{r+1}}, \cdots, \hat{x}_{B_m}, 0, \cdots, 0, \hat{x}_{B_j}, 0, \cdots, 0]^T$$

is a basic solution with $\hat{x}_{B_r} = 0$. Thus the old basic variable x_r is replaced by the new basic variable x_j.

Now we have to make sure that the new basic solution is a basic *feasible* solution with a *larger* objective value. We will see that we can assure the feasibility of the new solution by choosing a suitable $\mathbf{b}_r$ to be replaced, and we can improve the objective value by choosing a suitable $\mathbf{a}_j$ to be inserted in B.

2.3.1 Feasibility: restriction on $\mathbf{b}_r$.

We require that $\hat{x}_{B_i} \geq 0$ for all $i = 1, \cdots, n$. Since $x_{B_i} = 0$ for all $m < i \leq n$, we only need

$$\begin{cases} x_{B_i} - x_{B_r}\dfrac{y_{ij}}{y_{rj}} \geq 0 & i = 1, 2, \cdots, m \\ \qquad\qquad \dfrac{x_{B_r}}{y_{rj}} \geq 0 \end{cases}$$

Let r be chosen such that

$$\frac{x_{B_r}}{y_{rj}} = \min_{i=1,2,\cdots,m} \left\{ \frac{x_{B_i}}{y_{ij}} : y_{ij} > 0 \right\}. \tag{2.25}$$

Then it is easy to check that all $\hat{x}_{B_i} \geq 0$ for all i. Thus the corresponding column $\mathbf{b}_r$ is the column to be replaced from B. We call $\mathbf{b}_r$ the *leaving* column.

2.3.2 Optimality: restriction on $\mathbf{a}_j$.

Originally, the objective value is given by

$$z = \mathbf{c}_B^T \mathbf{x}_B = \sum_{i=1}^{m} c_{B_i} x_{B_i}.$$

After the removal and insertion of columns of B, the new objective value becomes

$$\hat{z} = \hat{\mathbf{c}}_B^T \hat{\mathbf{x}}_B = \sum_{i=1}^{m} \hat{c}_{B_i} \hat{x}_{B_i}$$

where $\hat{c}_{B_i} = c_{B_i}$ for all $i = 1, 2, \cdots, m$ and $i \neq r$, and $\hat{c}_{B_r} = c_j$. Thus we have

$$\begin{aligned}
\hat{z} &= \sum_{\substack{i=1 \\ i\neq r}}^{m} c_{B_i}\hat{x}_{B_i} + \hat{c}_{B_r}\hat{x}_{B_r} \\
&= \sum_{\substack{i=1 \\ i\neq r}}^{m} c_{B_i}\left(x_{B_i} - x_{B_r}\frac{y_{ij}}{y_{rj}}\right) + c_j\frac{x_{B_r}}{y_{rj}} \\
&= \sum_{i=1}^{m} c_{B_i}x_{B_i} - \frac{x_{B_r}}{y_{rj}}\sum_{i=1}^{m} c_{B_i}y_{ij} + c_j\frac{x_{B_r}}{y_{rj}} \\
&= z - \frac{x_{B_r}}{y_{rj}}\mathbf{c}_B^T\mathbf{y}_j + c_j\frac{x_{B_r}}{y_{rj}} \\
&= z + \frac{x_{B_r}}{y_{rj}}(c_j - z_j)
\end{aligned}$$

where

$$z_j = \mathbf{c}_B^T\mathbf{y}_j = \mathbf{c}_B^T B^{-1}\mathbf{a}_j. \tag{2.26}$$

Obviously, by our choice, if $c_j > z_j$, then $\hat{z} \geq z$. Hence we choose our $\mathbf{a}_j$ such that $c_j - z_j > 0$ and $y_{ij} > 0$ for some i. For if $y_{ij} \leq 0$ for all i, then the new solution will not be feasible. The $\mathbf{a}_j$ that we have chosen is called the *entering* column. We note that if $\mathbf{x}_B$ is non-degenerate, then $x_{B_r} > 0$, and hence $\hat{z} > z$.

Let us summarize the process of improving a basic feasible solution in the following theorem.

Theorem 2.6. *Let $\mathbf{x}_B$ be a BFS to an LPP with corresponding basic matrix B and objective value z. If (i) there exists a column $\mathbf{a}_j$ in A but not in B such that the condition $c_j - z_j > 0$ holds and if (ii) at least one $y_{ij} > 0$, then it is possible to obtain a new BFS by replacing one column in B by $\mathbf{a}_j$ and the new value of the objective function $\hat{z}$ is larger than or equal to z. Furthermore, if $\mathbf{x}_B$ is non-degenerate, then we have $\hat{z} > z$.*

The numbers $c_j - z_j$ are called the *reduced cost coefficients* with respect to the matrix B. We remark that if $\mathbf{a}_j$ already appears as $\mathbf{b}_i$, i.e. x_j is a basic variable and $x_j = x_{B_i}$, then $c_j - z_j = 0$. In fact

$$z_j = \mathbf{c}_B^T B^{-1}\mathbf{a}_j = \mathbf{c}_B^T\mathbf{y}_j = \mathbf{c}_B^T\mathbf{e}_i = c_{B_i} = c_j.$$

By Theorem 2.6, it is natural to think that when all $z_j - c_j \geq 0$, we have reached the optimal solution.

Theorem 2.7. *Let $\mathbf{x}_B$ be a BFS to an LPP with corresponding objective value $z_0 = \mathbf{c}_B^T\mathbf{x}_B$. If $z_j - c_j \geq 0$ for every column $\mathbf{a}_j$ in A, then $\mathbf{x}_B$ is optimal.*

Proof. We first note that given any feasible solution $\mathbf{x}$, then by the assumption that $z_j - c_j \geq 0$ for all $j = 1, 2, \cdots, n$ and by (2.26), we have

$$z = \sum_{j=1}^{n} c_j x_j \leq \sum_{j=1}^{n} z_j x_j = \sum_{j=1}^{n} (\mathbf{c}_B^T\mathbf{y}_j)x_j = \sum_{j=1}^{n}\left(\sum_{i=1}^{m} c_{B_i}y_{ij}\right)x_j = \sum_{i=1}^{m}\left(\sum_{j=1}^{n} y_{ij}x_j\right)c_{B_i}.$$

Thus

$$z \leq \sum_{i=1}^{m}\left(\sum_{j=1}^{n} y_{ij}x_j\right)c_{B_i}. \tag{2.27}$$

Now we claim that

$$\tilde{x}_i \equiv \sum_{j=1}^{n} y_{ij}x_j = x_{B_i}, \quad i = 1, 2, \cdots, m.$$

Since $\mathbf{x}$ is a feasible solution, $A\mathbf{x} = \mathbf{b}$. Thus by (2.24),

$$\mathbf{b} = \sum_{j=1}^{n} x_j \mathbf{a}_j = \sum_{j=1}^{n} x_j (B\mathbf{y}_j) = \sum_{j=1}^{n} \left(\sum_{i=1}^{m} y_{ij} \mathbf{b}_i \right) x_j = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} y_{ij} x_j \right) \mathbf{b}_i = \sum_{i=1}^{m} \tilde{x}_i \mathbf{b}_i = B\tilde{\mathbf{x}}.$$

Since B is non-singular and we already have $B\mathbf{x}_B = \mathbf{b}$, it follows that $\tilde{\mathbf{x}} = \mathbf{x}_B$. Thus by (2.27),

$$z \le \sum_{i=1}^{m} x_{B_i} c_{B_i} = z_0$$

for all $\mathbf{x}$ in the feasible region. □

Example 2.4. Let us consider the LPP with

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{c}^T = [2, 5, 6, 8] \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} 5 \\ 2 \end{bmatrix}.$$

Let us choose our starting B as

$$B = [\mathbf{a}_1, \mathbf{a}_4] = [\mathbf{b}_1, \mathbf{b}_2] = \begin{bmatrix} 1 & 4 \\ 1 & 1 \end{bmatrix}.$$

Then it is easily checked that the corresponding basic solution is $\mathbf{x}_B = [1, 1]^T$, which is clearly feasible with objective value

$$z = \mathbf{c}_B^T \mathbf{x}_B = [2, 8] \begin{bmatrix} 1 \\ 1 \end{bmatrix} = 10.$$

Since

$$B^{-1} = \frac{1}{3} \begin{bmatrix} -1 & 4 \\ 1 & -1 \end{bmatrix},$$

by (2.24), the y_{ij} are given by

$$\mathbf{y}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \mathbf{y}_2 = \begin{bmatrix} -\frac{2}{3} \\ \frac{2}{3} \end{bmatrix} \quad \mathbf{y}_3 = \begin{bmatrix} -1 \\ 1 \end{bmatrix} \quad \mathbf{y}_4 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Hence

$$z_2 = [2, 8] \begin{bmatrix} -\frac{2}{3} \\ \frac{2}{3} \end{bmatrix} = 4,$$

and

$$z_3 = [2, 8] \begin{bmatrix} -1 \\ 1 \end{bmatrix} = 6.$$

Since $z_2 - c_2 = -1$ and $z_3 - c_3 = 0$, we see that $\mathbf{a}_2$ is the *entering* column. As remarked above,

$$z_1 - c_1 = z_4 - c_4 = 0,$$

because x_1 and x_4 are basic variables. Looking at the column entries of $\mathbf{y}_2$, we find that y_{22} is the only positive entry. Hence $\mathbf{b}_2 = \mathbf{a}_4$ is the *leaving* column. Thus

$$\hat{B} = [\mathbf{a}_1, \mathbf{a}_2] = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix},$$

and the corresponding basic solution is found to be $\hat{\mathbf{x}}_B = [2, \frac{3}{2}]^T$, which is clearly feasible as expected. The new objective value is given by

$$\hat{z} = [2, 5] \begin{bmatrix} 2 \\ \frac{3}{2} \end{bmatrix} = 11.5 > z.$$

Since

$$\hat{B}^{-1} = \frac{1}{2}\begin{bmatrix} 0 & 2 \\ 1 & -1 \end{bmatrix},$$

by (2.24), the $\hat{y}_{ij}$ are given by

$$\hat{\mathbf{y}}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \hat{\mathbf{y}}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \quad \hat{\mathbf{y}}_3 = \begin{bmatrix} 0 \\ \frac{3}{2} \end{bmatrix} \quad \hat{\mathbf{y}}_4 = \begin{bmatrix} 1 \\ \frac{3}{2} \end{bmatrix}.$$

Hence

$$z_3 - c_3 = [2, 5]\begin{bmatrix} 0 \\ \frac{3}{2} \end{bmatrix} - 6 = 1.5,$$

and

$$z_4 - c_4 = [2, 5]\begin{bmatrix} 1 \\ \frac{3}{2} \end{bmatrix} - 8 = 1.5.$$

Since all $z_j - c_j \geq 0$, $1 \leq j \leq 4$, we see that the point $[2, \frac{3}{2}]$ is an optimal solution.

The last example illustrates how one can find the optimal solution by searching through the basic feasible solutions. That is exactly what *simplex method* does. However, the simplex method uses tableaus to minimize the book-keeping work that we encountered in the last example.