L. R. Foulds

Optimization Techniques

An Introduction

With 72 Illustrations

L. R. Foulds
Department of Economics
University of Canterbury
Christchurch 1
New Zealand

AMS Classification: 49-01, 90-01

Library of Congress Cataloging in Publication Data

Foulds, L. R., 1948–
Optimization techniques.

Bibliography: p.
Includes index.
1. Mathematical optimization. 2. Programming
(Mathematics) I. Title.
QA402.5.F68 519 81-5642
AACR2

Printed in the United States of America.

9 8 7 6 5 4 3 2 1

ISBN 0-387-90586-3 Springer-Verlag New York Heidelberg Berlin
ISBN 3-540-90586-3 Springer-Verlag Berlin Heidelberg New York

Contents

Chapter 1

Introduction

1.1 Motivation for Studying Optimization

There exist an enormous variety of activities in the everyday world which can usefully be described as systems, from actual physical systems such as chemical processing plants to theoretical entities such as economic models. The efficient operation of these systems often requires an attempt at the optimization of various indices which measure the performance of the system. Sometimes these indices are quantified and represented as algebraic variables. Then values for these variables must be found which maximize the gain or profit of the system and minimize the waste or loss. The variables are assumed to be dependent upon a number of factors. Some of these factors are often under the control, or partial control, of the analyst responsible for the performance of the system.

The process of attempting to manage the limited resources of a system can usually be divided into six phases: (i) mathematical analysis of the system; (ii) construction of a mathematical model which reflects the important aspects of the system; (iii) validation of the model; (iv) manipulation of the model to produce a satisfactory, if not optimal, solution to the model; (v) implementation of the solution selected; and (vi) the introduction of a strategy which monitors the performance of the system after implementation. It is with the fourth phase, the manipulation of the model, that the theory of optimization is concerned. The other phases are very important in the management of any system and will probably require greater total effort than the optimization phase. However, in the presentation of optimization theory here it will be assumed that the other phases have been, or will be, carried out. Because the theory of optimization provides this link in the chain of systems management it is an important body of mathematical knowledge.

1.2 The Scope of Optimization

One of the most important tools of optimization is *linear programming.* A linear programming problem is specified by a linear, multivariable function which is to be optimized (maximized or minimized) subject to a number of linear constraints. The mathematician G. B. Dantzig (1963) developed an algorithm called the *simplex method* to solve problems of this type. The original simplex method has been modified into an efficient algorithm to solve large linear programming problems by computer. Problems from a wide variety of fields of human endeavor can be formulated and solved by means of linear programming. Resource allocation problems in government planning, network analysis for urban and regional planning, production planning problems in industry, and the management of transportation distribution systems are just a few. Thus linear programming is one of the successes of modern optimization theory.

Integer programming is concerned with the solution of optimization problems in which at least some of the variables must assume only integer values. In this book only integer programming problems in which all terms are linear will be covered. This subtopic is often called *integer linear programming.* However, because little is known about how to solve nonlinear integer programming problems, the word linear will be assumed here for all terms. Many problems of a combinatorial nature can be formulated in terms of integer programming. Practical examples include facility location, job sequencing in production lines, assembly line balancing, matching problems, inventory control, and machine replacement. One of the important methods for solving these problems, due to R. E. Gomory (1958), is based in part on the simplex method mentioned earlier. Another approach is of a combinatorial nature and involves reducing the original problem to smaller, hopefully easier, problems and partitioning the set of possible solutions into smaller subsets which can be analyzed more easily. This approach is called *branch and bound* or *branch and backtrack.* Two of the important contributions to this approach have been by Balas (1965) and Dakin (1965). Although a number of improvements have been made to all these methods, there does not exist as yet a relatively efficient method for solving realistically-sized integer programming problems.

Another class of problems involves the *management of a network.* Problems in traffic flow, communications, the distribution of goods, and project scheduling are often of this type. Many of these problems can be solved by the methods mentioned previously—linear or integer programming. However because these problems usually have a special structure, more efficient specialized techniques have been developed for their solution. Outstanding contributions have been made in this field by Ford and Fulkerson (1962). They developed the *labelling method* for maximizing the flow of a commodity through a network and the *out-of-kilter method* for minimizing the cost of transporting a given quantity of a commodity through a network. These ideas can be combined with those of integer programming to analyze a whole host of practical network problems.

Some problems can be decomposed into parts, the decision processes of which are then optimized. In some instances it is possible to attain the optimum for the original problem solely by discovering how to optimize these constituent parts. This decomposition process is very powerful, as it allows one to solve a series of smaller, easier problems rather than one large, intractable problem. Systems for which this approach will yield a valid optimum are called *serial multistage systems.* One of the best known techniques to attack such problems was named *dynamic programming* by the mathematician who developed it, R. E. Bellman (1957). Serial multistage systems are characterized by a process which is performed in stages, such as manufacturing processes. Rather than attempting to optimize some performance measure by looking at the problem as a whole, dynamic programming optimizes one stage at a time to produce an optimal set of decisions for the whole process. Problems from all sorts of areas, such as capital budgeting, machine reliability, and network analysis, can be viewed as serial multistage systems. Thus dynamic programming has wide applicability.

In the formulation of many optimization problems the assumption of linearity cannot be made, as it was in the case of linear programming. There do not exist general procedures for nonlinear problems. A large number of specialized algorithms have been developed to treat special cases. Many of these procedures are based on the mathematical theory concerned with analysing the structure of such problems. This theory is usually termed *classical optimization.* One of the outstanding modern contributions to this theory has been made by Kuhn and Tucker (1951) who developed what are known as the Kuhn–Tucker conditions.

The collection of techniques developed from this theory is called *nonlinear programming.* Despite the fact that many nonlinear programming problems are very difficult to solve, there are a number of practical problems which can be formulated nonlinearly and solved by existing methods. These include the design of such entities as electrical transformers, chemical processes, vapour condensors, microwave matching networks, gallium–arsenic light sources, digital filters, and also problems concerning maximum likelihood estimation and optimal parts replacement.

1.3 Optimization as a Branch of Mathematics

It can be seen from the previous section that the theory of optimization is mathematical in nature. Typically it involves the maximization or minimization of a function (sometimes unknown) which represents the performance of some system. This is carried out by the finding of values for those variables (which are both quantifiable and controllable) which cause the function to yield an optimal value. A knowledge of linear algebra and differential multivariable calculus is required in order to understand how the algorithms operate. A sound knowledge of analysis is necessary for an understanding of the theory.

Some of the problems of optimization theory can be solved by the classical techniques of advanced calculus—such as Jacobian methods and the use of Lagrange multipliers. However, most optimization problems do not satisfy the conditions necessary for solution in this manner. Of the remaining problems many, although amenable to the classical techniques, are solved more efficiently by methods designed for the purpose. Throughout recorded mathematical history a collection of such techniques has been built up. Some have been forgotten and reinvented, others received little attention until modern-day computers made them feasible.

The bulk of the material of the subject is of recent origin because many of the problems, such as traffic flow, are only now of concern and also because of the large numbers of people now available to analyze such problems. When the material is catalogued into a meaningful whole the result is a new branch of applied mathematics.

1.4 The History of Optimization

One of the first recorded instances of optimization theory concerns the finding of a geometric curve of given length which will, together with a straight line, enclose the largest possible area. Archimedes conjectured correctly that the optimal curve is a semicircle. Some of the early results are in the form of principles which attempt to describe and explain natural phenomena. One of the earliest examples was presented approximately 100 years after Archimedes' conjecture. It was formulated by Heron of Alexandria in C. 100 B.C., who postulated that light always travels by the shortest path. It was not until 1657 that Fermat correctly generalized this postulate by stating that light always travels by the path which incurs least time rather than least distance.

The fundamental problem of another branch of optimization is concerned with the choosing of a function that minimizes certain functionals. (A functional is a special type of function whose domain is a set of real-valued functions.) Two problems of this nature were known at the time of Newton. The first involves finding a curve such that the solid of revolution created by rotating the curve about a line through its endpoints causes the minimum resistance when this solid is moved through the air at constant velocity. The second problem is called the *brachistochrone*. In this problem two points in space are given. One wishes to find the shape of a curve joining the two points, such that a frictionless bead travelling on the curve from one point to the other will cover the journey in least time. This problem was posed as a competiton by John Bernoulli in 1696. The problem was successfully solved by Bernoulli himself, de l'Hôpital, Leibniz, and Newton (who took less than a day!). Problems such as these led Euler to develop the ideas involved into a systematic discipline which he called the *calculus of variations* in 1766. Also at the time of Euler many laws of mechanics were first formulated in terms of principles of optimality (examples are the least action principle of Maupertuis, the principle of least restraint of Gauss, and Lagrange's kinetic principle). Lagrange and Gauss both made other contributions. In 1760 Lagrange invented a method for solving optimization problems that had equality constraints using his *Lagrange multipliers*. Lagrange transformations are, among other uses, employed to examine the behaviour of a function in the neighbourhood of a suspected optimum. And Gauss, who made contributions to many fields, developed the method of *least squares curve fitting* which is of interest to those working in optimization as well as statistics.

In 1834 W. R. Hamilton developed a set of functions called Hamiltonians which were used in the statement of a principle of optimality that unified what was known of optics and mechanics at that time. In 1875 J. W. Gibbs presented a further principle of optimality concerned with the equilibrium of a thermodynamical system. Between that time and the present there have been increasing numbers of contributions each year. Among the most outstanding recent achievements, the works of Dantzig and of Bellman have already been mentioned. Another is the work of Pontryagin (1962) and others, who developed the *maximum principle* which is used to solve problems in the theory of optimal control.

1.5 Basic Concepts of Optimization

This section introduces some of the basic concepts of optimization. Each concept is illustrated by means of the following example.

The problem is to:

$$\text{Maximize:} \quad x_0 = f(X) = f(x_1, x_2) \tag{1.1}$$

$$\text{subject to:} \quad h_1(X) \leq 0 \tag{1.2}$$

$$x_1 \geq 0 \tag{1.3}$$

$$x_2 \geq 0. \tag{1.4}$$

This is a typical problem in the theory of optimization—the maximization (or minimization) of a real-valued function of a number of real variables (sometimes just a single variable) subject to a number of constraints (sometimes the number is zero). The special case of functionals, where the domain of the function is a set of functions, will be dealt with under the section on the calculus of variations in Chapter 7.

The function f is called the *objective function.* The set of constraints, in this case a set of inequalities, is called the *constraint set.* The problem is to find real values for x_1 and x_2, satisfying (1.2), (1.3) and (1.4), which when inserted in (1.1) will cause $f(x_1, x_2)$ to take on a value no less than that for any other such x_1, x_2 pair. Hence x_1 and x_2 are called *independent variables.*

Three objective function contours are present in Figure 1.1. The objective function has the same value at all points on each line, so that the contours can be likened to isobar lines on a weather map. Thus it is not hard to see

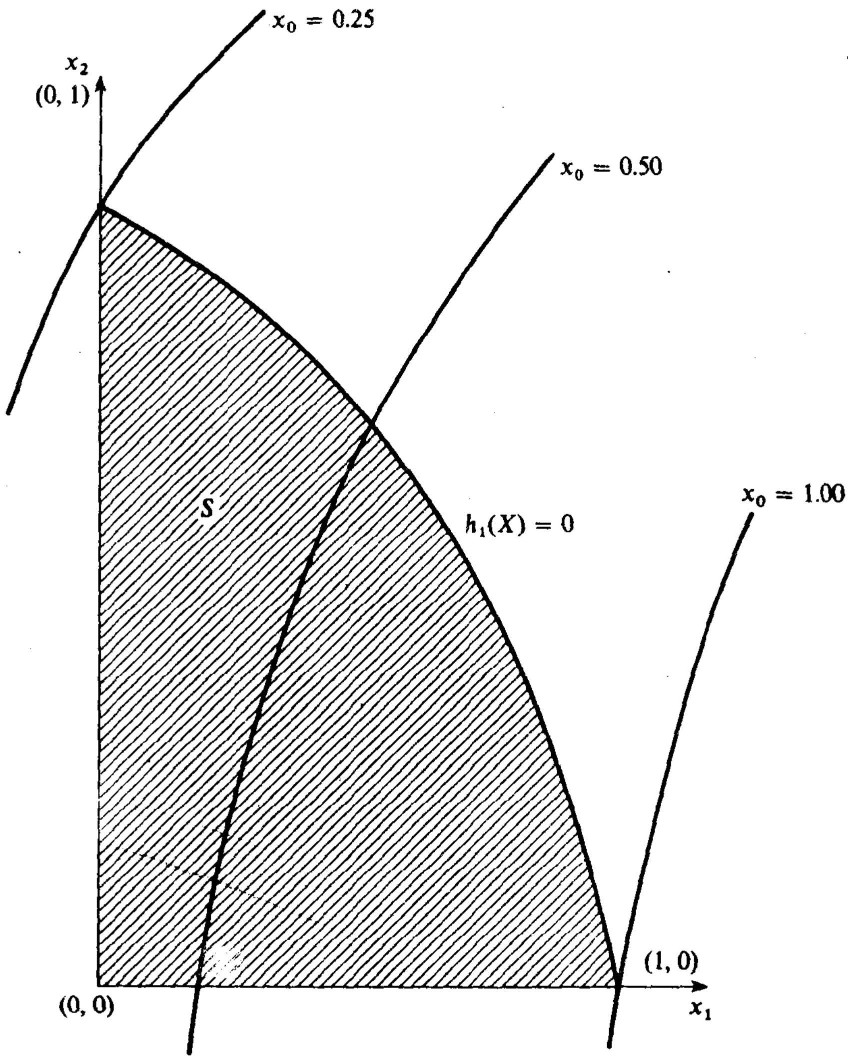

Figure 1.1. Objective function contours and the feasible region for an optimization problem.

that the solution to the problem is:

$$X^* = (x_1^*, x_2^*) = (1,0).$$

This means that

$$f(X^*) \geq f(X) \quad \text{for all } X \in S. \tag{1.5}$$

When a solution $X^* \in S$ satisfies (1.5) it is called the *optimal solution*, and in this case the *maximal solution*. If the symbol in (1.5) were "$\leq$", X^* would be called the *minimal solution*. Also, $f(X^*)$ is called the *optimum* and is written x_0^*.

On looking at Figure 1.1 it can be seen that greater values for f could be obtained by choosing certain x_1, x_2 outside S. Any ordered pair of real numbers is called a *solution* to the problem and the corresponding value of f is called the *value* of the solution. A solution X such that

$$X \in S$$

is called a *feasible solution*.

Let us examine which x_1, x_2 pairs are likely candidates to achieve this maximization. In Figure 1.1 the set of points which satisfy this constraint set has been shaded. The set is defined as S:

$$S = \{(x_1, x_2): h(x_1, x_2) \leq 0, x_1 \geq 0, x_2 \geq 0\}.$$

Such a set S for an optimization problem is often a connected region and is called the *feasible region*.

Many optimization problems do not have unique optimal solutions. For instance, suppose a fourth constraint

$$h_2(x_1, x_2) \leq 0 \tag{1.6}$$

is added to the problem. The feasible region is shown in Figure 1.2. In this case one of the boundaries of S coincides with an objective function contour. Thus all points on that boundary represent maximum solutions.

However, if it exists the optimal value is always unique.

As another example of a problem which does not have an optimal solution, suppose (1.2) is replaced by:

$$h_1(X) < 0. \tag{1.7}$$

On examining Figure 1.2, it becomes apparent that (1.7) does not hold for $X^* = (1,0)$, hence $X^* \notin S$. In fact, there is no solution which will satisfy (1.5), as points successively closer to (but a positive distance away from) (1, 0) correspond to successively larger x_0 values. To recognize this situation we called $f(X')$ an *upper bound for f under S* if

$$f(X') \geq f(X) \quad \text{for all } X \in S. \tag{1.8}$$

Also $f(X')$ is called a *least upper bound* or *supremum for f under S* if $f(X')$ is an upper bound for f under S and

$$f(X') \leq f(X) \quad \text{for all upper bounds } f(X) \text{ for } f \text{ under } S. \tag{1.9}$$

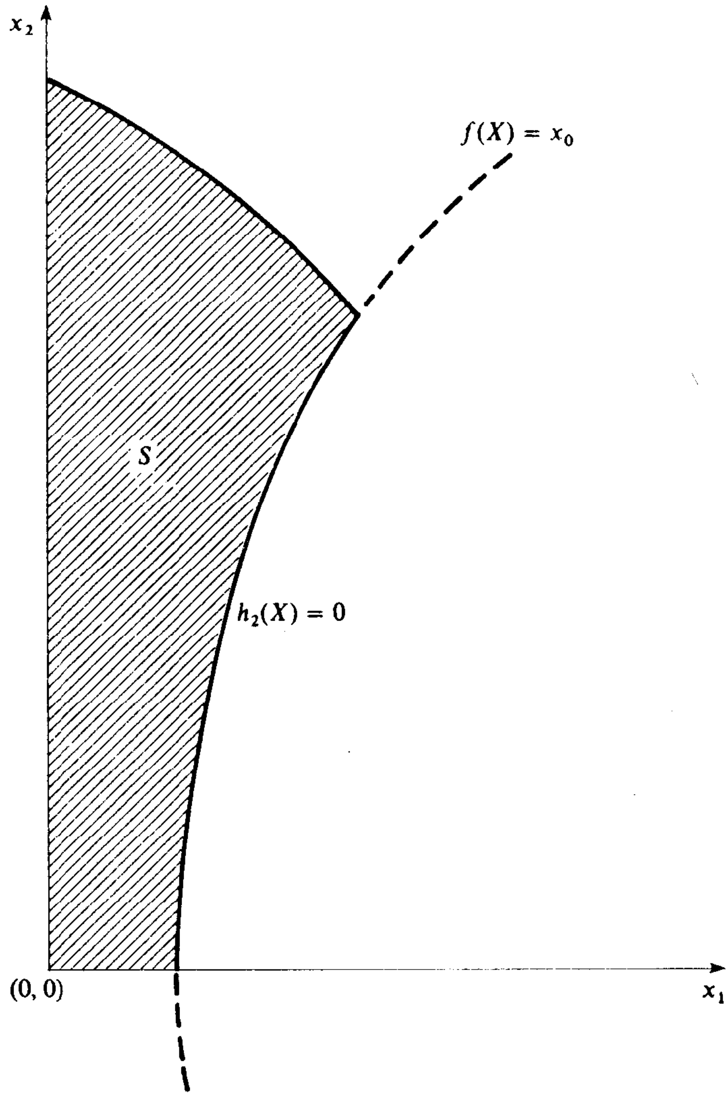

Figure 1.2. Feasible region for an optimization problem where one constraint is identical with an objective function contour.

Most of the preceding ideas have been concerned with maximization. Of course many optimization problems have the aim of minimization and each of the above concepts has a minimization counterpart. The sense of the inequalities in (1.7), (1.8), and (1.9) need to be reversed for minimization. The counterparts of the terms are:

minimum	maximum
lower bound	upper bound
greatest lower bound	least upper bound
infimum	supremum

Throughout the remainder of book we shall deal mainly with maximization problems only, because of the following theorem.

Theorem 1.1. *If X^* is the optimal solution to problem* P1:

$$\begin{aligned} \text{Maximize:}\quad & f(X), \\ \text{subject to:}\quad & g_j(X) = 0, \qquad j = 1, 2, \ldots, m \\ & h_j(X) \le 0, \qquad j = 1, 2, \ldots, k \end{aligned}$$

then X^* is the optimal solution to problem P2:

$$\begin{aligned} \text{Minimize:}\quad & -f(X), \\ \text{subject to:}\quad & g_j(X) = 0, \qquad j = 1, 2, \ldots, m \\ & h_j(X) \le 0, \qquad j = 1, 2, \ldots, k. \end{aligned}$$

PROOF. Because X^* is the optimal solution for P1, it is a feasible solution for P1, hence

$$\begin{aligned} g_j(X^*) = 0, &\qquad j = 1, 2, \ldots, m \\ h_j(X^*) \le 0, &\qquad j = 1, 2, \ldots, k. \end{aligned}$$

Hence X^* is a feasible solution for P2.

Also,

$$f(X^*) \ge f(X) \quad \text{for all } X \in S$$

where

$$S = \{X: g_j(X) = 0, j = 1, 2, \ldots, m; h_j(X) \le 0, j = 1, 2, \ldots, k\}.$$

Hence

$$-f(X^*) \le -f(X) \quad \text{for all } X \in S.$$

Hence X^* is optimal for P2. □

This result allows us to solve any minimization problem by multiplying its objective function by -1 and solving a maximization problem under the same constraints. Of course we could have just as easily proven another theorem concerning the conversion of any maximization problem into an equivalent minimization problem.

Chapter 2

Linear Programming

2.1 Introduction

This present chapter is concerned with a most important area of optimization, in which the objective function and all the constraints are linear. Problems in which this is not the case fall in the nonlinear programming category and will be covered in Chapters 7 and 8.

There are a large number of real problems that can be either formulated as linear programming (L.P.) problems or formulated as models which can be successfully approximated by linear programming. Relatively small problems can readily be solved by hand, as will be explained later in the chapter. Large problems can be solved by very efficient computer programs. The mathematical structure of L.P. allows important questions to be answered concerning the sensitivity of the optimum to data changes. L.P. is also used as a subroutine in the solving of more complex problems in nonlinear and integer programming.

This chapter will begin by introducing the basic ideas of L.P. with a simple example and then generalize. A very efficient method for solving L.P. problems, the simplex method, will be developed and it will be shown how the method deals with the different types of complications that can arise. Next the idea of a dual problem is introduced with a view to analyzing the behaviour of the optimal L.P. solution when the problem is changed. This probing is called postoptimal analysis. Algorithms for special L.P. problems will also be looked at.

2.2 A Simple L.P. Problem

A coal mining company producing both lignite and anthracite finds itself in the happy state of being able to sell all the coal it can process. The present profit is \$4.00 and \$3.00 (in hundreds of dollars) for a ton of lignite and anthracite, respectively. However, because of various restrictions the cutting machine at the coal face, the screens, and the washing plant can be operated for no more than 12, 10, and 8 hours per day, respectively. It requires 3, 3, and 4 hours for the cutting machine, the screens, and the washing plant, respectively, to process one ton of lignite. It requires 4, 3, and 2 hours for the cutting machine, the screens, and the washing plant, respectively, to process one ton of anthracite. The problem is to decide how many tons of each type of coal will be produced so as to maximize daily profits.

In order to solve this problem we need to express it in mathematical terms. Toward this end the decision (independent) variables are defined as follows. Let

x_1 = the daily production of lignite in tons,
x_2 = the daily production of anthracite in tons,
x_0 = the profit gained by producing x_1 and x_2 tons of lignite and anthracite, respectively.

If x_1 tons of lignite are produced each day, and the profit per ton is \$4.00 then the daily profit for lignite is

$$\$4x_1.$$

Similarly, if x_2 tons of anthracite are produced each day with a profit of \$3.00 per ton, then the daily profit is

$$\$3x_2.$$

Thus for a daily production schedule of x_1 and x_2 tons of lignite and anthracite, the total daily profit, in dollars, is:

$$4x_1 + 3x_2\ (=x_0).$$

It is this expression whose value we must maximize.

We can formulate similar expressions for the constraints of time on the various machines. For instance, consider the cutting operation. If x_1 tons of lignite are produced each day and each ton of lignite requires 3 hours' cutting time, then the total cutting time required to produce those x_1 tons of lignite is

$$3x_1 \text{ hours.}$$

Similarly, if x_2 tons of anthracite are produced each day with each ton taking 4 hours to cut, the total cutting time required to produce those x_1 tons of anthracite is

$$4x_2 \text{ hours.}$$

Thus the total cutting time for x_1 tons of lignite and x_2 tons of anthracite is

$$3x_1 + 4x_2.$$

But only 12 hours' cutting time are available each day. Hence we have the constraint:

$$3x_1 + 4x_2 \leq 12.$$

We can formulate similar constraints for the screening and washing times. This has been done below. The problem can now be stated mathematically:

$$\begin{aligned} \text{Maximize:} \quad & 4x_1 + 3x_2 = x_0 && (2.1)\\ \text{subject to:} \quad & 3x_1 + 4x_2 \leq 12 && (2.2)\\ & 3x_1 + 3x_2 \leq 10 && (2.3)\\ & 4x_1 + 2x_2 \leq 8 && (2.4)\\ & x_1 \geq 0 && (2.5)\\ & x_2 \geq 0. && (2.6) \end{aligned}$$

The above expressions are now explained:

(2.1): The objective is to maximize daily profit.
(2.2): A maximum of 12 hours cutting time is available each day.
(2.3): A maximum of 10 hours screening time is available each day.
(2.4): A maximum of 8 hours washing time is available each day.
(2.5), (2.6): A nonnegative amount of each type of coal must be produced.

Because only two independent variables are present it is possible to solve the problem graphically. This can be achieved by first plotting the constraints (2.2)–(2.6) in two-dimensional space. The origin can be used to test which half-plane created by each constraint contains feasible points. The feasible region is shown in Figure 2.1. It can be seen that constraint (2.3) is *redundant*, in the sense that it does not define part of the boundary of the feasible region. The arrow on constraint (2.3) denotes the feasible half-plane defined by the constraint. The problem now becomes that of selecting the point in the feasible region which corresponds to the maximum objective function value — the optimum. This point is found by setting the objective function equal to a number of values and plotting the resulting lines. Clearly, the maximum value corresponds to point $(\frac{4}{5}, \frac{12}{5})$. Thus the optimal solution is

$$x_1^* = \tfrac{4}{5} \quad \text{and} \quad x_2^* = \tfrac{12}{5},$$

with value $10\frac{2}{5}$. Hence the best profit the company can hope to make is \$1,040 by producing 0.8 tons of lignite and 2.4 tons of anthracite per day.

When more than two independent variables are present, linear programs are solved by analytic methods, as it is difficult to draw in three dimensions and impossible in higher dimensions. The next section introduces the general problem.

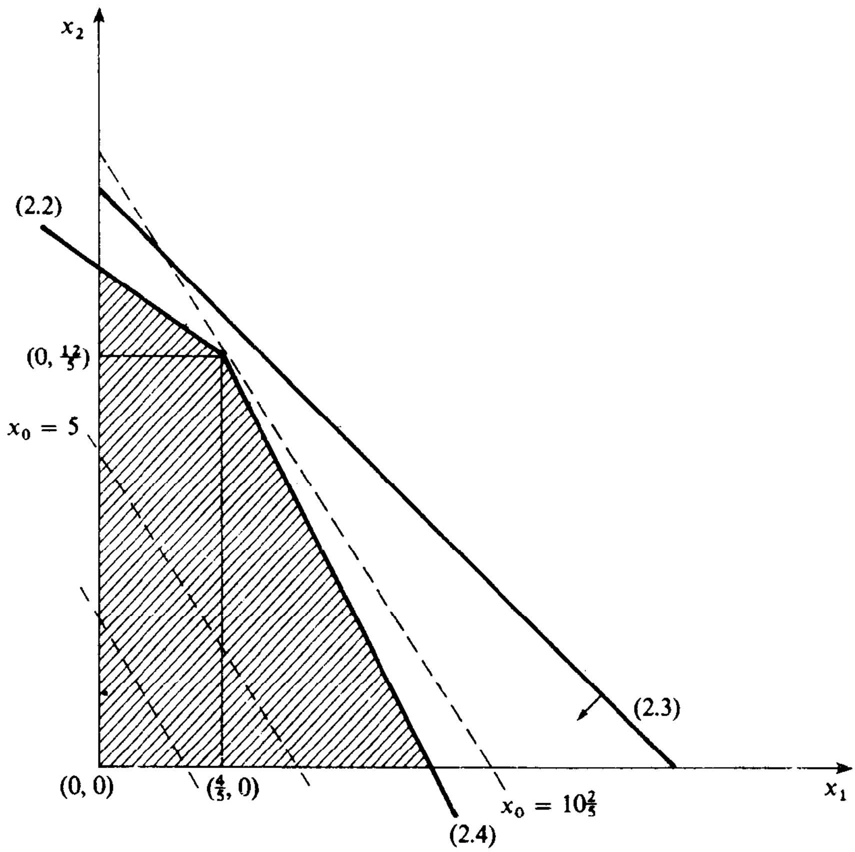

Figure 2.1. Graphical solution to the L.P. example problem.

2.3 The General L.P. Problem

The problem of (2.1)–(2.6) can be generalized as follows:

$$\begin{aligned}
\text{Maximize:} \quad & c_1x_1 + c_2x_2 + \cdots + c_nx_n = x_0 \\
\text{subject to:} \quad & a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \le b_1 \\
& a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \le b_2 \\
& \qquad \vdots \\
& a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \le b_m \\
& x_i \ge 0, \qquad i = 1, 2, \ldots, n.
\end{aligned}$$

Of course this problem can be stated in matrix form:

$$\begin{aligned}
\text{Maximize:} \quad & C^TX \\
\text{subject to:} \quad & AX \le B, \\
& X \ge 0,
\end{aligned}$$