OPTIMIZATION 1

CHAPTER 1

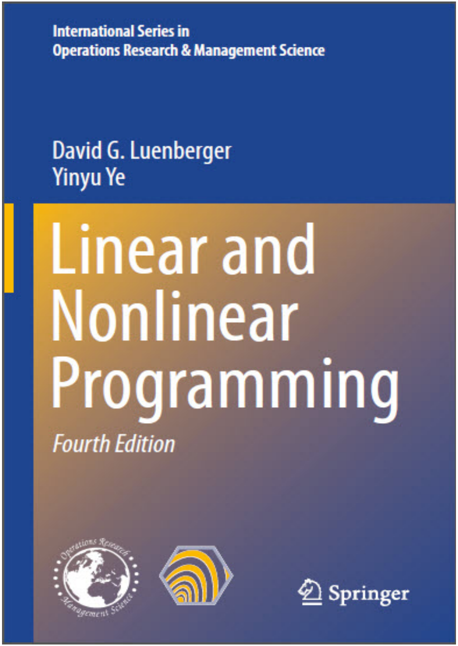
BASIC PROPERTIES OF LINEAR PROGRAMS

OPTIMIZATION I

References

Textbooks:

1. D. G. Luenberger, Y. Ye, *Linear and Nonlinear Programming*, 4th ed. Springer, 2016
2. I. Griva, S. G. Nash, A. Sofer, *Linear and Nonlinear Optimization*, 2nd ed. SIAM, 2009
3. G. Cornuejols, R. Tutuncu, *Optimization Methods in Finance*, Cambridge University Press, 2007


International Series in
Operations Research & Management Science
David G. Luenberger
Yinyu Ye
Linear and Nonlinear Programming
Fourth Edition
Springer

In daily life it is constantly necessary to choose the best possible (optimal) solution. A tremendous number of such problems arise in economics and in technology.

In such cases it is frequently useful to resort to mathematics.

Chapter 1 BASIC PROPERTIES OF LINEAR PROGRAMS

"Nothing takes place in the world whose meaning is not that of some maximum or minimum."

L. Euler

There are four stages that a real-world problem passes through from organization to conclusion.

- recognition of the problem;
- formulation of a mathematical model;
- solution of the mathematical problem; and
- translation of the results back into the context of the original problem.

- **Matrix Notation**

An ***m* × *n* matrix** is a rectangular array of m rows and n columns of the form

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

where each a_{ij} is a number called an **entry** or **element** of the matrix. The numbers m and n are called the **dimension** or **size** of the matrix.

The matrix itself is denoted by a capital boldface letter. When specific numbers are not used, the elements are denoted by lower-case letters, having a double subscript.

In a matrix $\mathbf{A}$, a_{ij} denotes the entry that occurs in row i and column j of $\mathbf{A}$, and we use the notation $\mathbf{A} = [a_{ij}]$.

An $m \times n$ matrix whose entries are all zeros is called the $m \times n$ **zero matrix** and is denoted by $\mathbf{O}$.

A matrix with n rows and n columns is called a **square matrix of order n**, and the entries $a_{11}, a_{22}, \ldots, a_{nn}$ form the **main diagonal** of $\mathbf{A}$.

An **identity matrix I** is an $n \times n$ matrix that has ones on the main diagonal and zeros elsewhere:

$$\mathbf{I} = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & & & & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix}.$$

If $\mathbf{A} = [a_{ij}]$ is an $m \times n$ matrix, then the **transpose** of $\mathbf{A}$, denoted $\mathbf{A}^T$, is the $n \times m$ matrix $\mathbf{A}^T = [b_{ij}]$, where $b_{ij} = a_{ji}$ for all i and j, $1 \leq j \leq m$, and $1 \leq i \leq n$.

A square matrix $\mathbf{A} = [a_{ij}]$ is **symmetric** if $a_{ij} = a_{ji}$ for all $i, j = 1, \ldots, n$, that is, $\mathbf{A}^T = \mathbf{A}$.

A square matrix $\mathbf{A}$ is **nonsingular** if there is a matrix $\mathbf{A}^{-1}$, called the **inverse** of $\mathbf{A}$, such that $\mathbf{A}^{-1}\mathbf{A} = \mathbf{I} = \mathbf{A}\mathbf{A}^{-1}$.

The **determinant** of a square matrix $\mathbf{A}$ is denoted by $\det(\mathbf{A})$.

The determinant is nonzero if and only if the matrix is nonsingular.

$$\exists\, \mathbf{A}^{-1} \iff \det(\mathbf{A}) \neq 0.$$

- Matrices with a single row or column are called **vectors**.
- Matrices having just one row are **row vectors**.
- Matrices having just one column are **column vectors**.
- Its entries are called the **components** of the vector.

- To economize page space, row vectors are written

$$\mathbf{a} = [a_1, a_2, \ldots, a_n]$$

and column vectors are written

$$\mathbf{a} = (a_1, a_2, \ldots, a_n)$$

Thus,

$$(a_1, a_2, \ldots, a_n) = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}.$$

- **The space $\mathbb{R}^n$**

- The scalar product of two vectors

$$\mathbf{x} = (x_1, x_2, \ldots, x_n) \quad \text{and} \quad \mathbf{y} = (y_1, y_2, \ldots, y_n)$$

is defined as

$$\boxed{\mathbf{x}^T\mathbf{y} = x_1y_1 + x_2y_2 + \cdots + x_ny_n = \sum_{i=1}^{n} x_iy_i}$$

The vectors $\mathbf{x}$ and $\mathbf{y}$ are said to be **orthogonal** if $\mathbf{x}^T\mathbf{y} = 0$.

- The **magnitude** or **norm** of vector $\mathbf{x}$ is

$$\boxed{|\mathbf{x}| = (\mathbf{x}^T\mathbf{x})^{1/2} = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}}$$

- The **distance** between $\mathbf{x}$ and $\mathbf{y}$ is

$$\boxed{d(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}| = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}}$$

- **Convex Sets**

Definition 1.1

A set C in $\mathbb{R}^n$ is said to be **convex** if for every $\mathbf{x}, \mathbf{y} \in C$ and every real number λ, $0 < \lambda < 1$, the point $(1-\lambda)\mathbf{x} + \lambda\mathbf{y} \in C$.

$$
\begin{aligned}
C \text{ convex} &\iff (\forall \mathbf{x}, \mathbf{y} \in C)(\forall \lambda \in (0,1))\big((1-\lambda)\mathbf{x} + \lambda\mathbf{y} \in C\big) \\
&\iff (\forall \mathbf{x}, \mathbf{y} \in C)(\forall \lambda \in [0,1])\big((1-\lambda)\mathbf{x} + \lambda\mathbf{y} \in C\big)
\end{aligned}
$$

- The agree that the empty set $\emptyset$ is convex.

Geometrically speaking, a set C is convex if for every pair of points $\mathbf{x}, \mathbf{y}$ in C the **line segment**

$$[\mathbf{x}, \mathbf{y}] := \{(1-\lambda)\mathbf{x} + \lambda\mathbf{y} : 0 \leq \lambda \leq 1\}$$

belongs to C.

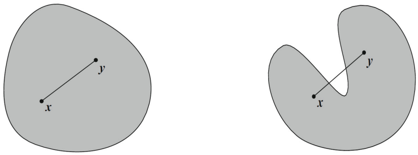


Figure 1.1 Convex and nonconvex sets

Theorem 1.1

(a) *If C is a convex set and α is a real number, the set*

$$\alpha C = \{\mathbf{x} : \mathbf{x} = \alpha \mathbf{c},\ \ \mathbf{c} \in C\}$$

is convex.

(b) *If C and D are convex sets, then the set*

$$C + D = \{\mathbf{x} : \mathbf{x} = \mathbf{y} + \mathbf{z},\ \ \mathbf{y} \in C,\ \ \mathbf{z} \in D\}$$

is convex.

(c) *The intersection of any collection of convex sets is convex.*

Definition 1.2

A point $\mathbf{x}$ is said to be a **convex combination** of points $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ in $\mathbb{R}^n$, if there exist nonnegative numbers $\lambda_1, \lambda_2, \ldots, \lambda_k$ with

$$\lambda_1 + \lambda_2 + \cdots + \lambda_k = 1$$

such that

$$\mathbf{x} = \lambda_1 \mathbf{x}_1 + \lambda_2 \mathbf{x}_2 + \cdots + \lambda_k \mathbf{x}_k.$$

Theorem 1.2

A set C is convex if and only if it contains all convex combinations of points in C.

Theorem 1.3

The closure and the interior of a convex set are convex.

$$\boxed{C \text{ is convex} \Longrightarrow \overline{C} \text{ and } \operatorname{int} C \text{ are convex}}$$

Note Empty set is convex.

Definition 1.3

A set C is a **cone** if $\mathbf{x} \in C$ implies $\lambda\mathbf{x} \in C$ for all $\lambda > 0$. A cone that is also convex is a **convex cone**.

The point $\mathbf{0}$, called the **vertex** of C, may or may not belong to C.

If $\mathbf{x} = (x_1, x_2, \ldots, x_n)$, $\mathbf{y} = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n$, we write

- $\mathbf{x} \leq \mathbf{y}$ if $x_i \leq y_i$ for all $i = 1, 2, \ldots, n$, and
- $\mathbf{x} < \mathbf{y}$ if $x_i < y_i$ for all $i = 1, 2, \ldots, n$.

In particular,

$\mathbf{x} \geq \mathbf{0}$ means that $x_i \geq 0$ for all $i = 1, 2, \ldots, n$.

Example 1.1 (a) The convex cones

$$\{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \geq \mathbf{0}\} \quad \text{and} \quad \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} > \mathbf{0}\}$$

are called **non-negative orthant** and **positive orthant**, respectively.

(b) The sets

$$\{\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2 : x_1 \cdot x_2 = 0\}$$

and

$$\{\mathbf{x} = (x_1, x_2) \in \mathbb{R}^2 : x_1 \cdot x_2 \geq 0\}$$

are cones but they are not convex.

- **Extreme Points**

Definition 1.4

A point $\mathbf{x}$ in a convex set C is said to be an **extreme point** of C if there are no two distinct points $\mathbf{x}_1$ and $\mathbf{x}_2$ in C such that

$$\mathbf{x} = (1-\lambda)\mathbf{x}_1 + \lambda\mathbf{x}_2 \quad \text{for some} \quad \lambda \in (0,1).$$

An extreme point is thus a point that does not lie strictly within a line segment connecting two other points of the set.

- **Hyperplanes and Half Spaces**

Definition 2.1

A set V in $\mathbb{R}^n$ is said to be a **linear variety** (or an **affine set**), if, given any $\mathbf{x}, \mathbf{y} \in V$, we have

$$(1-\lambda)\mathbf{x} + \lambda\mathbf{y} \in V \quad \text{for all} \quad \lambda \in \mathbb{R}.$$

$$V \text{ is affine } \iff (\forall \mathbf{x}, \mathbf{y} \in V)(\forall \lambda \in \mathbb{R})\big((1-\lambda)\mathbf{x} + \lambda\mathbf{y} \in V\big).$$

Remark If $A \subset \mathbb{R}^n$ is an affine subset of $\mathbb{R}^n$, and $a \in A$ is an arbitrary point, then $L = A - a$ is a linear subspace of $\mathbb{R}^n$, which is independent of $a \in A$. Conversely, if $a \in \mathbb{R}^n$ and L is a linear subspace of $\mathbb{R}^n$, then $A := a + L$ is an affine subset of $\mathbb{R}^n$. In this case, the **dimension** of A is defined as the dimension of L.

Definition 2.2

A **hyperplane** in $\mathbb{R}^n$ is an $(n-1)$-dimensional linear variety.

Theorem 2.1

(a) *Let* $\mathbf{a}$ *be a nonzero n-dimensional column vector, and let* α *be a real number. The set*

$$H = \left\{ \mathbf{x} \in \mathbb{R}^n : \mathbf{a}^T \mathbf{x} = \alpha \right\}$$

is a hyperplane in $\mathbb{R}^n$*.*

(b) *Let* H *be a hyperplane in* $\mathbb{R}^n$*. Then there is a* *nonzero* *n-dimensional vector* $\mathbf{a}$ *and a constant* α *such that*

$$H = \left\{ \mathbf{x} \in \mathbb{R}^n : \mathbf{a}^T \mathbf{x} = \alpha \right\}$$

Definition 2.3

Let **a** be a nonzero vector in $\mathbb{R}^n$ and let α be a real number. Corresponding to the hyperplane

$$H = \left\{\mathbf{x} \in \mathbb{R}^n : \mathbf{a}^T\mathbf{x} = \alpha\right\}$$

are the **positive** and **negative closed half spaces**

$$H_+ = \left\{\mathbf{x} : \mathbf{a}^T\mathbf{x} \geq \alpha\right\}, \quad H_- = \left\{\mathbf{x} : \mathbf{a}^T\mathbf{x} \leq \alpha\right\}$$

and the **positive** and **negative open half spaces**

$$\mathring{H}_+ = \left\{\mathbf{x} : \mathbf{a}^T\mathbf{x} > \alpha\right\}, \quad \mathring{H}_- = \left\{\mathbf{x} : \mathbf{a}^T\mathbf{x} < \alpha\right\}.$$

Remark H, H_+, H_-, $\overset{\circ}{H}_+$, and $\overset{\circ}{H}_-$ are convex sets. So

The solution set of a system of linear equalities and inequalities is a convex set.

Definition 2.4

Let B, C be nonempty subsets of $\mathbb{R}^n$, and let H be a hyperplane in $\mathbb{R}^n$.

(i) H is said to **separate** B and C if B lies in one of the closed halfspaces determined by H, and C lies in the other.

(ii) H is said to **separate** B and C **strictly** if B and C lie in opposite open halfspaces determined by H.

- The hyperplane $H = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{a}^T\mathbf{x} = \alpha\}$ separates sets B and C in $\mathbb{R}^n$ if

$$\sup_{\mathbf{u}\in B} \mathbf{a}^T\mathbf{u} \le \alpha \le \inf_{\mathbf{v}\in C} \mathbf{a}^T\mathbf{v} \quad \text{or} \quad \sup_{v\in C} \mathbf{a}^T\mathbf{v} \le \alpha \le \inf_{\mathbf{u}\in B} \mathbf{a}^T\mathbf{u}.$$

- H separates B and C strictly if either

$$\mathbf{a}^T\mathbf{u} < \alpha < \mathbf{a}^T\mathbf{v} \quad \text{for all } \mathbf{u} \in B,\ \mathbf{v} \in C$$

or

$$\mathbf{a}^T\mathbf{v} < \alpha < \mathbf{a}^T\mathbf{u} \quad \text{for all } \mathbf{u} \in B,\ \mathbf{v} \in C.$$

Theorem 2.2

Let C be a nonempty closed convex set and let $\mathbf{x}_0 \in \mathbb{R}^n$ do not belong to C. Then there is a nonzero vector $\mathbf{a} \in \mathbb{R}^n$ such that

$$\mathbf{a}^T \mathbf{x}_0 < \inf_{\mathbf{x} \in C} \mathbf{a}^T \mathbf{x}.$$

In other words, a nonempty closed convex set in $\mathbb{R}^n$ may be strictly separated from a point not belonging to it.

Definition 2.5

A hyperplane containing a convex set C in one of its closed half spaces and containing a boundary point of C is said to be a **supporting hyperplane** of C.

Theorem 2.3

Let C be a convex set and let $\mathbf{x}_0$ be a boundary point of C. Then there is a hyperplane supporting C at $\mathbf{x}_0$.

Corollary 2.4

Let $C \subset \mathbb{R}^n$ be a convex set and let $\mathbf{x}_0 \notin C$. Then there exists a nonzero vector $\mathbf{a}$ such that

$$\mathbf{a}^T \mathbf{x}_0 \leq \inf_{\mathbf{x} \in C} \mathbf{a}^T \mathbf{x}.$$

Theorem 2.5

Let B and C be nonempty convex sets with no common points. Then there is a hyperplane separating B and C. That is, there is a nonzero vector $\mathbf{a}$ *such that*

$$\sup_{\mathbf{u}\in B} \mathbf{a}^T\mathbf{u} \leq \inf_{\mathbf{v}\in C} \mathbf{a}^T\mathbf{v}.$$

1.3 LINEAR PROGRAMMING PROBLEMS AND EXAMPLES

In many business and economic problems we are asked to optimize (maximize or minimize) a function subject to a system of equalities or inequalities.

The function to be optimized is called the **objective function**.

In general, an optimization problem can then be written as

$$\begin{aligned} \text{minimize} \quad & f(\mathbf{x}) \\ \text{subject to} \quad & h_i(\mathbf{x}) = 0, \quad i \in \mathcal{I}, \\ & g_j(\mathbf{x}) \leq 0, \quad j \in \mathcal{J}, \\ & \mathbf{x} \in \Omega \subset \mathbb{R}^n, \end{aligned}$$

where f and each h_i, g_j are scalar-valued functions of the variables $\mathbf{x}$, and $\mathcal{I}, \mathcal{J}$ are sets of indices.

Minimization versus maximization

Given any function f and constraint set S, any vector $\bar{\mathbf{x}}$ that solves the problem

$$\max_{x \in S} f(x)$$

is also a solution of the problem

$$\min_{x \in S}(-f(x)),$$

and vice versa since

$$\max_{x \in S} f(x) = -\min_{x \in S}(-f(x)),$$

A **linear program (LP)** is an optimization problem in which the objective function is linear in the unknowns and the constraints consist of linear equalities and linear inequalities.

Linear Programming is the process of minimizing a linear objective function subject to a finite number of linear equality and inequality constraints.

A **linear function** is a function of the form

$$f(x_1, x_2, \ldots, x_n) = c_1x_1 + c_2x_2 + \cdots + c_nx_n$$

where $c_1, c_2, \ldots, c_n$ are constants and $x_1, x_2, \ldots, x_n$ are variables.

We shall say that a linear programming problem is in **standard form** if it is in the following form:

$$
\begin{array}{lrcl}
\text{minimize} & c_1x_1 + c_2x_2 + \cdots + c_nx_n & & \\
\text{subject to} & a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & b_1 \\
 & a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = & b_2 \\
 & \vdots \qquad\qquad\qquad\qquad \vdots & \vdots & \\
 & a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = & b_m \\
\text{and} & x_1 \geq 0,\ x_2 \geq 0,\ \ldots, x_n \geq 0, & &
\end{array}
$$

where the b_i's, c_i's and a_{ij}'s are fixed real constants, and the x_i's are real numbers to be determined.

This form is standard in the sense that the variables are all nonnegative and the other constraints are linear equations.

• By grouping the variables $x_1, x_2, \ldots, x_n$ into a column vector $\mathbf{x}$ and constructing the following matrix and vectors from the problem data,

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix}, \quad \mathbf{c} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix},$$

we can restate the standard form compactly as follows:

$$\begin{aligned} \text{minimize} \quad & \mathbf{c}^T\mathbf{x} \\ \text{subject to} \quad & \mathbf{A}\mathbf{x} = \mathbf{b} \\ & \mathbf{x} \geq \mathbf{0}. \end{aligned}$$

Here the vector inequality $\mathbf{x} \geq \mathbf{0}$ means that each component of $\mathbf{x}$ is nonnegative.

History

• In 1939 L. V. Kantorovich published a monograph entitled *Mathematical Methods in the Organization and Planning of Production*.

Kantorovich recognized that a broad class of production problems led to the same mathematical problem and that this problem was susceptible to solution by numerical methods.

However, Kantorovich's work went unrecognized.

1.3 LINEAR PROGRAMMING PROBLEMS AND EXAMPLES

- In 1941 Frank Hitchcock formulated the transportation problem.
- 1945 George Stigler considered the problem of determining an adequate diet for an individual at minimal cost.

  Through these problems and others, especially problems related to the World War II effort, it became clear that a feasible method for solving linear programming problems was needed.
- Then in 1951 George Dantzig developed the simplex method.

Converting LPs to standard form.

Every linear program can be put into this standard form.

Example 3.1 **(Slack variables).** Consider the problem

$$
\begin{array}{lrcl}
\text{minimize} & c_1x_1 + c_2x_2 + \cdots + c_nx_n & & \\
\text{subject to} & a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & \leq & b_1 \\
 & a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & \leq & b_2 \\
 & \vdots \qquad\qquad\qquad\qquad \vdots & \vdots & \\
 & a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & \leq & b_m \\
\text{and} & x_1 \geq 0,\ x_2 \geq 0,\ \ldots, x_n \geq 0. & &
\end{array}
$$

The problem may be alternatively expressed as

$$
\begin{array}{lllcl}
\text{minimize} & c_1x_1 + c_2x_2 + \cdots + c_nx_n & & & \\
\text{subject to} & a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n + y_1 & & = & b_1 \\
 & a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & + y_2 & = & b_2 \\
 & \vdots \qquad\qquad \vdots & & \vdots & \\
 & a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & + y_m & = & b_m \\
\text{and} & x_1 \geq 0, \ x_2 \geq 0, \ \ldots, x_n \geq 0, & & & \\
 & y_1 \geq 0, \ y_2 \geq 0, \ \ldots, y_m \geq 0. & & &
\end{array}
$$

The new positive variables y_i introduced to convert the inequalities to equalities are called **slack variables** (or more loosely, **slacks**).

Example 3.2 **(Surplus variables)**. If the linear inequalities of constraints are reversed so that a typical inequality is

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n \geq b_i,$$

then it is clear that this is equivalent to

$$a_{i1}x_1 + a_{i2}x_2 + \cdots + a_{in}x_n - y_i = b_i,$$

with $y_i \geq 0$.

- These variables y_i are called **surplus variables**.

It should be clear that by suitably multiplying by minus unity, and adjoining slack and surplus variables, any set of linear inequalities can be converted to standard form if the unknown variables are restricted to be nonnegative.

Example 3.3 The problem

$$\begin{array}{ll} \text{minimize} & 3x_1 + 2x_2 + 4x_3 \\ \text{subject to} & 30x_1 + 100x_2 + 85x_3 \leq 2500 \\ & 6x_1 + 2x_2 + 3x_3 \geq 90 \\ & x_1 \geq 0,\ x_2 \geq 0,\ x_3 \geq 0. \end{array}$$

This problem is equivalent to the following problem

$$\begin{array}{ll} \text{minimize} & 3x_1 + 2x_2 + 4x_3 \\ \text{subject to} & 30x_1 + 100x_2 + 85x_3 + x_4 = 2500 \\ & 6x_1 + 2x_2 + 3x_3 - x_5 = 90 \\ & x_1,\ x_2,\ x_3,\ x_4,\ x_5 \geq 0. \end{array}$$

A variable without specified lower or upper bounds is called a **free** or **unrestricted** variable.

Example 3.4 **(Free variables-first method)**
If a linear program is given in standard form except that one or more of the unknown variables is not required to be nonnegative, the problem can be transformed to standard form by either of two simple techniques.

Suppose, for example, that the restriction $x_1 \geq 0$ is not present and hence x_1 is *free* to take on either positive or negative values.

In the first method we write

$$x_1 = u_1 - v_1$$

where we require $u_1 \geq 0$ and $v_1 \geq 0$. If we substitute $u_1 - v_1$ for x_1 everywhere in the LP, the linearity of the constraints is preserved and all variables are now required to be nonnegative.

The problem is then expressed in terms of the $n + 1$ variables $u_1, v_1, x_2, x_3, \ldots, x_n$.

For instance, the problem of maximizing $3x_1 - 2x_2 - x_3 + x_4$ subject to

$$\begin{aligned} 4x_1 \;\; - x_2 \qquad\quad + x_4 &\leq 6 \\ -7x_1 + 8x_2 + x_3 \qquad\quad &\geq 7 \\ x_1,\ x_2,\ x_3 \geq 0,\ x_4 &\text{ unrestricted} \end{aligned}$$

is equivalent to

$$\begin{array}{ll} \text{minimize} & -3x_1 + 2x_2 + x_3 - u_4 + v_4 \\ \text{subject to} & 4x_1 - \ \ x_2 \qquad + u_4 - v_4 + x_5 \quad = 6 \\ & -7x_1 + 8x_2 + x_3 \qquad\qquad\qquad - x_6 = 7 \\ & \qquad x_1,\ x_2,\ x_3,\ u_4,\ v_4, x_5,\ x_6 \geq 0. \end{array}$$

A second approach for converting to standard form when x_1 is unconstrained in sign is to eliminate x_1 together with one of the constraint equations.

Examples of Linear Programming Problems

Example 3.5 **(The diet problem)**. Assume that there are available at the market n different foods and that the jth food sells at a price c_j per unit. In addition there are m basic nutritional ingredients and, to achieve a balanced diet, each individual must receive at least b_i units of the ith nutrient per day.

Finally, we assume that each unit of food j contains a_{ij} units of the ith nutrient.

If we denote by x_j the number of units of food j in the diet, the problem then is to select the x_j's to minimize the total cost

$$c_1x_1 + c_2x_2 + \cdots + c_nx_n$$

subject to the nutritional constraints

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &\geq b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &\geq b_2 \\ \vdots \qquad\qquad\qquad\qquad\quad \vdots \;\; &\quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &\geq b_m \\ x_1, x_2, \ldots, x_n &\geq 0. \end{aligned}$$

Example 3.6 **(The transportation problem).** Quantities $a_1, a_2, \ldots, a_m$, respectively, of a certain product are to be shipped from each of m locations and received in amounts $b_1, b_2, \ldots, b_n$, respectively, at each of n destinations. Associated with the shipping of a unit of product from origin i to destination j is a unit shipping cost c_{ij}. It is desired to determine the amounts x_{ij} to be shipped between each origin-destination pair $i = 1, 2, \ldots, m; j = 1, 2, \ldots, n$; so as to satisfy the shipping requirements and minimize the total cost of transportation.

The total cost is

$$\sum_{j=1}^{n}\sum_{i=1}^{m} c_{ij}x_{ij}.$$

Thus, we have the linear programming problem:

$$\begin{array}{llll}
\text{minimize} & \sum_{i,j} c_{ij}x_{ij} & & \\
\text{subject to} & \sum_{j=1}^{n} x_{ij} = a_i & \text{for} \quad i = 1, 2, \ldots, m & (1) \\
& \sum_{j=1}^{n} x_{ij} = b_j & \text{for} \quad j = 1, 2, \ldots, n & (2) \\
& x_{ij} \geq 0 & \text{for} \quad i = 1, 2, \ldots, m; & \\
& & \qquad j = 1, 2, \ldots, n. &
\end{array}$$

In order that the constraints (1), (2) be consistent, we must assume that

$$\sum_{i=1}^{m} a_i = \sum_{j=1}^{n} b_j$$

which corresponds to assuming that the total amount shipped is equal to the total amount received.

Example 3.7 **(Manufacturing problem).** Suppose we own a facility that is capable of manufacturing n different products, each of which may require various amounts of m different resources. Each product can be produced at any level $x_j \geq 0$, $j = 1, 2, \ldots, n$, and each unit of the jth product can sell for p_j dollars and needs a_{ij} units of the ith resource, $i = 1, 2, \ldots, m$. Assume linearity of the production facility and we are given a set of m numbers $b_1, b_2, \ldots, b_m$ describing the available quantities of the m resources, and we wish to manufacture products at maximum revenue.

Ours decision problem is a linear program:

$$
\begin{array}{rr}
\text{maximize} & p_1x_1 + p_2x_2 + \cdots + p_nx_n \\
\text{subject to} & a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \leq b_1 \\
& a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \leq b_2 \\
& \vdots \\
& a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \leq b_m \\
& x_1,\ x_2, \ldots, x_n \geq 0.
\end{array}
$$

- Linear programs can be studied both algebraically and geometrically.
- The two approaches are equivalent, but one or the other may be more convenient for answering a particular question about a linear program.
- The algebraic point of view is based on writing the linear program in standard form.
- Then the coefficient matrix of the constraints of the linear program can be analyzed using the tools of linear algebra.

Consider the system of equalities

$$\mathbf{Ax} = \mathbf{b} \tag{3}$$

where $\mathbf{x}$ is an n-vector, $\mathbf{b}$ an m-vector, and $\mathbf{A}$ is an $m \times n$ matrix. Suppose that from the n columns of $\mathbf{A}$ we select a set of m linearly independent columns. For simplicity assume that we select the first m columns of $\mathbf{A}$ and denote the $m \times m$ matrix determined by these columns by $\mathbf{B}$. The matrix $\mathbf{B}$ is then nonsingular and we may uniquely solve the equation.

$$\mathbf{Bx_B} = \mathbf{b}$$

for the m-vector $\mathbf{x_B}$. By putting $\mathbf{x} = (\mathbf{x_B}, \mathbf{0})$, obtain a solution to $\mathbf{Ax} = \mathbf{b}$.

Here is an example. Let

$$\mathbf{A} = \begin{bmatrix} 2 & -1 & 0 & -1 \\ -1 & 2 & -1 & 1 \\ 0 & -1 & 1 & 0 \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}.$$

Then

$$\mathbf{B} = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

is a basis in $\mathbf{A}$ and the equation $\mathbf{B}\mathbf{x_B} = \mathbf{b}$ has a unique solution $\mathbf{x_B} = (x_1, x_2, x_3) = (3, 5, 6)$. Thus $\mathbf{x} = (3, 5, 6, 0, 0)$ is a solution to $\mathbf{A}\mathbf{x} = \mathbf{b}$.

Definition 3.1

Given the set of m simultaneous linear equations in n unknowns (3), let $\mathbf{B}$ be any nonsingular $m \times m$ submatrix made up of columns of $\mathbf{A}$. Then, if all $n - m$ components of $\mathbf{x}$ not associated with columns of $\mathbf{B}$ are set equal to zero, the solution to the resulting set of equations is said to be a **basic solution** to (3) with respect to the **basis B**. The components of $\mathbf{x}$ associated with columns of $\mathbf{B}$ are called **basic variables**.

Example 4.1 Consider the linear system

$$x_1 + x_2 - 3x_3 + x_4 = 5$$
$$2x_1 - x_2 + 3x_3 + x_4 = 10.$$

One basis is $\{(1,2),(1,-1)\}$. The basic solution corresponding to this basis is the vector

$$\mathbf{x} = (5,0,0,0).$$

We explicitly make the following assumption in our development, unless noted otherwise.

Full rank assumption

The $m \times n$ matrix **A** *has $m < n$, and the m rows of* **A** *are linearly independent, that is,* $\text{rank}(A) = m$.

In this case we also say that the matrix **A** has **full rank**.

Definition 4.2

If one or more of the basic variables in a basic solution has value zero, that solution is said to be a **degenerate basic solution**.

For instance, in Example 4.1, $\mathbf{x} = (5, 0, 0, 0)$ is degenerate basic solution. Different bases

$$\begin{bmatrix} 1 & 1 \\ 2 & -1 \end{bmatrix}, \quad \begin{bmatrix} 1 & -3 \\ 2 & 3 \end{bmatrix}, \quad \text{and} \quad \begin{bmatrix} 1 & 1 \\ 2 & 1 \end{bmatrix}$$

give rise to this degenerate basic feasible solution.

Note A basic solution is a solution to the system $\mathbf{Ax} = \mathbf{b}$; it *does not* necessarily satisfy $\mathbf{x} \geq \mathbf{0}$.

Consider now the system of constraints

$$\begin{aligned} \mathbf{Ax} &= \mathbf{b} \\ \mathbf{x} &\geq \mathbf{0}, \end{aligned} \tag{4}$$

which represent the constraints of a linear program in standard form.

Definition 4.3

A vector $\mathbf{x}$ satisfying (4) is called a **feasible solution** for these constraints.

A feasible solution to the constraints (4) that is also basic is said to be a **basic feasible solution**; if this solution is also a degenerate basic solution, it is called a **degenerate basic feasible solution**.

The **feasible region** of a linear program is the set of all its feasible solutions.

If a problem has no feasible solution, then the problem itself is called **infeasible**.

Remark A feasible solution is a *basic feasible solution* if the set of column vectors corresponding to positive components of the solution is independent.

The Fundamental theorem of LP

Consider a linear program in standard form

$$\begin{array}{lrcl} \text{minimize} & \mathbf{c}^T\mathbf{x} & & \\ \text{subject to} & \mathbf{A}\mathbf{x} & = & \mathbf{b} \\ & \mathbf{x} & \geq & \mathbf{0}. \end{array} \tag{5}$$

Definition 5.1

A feasible solution to the constraints that achieves the minimum value of the objective function subject to those constraints is said to be an **optimal feasible solution**. If this solution is basic, it is an **optimal basic feasible solution**.

Theorem 4.1 (Fundamental theorem of linear programming)

Given a linear program in standard form (5) where $\mathbf{A}$ *is an* $m \times n$ *matrix of rank* m,

(i) *if there is a feasible solution, there is a basic feasible solution;*

(ii) *if there is an optimal feasible solution, there is an optimal basic feasible solution.*

♣ **The idea of the proof.**

Basic feasible solutions and extreme points

Theorem 5.1 (Equivalence of vertices and basic solutions)

A point $\mathbf{x}$ *is an extreme point of the set*

$$\{\mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0}\}$$

if and only if it is a basic feasible solution.

An extreme point of

$$\{\mathbf{x} : \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0}\}$$

is also called a **vertex**.

Example 5.1 Consider the constraint set in $\mathbb{R}^3$ defined by

$$\begin{aligned} x_1 + x_2 + x_3 &= 1 \\ 2x_1 + 3x_2 \qquad &= 1 \\ x_1 \geq 0,\ x_2 \geq 0,\ x_3 &\geq 0. \end{aligned}$$

(a) Determine basic solutions of the system of equations.

(b) Determine vertices of the constraint set.

Corollary 5.2

If the convex set $\{\mathbf{x} : \mathbf{A}\mathbf{x} = \mathbf{b},\ \mathbf{x} \geq \mathbf{0}\}$ *is nonempty, it has at least one vertex.*

Corollary 5.3

If there is an optimal solution to a linear programming problem, there is an optimal solution which is a vertex of the constraint set.

Corollary 5.4

(i) *There are only a finite number of basic feasible solutions.*

(ii) *There are only finitely many vertices of the feasible set.*

In deed, if $\mathbf{A}$ is a $m \times n$ matrix of rank m, there are exactly

$$\binom{n}{m} = \frac{n!}{m!(n-m)!}$$

ways to choose m columns from the n columns of $\mathbf{A}$.

Example 5.2 Consider the following two-dimensional linear programming problem:

$$\begin{aligned} \text{maximize} \quad & z = 3x + 2y \\ \text{subject to} \quad & 2x + 3y \leq 12 \\ & 2x + \ \ y \leq 8 \\ & x \geq 0, \ \ y \geq 0. \end{aligned}$$

ANS. The solution is $(3, 2)$ and $z_{\max} = 13$.

Summary

- In searching for an optimal solution of a linear program in standard form, we may restrict our attention to basic feasible solutions of the constraints.
- A linear program in standard form can have only finitely many basic feasible solutions.

But, this finite number can be extremely large, even for problems of modest size.

For example, a problem with $m = 200$ and $n = 500$ would have about 10^{144} basic solutions.

In the next chapter we will develop a method that determines basic feasible solutions in a particular order that allows us to find an optimal solution in a *small number of trials*.