Quantitative Methods for the Management Sciences 45-760 Course Notes

INSTRUCTORS:
GERARD CORNUEJOLS
MICHAEL TRICK
Graduate School of Industrial Administration
Carnegie Mellon University, Pittsburgh, PA 15213 USA

Fall 1998

Contents

Course Syllabus

0.1 Introduction

Objectives of the Course:

- to provide a formal quantitative approach to problem solving;
- to give an intuition for managerial situations where a quantitative approach is appropriate;
- to introduce some widely used quantitative models;
- to introduce software for solving such models.

To accomplish these objectives we provide:

- background in linear algebra and nonlinear optimization;
- description of the linear programming model, decision theory, etc.;
- explanation of the mathematical ideas behind these widely used models;
- hands-on experience with software such as SOLVER;
- examples of business applications;
- a case study that shows how models are used in practice.

Instructors

Gérard Cornuéjols, GSIA 232A, phone 268-2284, fax 268-7357
email: gc0v@andrew.cmu.edu
secretary: Barbara Carlson 268-1342, GSIA 232.

Michael Trick, GSIA 257C, phone 268-3697, fax 268-7057
email: trick@andrew.cmu.edu

Required Material for the Course

The course notes are self contained. The software for this course is installed in your notebooks. Instruction on using it will be provided in class.

0.2 Grading Components

The course grade will be based on homework, a case study, and two exams.

Homework

The only way to learn mathematical material is by working problems.

- Homework is due at the beginning of class on the due date (see schedule). No late homework will be accepted.
- You are encouraged to work exercises that are not assigned. Some can be found in the course notes. More can be found in the textbook by Winston, *Operations Research: Applications and Algorithms*, Duxbury Press (1994). Copies are on reserve in the Hunt library.
- You should do the homework without consulting other students. If you need help, you may consult the instructor or a teaching assistant. For *nonassigned* exercises, it is a good idea to work them out with other students.
- Every attempt will be made to return graded homeworks promptly. Solutions to the homework will be distributed in class after the due date.
- The following rules apply to all homework:
  - Use only 11 by 8.5 in. paper.
  - Start each problem on a new page.
  - Assemble the homework in proper order and staple it together.
  - PRINT your name on the top right hand corner.

Case Study

The case study has two parts. The first part will require formulating a linear program. The second part requires solving the linear program and analyzing the results for managerial purposes.

Exams

- There will be one midterm and one final exam. The precise dates of the exams will be announced by the GSIA administration.
- The questions in the exams will be similar to the homework problems and the problems discussed in class both in form and content. The lectures may occasionally range over peripheral topics, but the content of the exercises is a reliable guide to which topics can appear on the exam.
- All exams will be **open-notes**.
- Except for provable unforeseen circumstances, makeup exams will not be given. A student is required to inform the instructor as soon as possible in such a case. In case of an unavoidable conflict, it is the student's responsibility to schedule a makeup exam *before* the exam date after discussing with the instructor.

Grading Weights

Midterm exam = 25%, homework = 15%, case study = 25%, final exam = 35%.

0.3 Schedule

It is advisable to read through the material in the course notes prior to when it is discussed in class.

Date		Topic	Chapter	Work Due
Aug	25	Linear Algebra	Chapter 1	
Aug	27	Functions of One Variable	Chapter 2	
Sept	1	Functions of Several Variables	Chapter 3	HW 1 due
Sept	3	Optimization with Constraints	Chapter 4	
Sept	8	LP modeling / Solver	Chapter 5	HW 2 due
Sept	10	LP modeling	Chapter 5	
Sept	??	MIDTERM EXAM		TBA
Sept	15	The Simplex Method	Chapter 7	HW 3 due
Sept	17	Sensitivity Analysis	Chapter 8	
Sept	22	Case Study Part I	Chapter 6	Case Part I due
Sept	24	Decision Theory	Chapter 9	
Sept	29	Game Theory	Chapter 10	HW 4 due
Oct	1	Network Models	Chapter 11	
Oct	6	Data Envelopment Analysis	Chapter 12	Case Part II due
Oct	8	Problem Session		HW 5 due
Oct	??	FINAL EXAM		TBA

0.4 Mathematical Methods in Business

0.4.1 History

Mathematical methods have long played an important role in management and economics but have become increasingly important in the last few decades. “Business mathematics,” such as the computation of compound interest, appeared in ancient Mesopotamia but became prevalent in the mercantile economy of the Renaissance. Mathematics later contributed to the engineering advances that made the Industrial Revolution possible. Mathematics crept into economics during the 18th and 19th centuries and became firmly established in the 20th.

It is only in the last few decades that management per se has received the sort of rigorous study that permits the application of mathematics. In the early part of this century “scientific management,” called “Taylorism” after its founder F. Taylor, was fashionable. This movement made some worthwhile contributions, such as time-and-motion studies, but it was coupled with dreadful labor relations practices. The symbol of the movement was the efficiency expert policing the shop floor with stopwatch and clipboard.

After Taylorism waned, interest shifted more to making efficient use of labor and other factors than making employees work harder. Shortly before the Second World War a team of scientists solved the operational problems of coordinating Britain’s radar stations and thereby created “operational research,” called “operations research” in the U.S. During the war this practice of putting scientists and mathematicians to work on operational problems was surprisingly successful in both the British and American military, partly because members of the professional class who had not previously “dirtied their hands” had their first opportunity to apply their training to operational problems. Their success attracted attention, and operations research spread to industry during the fifties.

After the war, G. Dantzig and others developed linear programming at about the same time that computers became available to solve linear programming models. Linear programming proved a remarkably useful tool for the efficient allocation of resources, and this gave a great boost to operations research. In the early postwar years J. von Neumann and O. Morgenstern invented game theory (closely related to linear programming), and H. W. Kuhn and A. W. Tucker broke ground in nonlinear programming. W. E. Deming and others developed statistical techniques for quality control, and the statistical methods first designed for psychometrics in the early 20th century became the mathematical basis for econometrics. Meanwhile business schools began teaching many of these techniques along with microeconomics and other quantitative fields. As a result of all this, mathematical modeling has played an increasingly important role in management and economics.

In fact one can make a case that overreliance on a mathematical approach has sometimes led to neglect of other approaches. After its first few years, operations research focused almost exclusively on mathematical modeling. The reigning orthodoxy in much of economics is the "neoclassical" paradigm, which is heavily mathematical. There is little communication between people in quantitative fields and those in such "soft" fields as organizational science and general management.

In fact both approaches appear to be in a crisis. Economics and operations research have achieved much, but neither can adequately understand phenomena that involve human beings. The soft sciences take on the difficult, ill-structured, human-oriented problems, but it is unclear even what counts as a satisfactory theory in these areas.

One can hope that this double crisis will lead in the coming years to a "paradigm shift" that will supplant the "hard" and "soft" approaches with an integrated science with the rigor of one and the breadth of the other. There are faint signs of such a shift, but they are not sufficiently crystallized to show up in a business school curriculum.

Despite the limitations of the quantitative methods discussed in this course, in many specific areas they are quite useful. In fact the potential of most mathematical techniques, particularly those of operations research, is only beginning to be realized in practice. Affordable personal computers and user-friendly software are motivating managers to learn about mathematical models and use them. Also the line between managerial and technical staff is blurring, so that managers now run models on their notebook computers that systems analysts once ran on mainframes. This again leads to more widespread use.

0.4.2 Mathematical Modeling Today

The applications of mathematical methods in management and economics today are so manifold that it is difficult to find a single person who is aware of their full scope. The following list can only be incomplete.

- *Economists* use linear and nonlinear programming, the theory of variational inequalities, optimal control theory, dynamic programming, game theory, probability choice models, utility theory, regression and factor analysis and other techniques to study equilibrium, optimal investment, competition, consumer behavior, and a host of other phenomena.

- People in *operations management* use statistical sampling and estimation theory, linear and integer programming, network programming, dynamic programming and optimal control theory, queuing theory, simulation, artificial intelligence techniques, and combinatorial optimization methods to solve problems in quality control, allocation of resources, logistics, project scheduling, labor and machine scheduling, job shop scheduling and assembly line balancing, and facility layout and location. The introduction of flexible manufacturing systems, robots and other automated devices has posed a whole new array of unsolved mathematical problems.

- People in *finance* use linear, nonlinear and integer programming, optimal control theory and dynamic programming, Markov decision theory, regression and time series to determine optimal resource allocation, multiperiod investments, capital budgeting, and investment and loan portfolio design, and to try to forecast market behavior.

- People in *marketing* use regression and factor analysis, time series, game theory, Markov decision theory, location theory, mathematical programming, probability choice models and utility theory to study consumer preferences, determine optimal location in product space, allocate advertising resources, design distribution systems, forecast market behavior, and study competitive strategy.

- People in *information systems* and decision support systems use artificial intelligence techniques, propositional and quantified logic, Bayesian methods, probabilistic logic, data structures and other computer science techniques, mathematical programming, and statistical decision theory to design expert and other knowledge-based systems, develop efficient inference and retrieval methods, and evaluate the economic and organizational effects of information systems.

The peculiar nature of mathematics unfortunately raises two obstacles to learning it. One is that students often believe that mathematics can be learned simply by studying a book. Mathematics is learned by working through problems. A second obstacle is that students often believe that they can learn to work problems by studying a book. A book can get one started, but learning to work problems is like learning to play basketball or play the piano—it requires practice, practice, practice. Mathematical skills are very athletic in this sense. (The phrase "quant jock" is appropriate.) That is why this course assigns a lot of homework problems.

Acknowledgements: We would like to thank John Hooker, Bill Hrusa, Rick Green and Anuj Mehrotra for their inputs.

Chapter 1

Basic Linear Algebra

Linear equations form the basis of linear programming. If you have a good understanding of the Gauss-Jordan method for solving linear equations, then you can also understand the solution of linear programs. In addition, this chapter introduces matrix notation and concepts that will be used in Chapters 3 and 4 (optimizing functions of several variables).

1.1 Linear Equations

The Gauss-Jordan elimination procedure is a systematic method for solving systems of linear equations. It works one variable at a time, eliminating it in all rows but one, and then moves on to the next variable. We illustrate the procedure on three examples.

Example 1.1.1 (Solving linear equations)

$$\begin{array}{rcrcrcr} x_1 & + & 2x_2 & + & x_3 & = & 4 \\ 2x_1 & - & x_2 & + & 3x_3 & = & 3 \\ x_1 & + & x_2 & - & x_3 & = & 3 \end{array}$$

In the first step of the procedure, we use the first equation to eliminate x_1 from the other two. Specifically, in order to eliminate x_1 from the second equation, we multiply the first equation by 2 and subtract the result from the second equation. Similarly, to eliminate x_1 from the third equation, we subtract the first equation from the third. Such steps are called *elementary row operations*. We keep the first equation and the modified second and third equations. The resulting equations are

$$\begin{array}{rcrcrcr} x_1 & + & 2x_2 & + & x_3 & = & 4 \\ & - & 5x_2 & + & x_3 & = & -5 \\ & - & x_2 & - & 2x_3 & = & -1 \end{array}$$

Note that only one equation was used to eliminate x_1 in all the others. This guarantees that the new system of equations has exactly the same solution(s) as the original one. In the second step of the procedure, we divide the second equation by -5 to make the coefficient of x_2 equal to 1. Then, we use this equation to eliminate x_2 from equations 1 and 3. This yields the following new system of equations.

$$\begin{array}{rcrcrcr} x_1 & & & + & \frac{7}{5}x_3 & = & 2 \\ & & x_2 & - & \frac{1}{5}x_3 & = & 1 \\ & & & - & \frac{11}{5}x_3 & = & 0 \end{array}$$

Once again, only one equation was used to eliminate x_2 in all the others and that guarantees that the new system has the same solution(s) as the original one. Finally, in the last step of the procedure, we use equation 3 to eliminate x_3 in equations 1 and 2.

$$\begin{array}{rcl} x_1 & & & = 2 \\ & x_2 & & = 1 \\ & & x_3 & = 0 \end{array}$$

So, there is a unique solution. Note that, throughout the procedure, we were careful to keep three equations that have the same solution(s) as the original three equations. Why is it useful? Because, linear systems of equations do not always have a unique solution and it is important to identify such situations.

Another example:

$$\begin{array}{rcrcrcr} x_1 & + & 2x_2 & + & x_3 & = & 4 \\ x_1 & + & x_2 & + & 2x_3 & = & 1 \\ 2x_1 & + & 3x_2 & + & 3x_3 & = & 2 \end{array}$$

First we eliminate x_1 from equations 2 and 3.

$$\begin{array}{rcrcrcr} x_1 & + & 2x_2 & + & x_3 & = & 4 \\ & - & x_2 & + & x_3 & = & -3 \\ & - & x_2 & + & x_3 & = & -6 \end{array}$$

Then we eliminate x_2 from equations 1 and 3.

$$\begin{array}{rrcrcr} x_1 & & + & 3x_3 & = & -2 \\ & x_2 & - & x_3 & = & 3 \\ & & & 0 & = & -3 \end{array}$$

Equation 3 shows that the linear system has *no solution.*

A third example:

$$\begin{array}{rcrcrcr} x_1 & + & 2x_2 & + & x_3 & = & 4 \\ x_1 & + & x_2 & + & 2x_3 & = & 1 \\ 2x_1 & + & 3x_2 & + & 3x_3 & = & 5 \end{array}$$

Doing the same as above, we end up with

$$\begin{array}{rrcrcr} x_1 & & + & 3x_3 & = & -2 \\ & x_2 & - & x_3 & = & 3 \\ & & & 0 & = & 0 \end{array}$$

Now equation 3 is an obvious equality. It can be discarded to obtain

$$\begin{array}{rrcrcr} x_1 & & = & -2 & - & 3x_3 \\ & x_2 & = & 3 & + & x_3 \end{array}$$

The situation where we can express some of the variables (here x_1 and x_2) in terms of the remaining variables (here x_3) is important. These variables are said to be *basic* and *nonbasic* respectively. Any choice of the nonbasic variable x_3 yields a solution of the linear system. Therefore the system has infinitely many solutions.

It is generally true that a system of m linear equations in n variables has either:

(a) no solution,

(b) a unique solution,

(c) infinitely many solutions.

The Gauss-Jordan elimination procedure solves the system of linear equations using two elementary row operations:

- modify some equation by multiplying it by a nonzero scalar (a *scalar* is an actual real number, such as $\frac{1}{2}$ or -2; it cannot be one of the variables in the problem),
- modify some equation by adding to it a scalar multiple of another equation.

The resulting system of m linear equations has the same solution(s) as the original system. If an equation $0 = 0$ is produced, it is discarded and the procedure is continued. If an equation $0 = a$ is produced where a is a nonzero scalar, the procedure is stopped: in this case, the system has no solution. At each step of the procedure, a new variable is made basic: it has coefficient 1 in one of the equations and 0 in all the others. The procedure stops when each equation has a basic variable associated with it. Say p equations remain (remember that some of the original m equations may have been discarded). When $n = p$, the system has a unique solution. When $n > p$, then p variables are basic and the remaining $n - p$ are nonbasic. In this case, the system has infinitely many solutions.

Exercise 1 Solve the following systems of linear equations using the Gauss-Jordan elimination procedure and state whether case (a), (b) or (c) holds.

(1)

$$\begin{array}{rcrcrcr} 3x_1 & & & - & 4x_3 & = & 2 \\ x_1 & + & x_2 & + & x_3 & = & 4 \\ & & 2x_2 & + & x_3 & = & 3 \end{array}$$

(2)

$$\begin{array}{rcrcrcr} 2x_1 & + & 2x_2 & - & x_3 & = & 1 \\ 4x_1 & & & + & x_3 & = & 2 \\ x_1 & - & x_2 & + & x_3 & = & 2 \end{array}$$

(3)

$$\begin{array}{rcrcrcr} x_1 & - & x_2 & + & x_3 & = & 1 \\ -2x_1 & + & 2x_2 & - & 2x_3 & = & -2 \\ -x_1 & + & x_2 & - & x_3 & = & -1 \end{array}$$

Exercise 2 Indicate whether the following linear system of equations has 0, 1 or infinitely many solutions.

$$\begin{array}{rcrcrcrcrcl} x_1 & + & 2x_2 & + & 4x_3 & + & x_4 & + & 3x_5 & = & 2 \\ 2x_1 & + & x_2 & + & x_3 & + & 3x_4 & + & x_5 & = & 1 \\ & & 3x_2 & + & 7x_3 & - & x_4 & + & 5x_5 & = & 6 \end{array}$$

1.2 Operations on Vectors and Matrices

It is useful to formalize the operations on vectors and matrices that form the basis of linear algebra. For our purpose, the most useful definitions are the following.

A *matrix* is a rectangular array of numbers written in the form

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}.$$

The matrix A has *dimensions* $m \times n$ if it has m rows and n columns. When $m = 1$, the matrix is called a *row vector*; when $n = 1$, the matrix is called a *column vector*. A *vector* can be represented either by a row vector or a column vector.

Equality of two matrices of *same* dimensions:

Let $A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$ and $B = \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \dots & b_{mn} \end{pmatrix}$.

Then $A = B$ means that $a_{ij} = b_{ij}$ for all i, j.

Multiplication of a matrix A by a scalar k:

$$kA = \begin{pmatrix} ka_{11} & ka_{12} & \dots & ka_{1n} \\ \vdots & \vdots & & \vdots \\ ka_{m1} & ka_{m2} & \dots & ka_{mn} \end{pmatrix}.$$

Addition of two matrices of *same* dimensions:

Let $A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$ and $B = \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & & \vdots \\ b_{m1} & \dots & b_{mn} \end{pmatrix}$.

Then $A + B = \begin{pmatrix} a_{11} + b_{11} & \dots & a_{1n} + b_{1n} \\ \vdots & & \vdots \\ a_{m1} + b_{m1} & \dots & a_{mn} + b_{mn} \end{pmatrix}$.

Note that $A + B$ is *not defined* when A and B have different dimensions.

Exercise 3 Compute $3\begin{pmatrix} 2 & 3 & 1 \\ 0 & 5 & 4 \end{pmatrix} - 2\begin{pmatrix} 0 & 4 & 1 \\ 1 & 6 & 6 \end{pmatrix}$.

Multiplication of a matrix of dimensions $m \times n$ by a matrix of dimensions $n \times p$:

Let $A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{i1} & \dots & a_{in} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$ and $B = \begin{pmatrix} b_{11} & \dots & b_{1j} & \dots & b_{1n} \\ \vdots & & \vdots & & \vdots \\ b_{m1} & \dots & b_{mj} & \dots & b_{mn} \end{pmatrix}$.

Then AB is a matrix of dimensions $m \times p$ computed as follows.

$$AB = \begin{pmatrix} a_{11}b_{11} + \ldots + a_{1n}b_{n1} & \dots & a_{11}b_{1p} + \ldots + a_{1n}b_{np} \\ \vdots & a_{i1}b_{1j} + \ldots + a_{in}b_{nj} & \vdots \\ a_{m1}b_{11} + \ldots + a_{mn}b_{n1} & \dots & a_{m1}b_{1p} + \ldots a_{mn}b_{np} \end{pmatrix}.$$

As an example, let us multiply the matrices

$$A = \begin{pmatrix} 2 & 3 & 4 \\ 5 & 1 & 0 \end{pmatrix} \text{ and } B = \begin{pmatrix} 11 & 12 \\ 14 & 15 \\ 17 & 18 \end{pmatrix}.$$

The result is $AB = \begin{pmatrix} 2 \times 11 + 3 \times 14 + 4 \times 17 & 2 \times 12 + 3 \times 15 + 4 \times 18 \\ 5 \times 11 + 1 \times 14 + 0 \times 17 & 5 \times 12 + 1 \times 15 + 0 \times 18 \end{pmatrix} = \begin{pmatrix} 132 & 141 \\ 69 & 75 \end{pmatrix}$.

Note that AB is defined *only* when the number of columns of A equals the number of rows of B. An important remark: even when both AB and BA are defined, the results are usually different. A property of matrix multiplication is the following:

$$(AB)C = A(BC).$$

That is, if you have three matrices A, B, C to multiply and the product is legal (the number of columns of A equals the number of rows of B and the number of columns of B equals the number of rows of C), then you have two possibilities: you can first compute AB and multiply the result by C, or you can first compute BC and multiply A by the result.

Exercise 4 Consider the following matrices

$$x = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix} \quad y = \begin{pmatrix} 3 & 2 & 1 \end{pmatrix} \quad z = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

When possible, compute the following quantities:

(a) xz

(b) zx

(c) yzx

(d) xzy

(e) $(x + y)z$

(f) $(xz) + (yz)$.

Remark: A system of linear equations can be written conveniently using matrix notation. Namely,

$$\begin{array}{ccccccc} a_{11}x_1 & + & \dots & + & a_{1n}x_n & = & b_1 \\ \dots & & \dots & & \dots & & \\ a_{m1}x_1 & + & \dots & + & a_{mn}x_n & = & b_m \end{array}$$

can be written as

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$$

or as

$$\begin{pmatrix} a_{11} \\ \vdots \\ a_{m1} \end{pmatrix} x_1 + \dots + \begin{pmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{pmatrix} x_n = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}.$$

So a matrix equation $Ax = b$ where A is a given $m \times n$ matrix, b is a given m-column vector and x is an unknown n-column vector, is a linear system of m equations in n variables. Similarly, a vector equation $a_1x_1 + \dots + a_nx_n = b$ where $a_1, \dots, a_n, b$ are given m-column vectors and $x_1, \dots, x_n$ are n unknown real numbers, is also a system of m equations in n variables.

Exercise 5 (a) Solve the matrix equation

$$\begin{pmatrix} 1 & 2 & 1 \\ 2 & -1 & 3 \\ 1 & 1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 4 \\ 3 \\ 3 \end{pmatrix}.$$

(b) Solve the vector equation

$$\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} x_1 + \begin{pmatrix} 2 \\ -1 \\ 1 \end{pmatrix} x_2 + \begin{pmatrix} 1 \\ 3 \\ -1 \end{pmatrix} x_3 = \begin{pmatrix} 4 \\ 3 \\ 3 \end{pmatrix}.$$

[Hint: Use Example 1.1.1]

The following standard definitions will be useful:

A *square* matrix is a matrix with same number of rows as columns.

The *identity matrix* I is a square matrix with 1's on the main diagonal and 0's elsewhere, i.e.
$$I = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

The *zero matrix* contains only 0's. In particular, the *zero vector* has all components equal to 0. A *nonzero vector* is one that contains at least one nonzero component.

The *transpose* of matrix $A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$ is the matrix $A^T = \begin{pmatrix} a_{11} & \dots & a_{m1} \\ a_{12} & \dots & a_{m2} \\ \vdots & & \vdots \\ a_{1n} & \dots & a_{mn} \end{pmatrix}$.

A square matrix A is *symmetric* if its elements satisfy $a_{ij} = a_{ji}$ for all i and j. So a symmetric matrix satisfies $A^T = A$.

1.3 Linear Combinations

Suppose that the vector $(1,3)$ represents the contents of a "Regular" can of mixed nuts (1 lb cashews and 3 lb peanuts) while $(1,1)$ represents a "Deluxe" can (1 lb cashews and 1 lb peanuts). Can you obtain a mixture of 2 lb cashews and 3 lb peanuts by combining the two mixtures in appropriate amounts? The answer is to use x_1 cans of Regular and x_2 cans of Deluxe in order to satisfy the equality

$$x_1 \begin{pmatrix} 1 \\ 3 \end{pmatrix} + x_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}.$$

This is none other than a system of two linear equations:

$$\begin{aligned} x_1 + x_2 &= 2 \\ 3x_1 + x_2 &= 3 \end{aligned}$$

The solution of these equations (obtained by the Gauss-Jordan procedure) is $x_1 = 1/2$, $x_2 = 3/2$. So the desired combination is to mix 1/2 can of Regular nuts with 3/2 cans of Deluxe nuts. Thus if some recipe calls for the mixture $(2,3)$, you can substitute 1/2 can of Regular mix and 3/2 can of Deluxe mix.

Suppose now that you want to obtain 1 lb cashews and no peanuts. This poses the equations,

$$x_1 \begin{pmatrix} 1 \\ 3 \end{pmatrix} + x_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

The solution is $(x_1, x_2) = (-1/2, 3/2)$. Thus you can obtain a pound of pure cashews by buying 3/2 cans of Deluxe mix and removing enough nuts to form 1/2 can Regular mix, which can be sold. In this case it is physically possible to use a negative amount of some ingredient, but in other cases it may be impossible, as when one is mixing paint.

A vector of the form $x_1 \begin{pmatrix} 1 \\ 3 \end{pmatrix} + x_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ is called a *linear combination* of the vectors $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. In particular we just saw that the vector $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$ is a linear combination of the vectors $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$. And so is $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

The question arises: can one obtain any mixture whatever by taking the appropriate combination of Regular and Deluxe cans? Is any vector $\begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ a linear combination of $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$ and $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$?

To answer the question, solve the equations in general. You want a mixture of b_1 lb cashews and b_2 lb peanuts and set

$$x_1 \begin{pmatrix} 1 \\ 3 \end{pmatrix} + x_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix},$$

or equivalently,

$$\begin{aligned} x_1 + x_2 &= b_1 \\ 3x_1 + x_2 &= b_2. \end{aligned}$$

These equations have a unique solution, no matter what are the values of b_1 and b_2, namely,

$$x_1 = \frac{b_2 - b_1}{2}, \quad x_2 = \frac{3b_1 - b_2}{2}.$$

No matter what vector (b_1, b_2) you want, you can get it as a linear combination of $(1,3)$ and $(1,1)$. The vectors $(1,3)$ and $(1,1)$ are said to be *linearly independent.*

Not all pairs of vectors can yield an arbitrary mixture (b_1, b_2). For instance, no linear combination of $(1,1)$ and $(2,2)$ yields (2,3). In other words, the equations

$$x_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + x_2 \begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

have no solution.

The reason is that $(2,2)$ is already a multiple of $(1,1)$, that is $(2,2)$ is a linear combination of $(1,1)$. The vectors $(1,1)$ and $(2,2)$ are said to be *linearly dependent.*

If (2,2) represents, for instance, a large Deluxe can of nuts and (1,1) a small one, it is clear that you cannot obtain any mixture you want by combining large and small Deluxe cans. In fact, once you have small cans, the large cans contribute nothing at all to the mixtures you can generate, since you can always substitute two small cans for a large one.

A more interesting example is (1,2,0), (1,0,1) and (2,2,1). The third vector is clearly a linear combination of the other two (it equals their sum). Altogether, these three vectors are said to be linearly dependent. For instance, these three vectors might represent mixtures of nuts as follows:

Brand	A	B	C
cashews	1	1	2
peanuts	2	0	2
almonds	0	1	1

Then once you have brands A and B, brand C adds nothing whatever to the variety of mixtures you can concoct. This is because you can already obtain a can of brand C from brands A and B anyway. In other words, if a recipe calls for brand C, you can always substitute a mixture of 1 can brand A and 1 can brand B.

Suppose you want to obtain the mixture $(1,2,1)$ by combining Brands A, B, and C. The equations are

$$\begin{array}{rcrcrcl} x_1 & + & x_2 & + & 2x_3 & = & 1 \\ 2x_1 & & & + & 2x_3 & = & 2 \\ & & x_2 & + & x_3 & = & 1 \end{array}$$

If you try to solve these equations, you will find that there is no solution. So the vector $(1,2,1)$ is **not** a linear combination of $(1,2,0)$, $(1,0,1)$ and $(2,2,1)$.

The concepts of linear combination, linear dependence and linear independence introduced in the above examples can be defined more formally, for any number of n-component vectors. This is done as follows.

- A vector b having n components is a *linear combination* of the k vectors $v_1, \ldots, v_k$, each having n components, if it is possible to find k real numbers $x_1, \ldots, x_k$ satisfying:

$$x_1 v_1 + \ldots x_k v_k = b \tag{1.1}$$

  To find the numbers $x_1, \ldots, x_k$, view (1.1) as a system of linear equations and solve by the Gauss-Jordan method.

- A set of vectors (all having n components) is *linearly dependent* if at least one vector is a linear combination of the others. Otherwise they are *linearly independent.*

- Given n linearly independent vectors $v_1, \ldots, v_n$, each having n components, any desired vector b with n components can be obtained as a linear combination of them:

$$x_1 v_1 + \ldots x_n v_n = b \tag{1.2}$$

  The desired weights $x_1, \ldots, x_n$ are computed by solving (1.2) with the Gauss-Jordan method: there is a unique solution whenever the vectors $v_1, \ldots, v_n$ are linearly independent.

Exercise 6 A can of Brand A mixed nuts has 1 lb cashews, 1 lb almonds, 2 lb peanuts. Brand B has 1 lb cashews and 3 lb peanuts. Brand C has 1 lb almonds and 2 lb peanuts. Show how much of each brand to buy/sell so as to obtain a mixture containing 5 lb of each type of nut.

Exercise 7 Determine whether the vector $\begin{pmatrix} 1 \\ 2 \\ -2 \end{pmatrix}$ is a linear combination of

(a) $\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$

(b) $\begin{pmatrix} 2 \\ 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ -2 \end{pmatrix}.$

In linear programming, there are typically many more variables than equations. So, this case warrants looking at another example.

Example 1.3.1 (Basic Solutions.) A craft shop makes deluxe and regular belts. Each deluxe belt requires a strip of leather and 2 hours of labor. Each regular belt requires a leather strip and 1 hour of labor. 40 leather strips and 60 hours of labor are available. How many belts of either kind can be made?

This is really a mixing problem. Rather than peanuts and cashews, the mixture will contain leather and labor. The items to be mixed are four *activities*: manufacturing a deluxe belt, manufacturing a regular belt, leaving a leftover leather strip in inventory, and leaving an hour of labor idle (or for other work). Just as each Regular can of mixed nuts contributes 1 pound of cashews and 3 pounds of peanuts to the mixture, each regular belt will consume 1 leather strip and 2 hours of labor. The aim is to combine the four activities in the right proportion so that 40 strips and 60 hours are accounted for:

$$x_1 \begin{pmatrix} 1 \\ 2 \end{pmatrix} + x_2 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + s_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + s_2 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 40 \\ 60 \end{pmatrix}.$$

So,

x_1 = number of deluxe belts made
x_2 = number of regular belts made
s_1 = number of leather strips left over
s_2 = number of labor hours left over

In tableau form, the equations are:

x_1	x_2	s_1	s_2	RHS
1	1	1	0	40
2	1	0	1	60

(1.3)

Since there are only two equations, you can only solve for two variables. Let's solve for x_1 and x_2, using Gauss-Jordan elimination. After the first iteration you get,

x_1	x_2	s_1	s_2	RHS
1	1	1	0	40
0	−1	−2	1	−20

(1.4)

A second iteration yields the solution,

x_1	x_2	s_1	s_2	RHS
1	0	−1	1	20
0	1	2	−1	20

(1.5)

This tableau represents the equations

$$\begin{aligned} x_1 \qquad - s_1 + s_2 &= 20 \\ x_2 + 2s_1 - s_2 &= 20 \end{aligned}$$

or,

$$\begin{aligned} x_1 &= 20 + s_1 - s_2 \\ x_2 &= 20 - 2s_1 + s_2 \end{aligned} \qquad (1.6)$$

You can't say how many deluxe belts x_1 and regular belts x_2 the plant will make until you specify how much leather s_1 and labor s_2 will be left over. But (1.6) is a formula for computing how many belts of either type must be made, for any given s_1 and s_2. So the equations have many solutions (infinitely many).

For example, if you want to have nothing left over ($s_1 = s_2 = 0$), you will make 20 of each. If you want to have 5 strips and 5 hours left over, you will make $x_1 = 20$ deluxe and $x_2 = 15$ regular belts.

The variables x_1, x_2 you solved for are called *basic variables*. The other variables are *nonbasic*. You have control of the nonbasic variables. Once you give them values, the values of the basic variables follow. A solution in which you make all the nonbasic variables zero is called a *basic solution.*

Can you have basic variables other than x_1, x_2? Sure. Any pair of variables can be basic, provided the corresponding columns in (1.3) are linearly independent (otherwise, you can't solve for the basic variables).

Equations (1.3), for instance, are already solved in (1.3) for basic variables s_1, s_2. Here the two basic activities are having leftover leather and having leftover labor. The basic solution is $(s_1, s_2) = (40, 60)$. This means that if you decide to produce no belts ($x_1 = x_2 = 0$), you must have 40 leftover leather strips and 60 leftover labor hours.

The intermediate step (1.4) solves the equations with basic variables x_1 and s_2. Here the basic solution is unrealizable. If you decide to participate only in the basic activities (making deluxe belts and having leftover labor), you must make 40 belts and have -20 leftover hours (i.e., use 20 more than you have), which you can't do within your current resources.

Exercise 8 Consider the system of equations

x_1	x_2	x_3	x_4	RHS
1	1	2	4	100
3	1	1	2	200

where the first four columns on the left represent a Regular mixture of nuts, a Deluxe mixture, a small can of Premium mixture, and a large can of Premium.

a) Solve the system with x_1 and x_2 basic.

b) You want 100 cans of mixed nuts, each of which contains 1 lb cashews and 2 lb peanuts (i.e., you want the right-hand side of the above equation). How can you get them by blending 10 small cans of Premium with proper amounts of the Regular and Deluxe blends? *Hint.* Set $x_3 = 10$ and $x_4 = 0$ in the expression for the solution values of x_1, x_2 found in (a).

c) How can you obtain a small can of Premium mix by combining (and decombining) the Regular and Deluxe blends?

d) How can you obtain one can of Regular mix by combining large and small cans of Premium? If you cannot do it, why not? *Hint.* It has to do with linear dependence.

Exercise 9 A can of paint A has 1 quart red, 1 quart yellow. A can of paint B has 1 quart red, 1 quart blue. A can of paint C has 1 quart yellow, 1 quart blue.

a) How much of each paint must be mixed to obtain a mixture of 1 quart red, 1 quart yellow and 1 quart blue?

b) How much of each paint must be mixed to obtain one quart of pure red? What do you conclude about the physical feasibility of such a mixture?

Exercise 10 An eletronics plant wants to make stereos and CB's. Each stereo requires 1 power supply and 3 speakers. Each CB requires 1 power supply and 1 speaker. 100 power supplies and 200 speakers are in stock. How many stereos and CB's can it make if it wants to use all the power supplies and all but 10 of the speakers? *Hint.* Use the following equations.

x_1	x_2	s_1	s_2	RHS
1	1	1	0	100
3	1	0	1	200

Exercise 11 A construction foreman needs cement mix containing 8 cubic yards (yd^3) cement, 12 yd^3 sand and 16 yd^3 water. On the site are several mixer trucks containing mix A (1 yd^3 cement, 3 yd^3 sand, 3 yd^3 water), several containing mix B (2 yd^3 cement, 2 yd^3 sand, 3 yd^3 water) and several containing mix C (2 yd^3 cement, 2 yd^3 sand, 5 yd^3 water).

How many truckloads of each mix should the foreman combine to obtain the desired blend?

1.4 Inverse of a Square Matrix

If A and B are square matrices such that $AB = I$ (the identity matrix), then B is called the *inverse* of A and is denoted by A^{-1}. A square matrix A has either no inverse or a unique inverse A^{-1}. In the first case, it is said to be *singular* and in the second case *nonsingular*. Interestingly, linear independence of vectors plays a role here: a matrix is singular if its columns form a set of linearly dependent vectors; and it is nonsingular if its columns are linearly independent. Another property is the following: if B is the inverse of A, then A is the inverse of B.

Exercise 12 (a) Compute the matrix product $\begin{pmatrix} 1 & 3 & -2 \\ 0 & -5 & 4 \\ 2 & -3 & 3 \end{pmatrix} \begin{pmatrix} -3 & -3 & 2 \\ 8 & 7 & -4 \\ 10 & 9 & -5 \end{pmatrix}$.

(b) What is the inverse of $\begin{pmatrix} -3 & -3 & 2 \\ 8 & 7 & -4 \\ 10 & 9 & -5 \end{pmatrix}$?

(c) Show that the matrix $A = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}$ is singular.

[Hint: Assume that the inverse of A is $B = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}$ and perform the matrix product AB. Then show that no choice of b_{ij} can make this product equal to the identity matrix $I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.]

An important property of nonsingular square matrices is the following. Consider the system of linear equations

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}, \text{ simply written as } Ax = b.$$

When A is a square nonsingular matrix, this linear system has a unique solution, which can be obtained as follows. Multiply the matrix equation $Ax = b$ by A^{-1} on the left:

$$A^{-1}Ax = A^{-1}b.$$

This yields $Ix = A^{-1}b$ and so, the unique solution to the system of linear equations is

$$x = A^{-1}b.$$

Exercise 13 Solve the system

$$\begin{array}{rcrcrcr} -3x_1 & - & 3x_2 & + & 2x_3 & = & 1 \\ 8x_1 & + & 7x_2 & - & 4x_3 & = & -1 \\ 10x_1 & + & 9x_2 & - & 5x_3 & = & 1 \end{array}$$

using the result of Exercise 12(b).

Finding the Inverse of a Square Matrix

Given $A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$, we must find $B = \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \dots & b_{nn} \end{pmatrix}$ such that $AB = I$ (the identity matrix). Therefore, the first column of B must satisfy $A \begin{pmatrix} b_{11} \\ \vdots \\ b_{n1} \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ (this vector is the 1st column of I). Similarly, for the other columns of B. For example, the jth column of B satisfies $A \begin{pmatrix} b_{ij} \\ \vdots \\ b_{nj} \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ (the jth column of I). So in order to get the inverse of an $n \times n$ matrix, we must solve n linear systems. However, the *same* steps of the Gauss-Jordan elimination procedure are needed for all of these systems. So we solve them all at once, using the matrix form.

Example: Find the inverse of $A = \begin{pmatrix} 3 & -2 \\ -4 & 3 \end{pmatrix}$.

We need to solve the following matrix equation

$$\begin{pmatrix} 3 & -2 \\ -4 & 3 \end{pmatrix} B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

We divide the first row by 3 to introduce a 1 in the top left corner.

$$\begin{pmatrix} 1 & -\frac{2}{3} \\ -4 & 3 \end{pmatrix} B = \begin{pmatrix} \frac{1}{3} & 0 \\ 0 & 1 \end{pmatrix}$$

Then we add four times the first row to the second row to introduce a 0 in the first column.

$$\begin{pmatrix} 1 & -\frac{2}{3} \\ 0 & \frac{1}{3} \end{pmatrix} B = \begin{pmatrix} \frac{1}{3} & 0 \\ \frac{4}{3} & 1 \end{pmatrix}$$

Multiply the second row by 3.

$$\begin{pmatrix} 1 & -\frac{2}{3} \\ 0 & 1 \end{pmatrix} B = \begin{pmatrix} \frac{1}{3} & 0 \\ 4 & 3 \end{pmatrix}$$

Add $\frac{2}{3}$ the second row to the first. (All this is classical Gauss-Jordan elimination.)

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} B = \begin{pmatrix} 3 & 2 \\ 4 & 3 \end{pmatrix}$$

As $IB = B$, we get

$$B = \begin{pmatrix} 3 & 2 \\ 4 & 3 \end{pmatrix}.$$

It is important to note that, in addition to the two elementary row operations introduced earlier in the context of the Gauss-Jordan elimination procedure, a third elementary row operation may sometimes be needed here, namely permuting two rows.

Example: Find the inverse of $A = \begin{pmatrix} 0 & 1 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

$$\begin{pmatrix} 0 & 1 & 0 \\ 2 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Because the top left entry of A is 0, we need to permute rows 1 and 2 first.

$$\begin{pmatrix} 2 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} B = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Now we divide the first row by 2.

$$\begin{pmatrix} 1 & \frac{1}{2} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} B = \begin{pmatrix} 0 & \frac{1}{2} & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Next we add $-\frac{1}{2}$ the second row to the first.

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} B = \begin{pmatrix} -\frac{1}{2} & \frac{1}{2} & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and we are done, since the matrix in front of B is the identity.

Exercise 14 Find the inverse of the matrix

$$A = \begin{pmatrix} 1 & 4 & 2 \\ 0 & 1 & 0 \\ -1 & -4 & 2 \end{pmatrix}$$

.

Exercise 15 Find the inverse of the matrix

$$A = \begin{pmatrix} 0 & 1 & -2 \\ 4 & 0 & 5 \\ 3 & 0 & 4 \end{pmatrix}$$

.

1.5 Determinants

To each square matrix, we associate a number, called its *determinant*, defined as follows:

If $A = (a_{11})$, then $\det(A) = a_{11}$,

If $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, then $\det(A) = a_{11}a_{22} + a_{12}a_{21}$.

For a square matrix A of dimensions $n \times n$, the determinant can be obtained as follows. First, define A_{1j} as the matrix of dimensions $(n-1) \times (n-1)$ obtained from A by deleting row 1 and column j. Then

$$\det(A) = a_{11}\det(A_{11}) - a_{12}\det(A_{12}) + a_{13}\det(A_{13}) - \ldots a_{1n}\det(A_{1n}).$$

Note that, in this formula, the signs alternate between + and −.

For example, if $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$, then

$$\det(A) = a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}).$$

Determinants have several interesting properties. For example, the following statements are equivalent for a square matrix A:

- $\det(A) = 0$,
- A has no inverse, i.e. A is singular,
- the columns of A form a set of linearly dependent vectors,
- the rows of A form a set of linearly dependent vectors.

For our purpose, however, determinants will be needed mainly in our discussion of classical optimization, in conjunction with the material from the following section.

Exercise 16 Compute the determinant of $A = \begin{pmatrix} 4 & 2 & 3 \\ 1 & 0 & 1 \\ 0 & 1 & 2 \end{pmatrix}$.

1.6 Positive Definite Matrices

When we study functions of several variables (see Chapter 3!), we will need the following matrix notions.

A square matrix A is *positive definite* if $x^T Ax > 0$ for all nonzero column vectors x. It is *negative definite* if $x^T Ax < 0$ for all nonzero x. It is *positive semidefinite* if $x^T Ax \geq 0$ and *negative semidefinite* if $x^T Ax \leq 0$ for all x. These definitions are hard to check directly and you might as well forget them for all practical purposes.

More useful in practice are the following properties, which hold when the matrix A is symmetric (that will be the case of interest to us), and which are easier to check.

The ith *principal minor* of A is the matrix A_i formed by the first i rows and columns of A. So, the first principal minor of A is the matrix $A_1 = (a_{11})$, the second principal minor is the matrix $A_2 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, and so on.

- The matrix A is positive definite if all its principal minors A_1, $A_2, \ldots A_n$ have strictly positive determinants.
- If these determinants are nonzero and alternate in signs, starting with $\det(A_1) < 0$, then the matrix A is negative definite.
- If the determinants are all nonnegative, then the matrix is positive semidefinite,
- If the determinant alternate in signs, starting with $\det(A_1) \leq 0$, then the matrix is negative semidefinite.

To fix ideas, consider a 2×2 symmetic matrix $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix}$.

It is positive definite if:

(i) $\det(A_1) = a_{11} > 0$

(ii) $\det(A_2) = a_{11}a_{22} - a_{12}a_{12} > 0$

and negative definite if:

(i) $\det(A_1) = a_{11} < 0$

(ii) $\det(A_2) = a_{11}a_{22} - a_{12}a_{12} > 0$.

It is positive semidefinite if:

(i) $\det(A_1) = a_{11} \geq 0$

(ii) $\det(A_2) = a_{11}a_{22} - a_{12}a_{12} \geq 0$

and negative semidefinite if:

(i) $\det(A_1) = a_{11} \leq 0$

(ii) $\det(A_2) = a_{11}a_{22} - a_{12}a_{12} \geq 0$.

Exercise 17 Check whether the following matrices are positive definite, negative definite, positive semidefinite, negative semidefinite or none of the above.

(a) $A = \begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix}$

(b) $A = \begin{pmatrix} -2 & 4 \\ 4 & -8 \end{pmatrix}$

(c) $A = \begin{pmatrix} -2 & 2 \\ 2 & -4 \end{pmatrix}$

(d) $A = \begin{pmatrix} 2 & 4 \\ 4 & 3 \end{pmatrix}$.

Chapter 2

Unconstrained Optimization: Functions of One Variable

This chapter reviews basic results about functions of one variable, including the notions of derivative, extremum and convexity. It also describes a numerical method for finding x such that $f(x) = 0$, known as "binary search".

2.1 Derivatives

It is important to know the following differentiation formulas:

Function	Derivative
$f(x) + g(x)$	$f'(x) + g'(x)$
$af(x)$	$af'(x)$
$f(x)g(x)$	$f'(x)g(x) + f(x)g'(x)$
$f(x)/g(x)$	$[g(x)f'(x) - f(x)g'(x)]/g^2(x)$
$f(g(x))$	$f'(g(x))g'(x)$
x^a	ax^{a-1}
e^x	e^x
$\ln x \quad (x > 0)$ where $\ln x = \log_e x$	$1/x$

Exercise 18 Compute the derivatives of the following functions:

(a) $x^3 - 3x + 1$

(b) $\ln(1 + x^2)$

(c) $\ln(1 + x)^2$

(d) $(3 - 2x)e^{x^2}$

For a function f of one variable x, recall that the derivative $f'(x)$ is equal to the slope of a tangent line at point x. So, if the function has a positive derivative at point x, then the function is increasing, and if it has a negative derivative, it is decreasing. Since the function and its tangent line are close around point x, the following formula can be used when Δ is small.

$$f(x+\Delta) \approx f(x) + \Delta f'(x)$$

Example 2.1.1 *Suppose that the demand for gasoline at price x (in dollars) is*

$$ae^{-0.2x}$$

gallons, where a is a constant. Suppose the price of gas rises by 10 cents. By how much does demand fall?

Let $f(x) = ae^{-0.2x}$ denote the demand for gas at price x. The rate of change is given by the derivative

$$f'(x) = -0.2f(x).$$

Since $\Delta = 0.10$, we get

$$f(x+0.10) \approx f(x) + 0.10(-0.2f(x)) = 0.98f(x).$$

So demand drops by 2%. The factor relating change in demand to change in price is known as "price elasticity of demand" in economics (You will learn more about this in 45-749 Managerial Economics and in marketing courses such as 45-720 Marketing Management and 45-834 Pricing). Here $f'(x) = -0.2f(x)$, so price elasticity of demand is -0.2.

Exercise 19 Suppose that, if x dollars are spent on advertising for a product during a given year, $f(x) = k(1 - e^{-cx})$ customers will purchase the product ($k > 0$ and $c > 0$ are two constants).

(a) As x grows large, the number of customers purchasing the product approaches a limit. Find this limit. Can you give an interpretation for the constant k?

(b) The *sales response* from a dollar of advertising is $f(x+1) - f(x)$. Using the formula for small changes, show that the sales response from a dollar of advertising is proportional to the number of potential customers who are not purchasing the product at present.

2.2 Maximum and Minimum

Let f be a function of one variable defined for all x in some domain D. A *global maximum* of f is a point x_0 in D such that $f(x) \leq f(x_0)$ for all x in D.

For a constant $\Delta > 0$, the *neighborhood* $N_\Delta(x_0)$ of a point x_0 is the set of all points x such that $x_0 - \Delta < x < x_0 + \Delta$. A point x_0 is a *local maximum* of f if there exists $\Delta > 0$ such that $f(x) \leq f(x_0)$ for all x in $N_\Delta(x_0)$ where $f(x)$ is defined.

Similarly one can define local and global minima. In Figure 2.1, the function f is defined for x in domain $[a, e]$. Points b and d are local maxima and b is a global maximum, whereas a, c, and e are local minima and e is a global minimum.

Finding extrema

Extrema, whether they are local or global, can occur in three places:

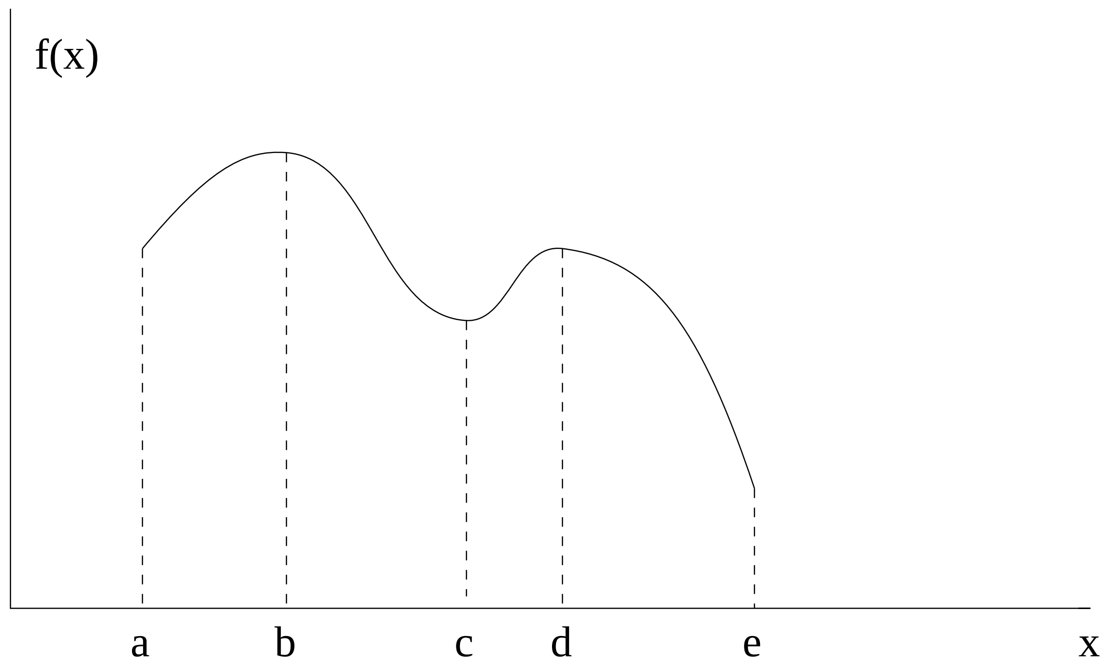


Figure 2.1: local maxima and minima

1. at the boundary of the domain,
2. at a point without a derivative, or
3. at a point x_0 with $f'(x_0) = 0$.

The last case is particularly important. So we discuss it further. Let f be differentiable in a neighborhood of x_0. If x_0 is a local extremum of f, then $f'(x_0) = 0$. Conversely, if $f'(x_0) = 0$, three possibilities may arise: x_0 is a local maximum, x_0 is a local minimum, or neither!!! To decide between these three possibilities, one may use the second derivative. Let f be twice differentiable in a neighborhood of x_0.

- If $f'(x_0) = 0$ and $f''(x_0) > 0$ then x_0 is a local minimum.
- If $f'(x_0) = 0$ and $f''(x_0) < 0$ then x_0 is a local maximum.
- If $f'(x_0) = 0$ and $f''(x_0) = 0$ then x_0 may or may not be a local extremum.

Figure 2.2 illustrates these three possibilities.

Example 2.2.1 *Suppose an oil cartel can price gasoline as it wishes. How will it maximize revenue, given that demand for gasoline at price x is*

$$f(x) = ae^{-0.2x}$$

The revenue at price x is

$$g(x) = xf(x).$$

We compute the derivative of g and set it to 0.

$$g'(x) = f(x) + x(-0.2f(x)) = (1 - 0.2x)f(x)$$

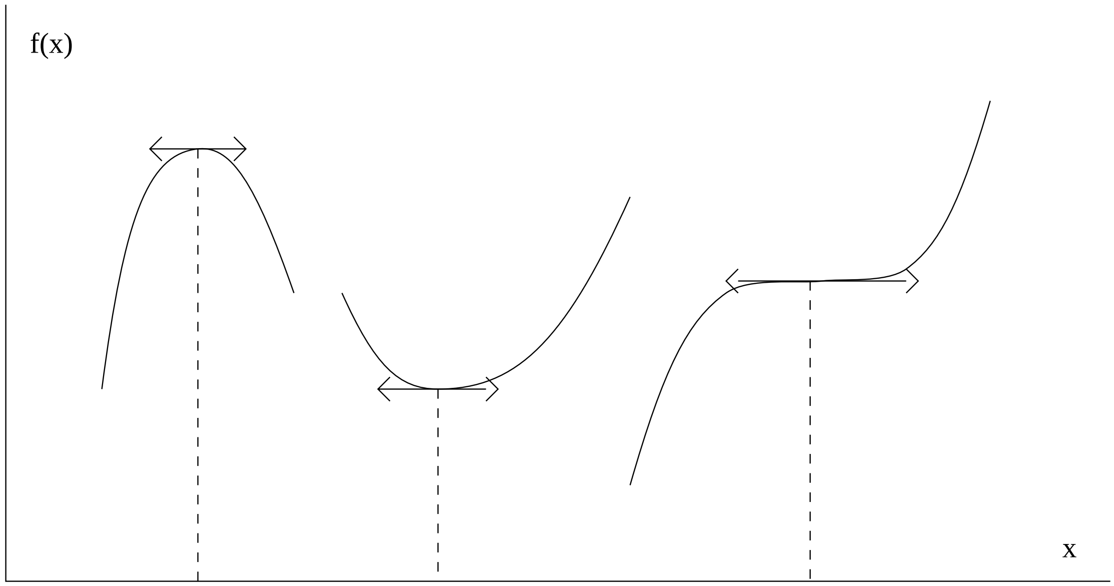


Figure 2.2: $f''(x_0) > 0$, $f''(x_0) < 0$ and one of the possibilities with $f''(x_0) = 0$

Since $f(x) > 0$ for all x, setting $g'(x) = 0$ implies $1 - 0.2x = 0$. So $x = 5$. This is the only possible local optimum. To determine whether it is a maximum or a minimum, we compute the second derivative.

$$g''(x) = 2(-0.2f(x)) + x(-0.2)^2 f(x) = -0.2(2 - 0.2x)f(x).$$

Putting in $x = 5$ shows $g''(x) < 0$, so this is a maximum: the oil cartel maximizes its revenue by pricing gasoline at \$5 per gallon.

Example 2.2.2 Economic Order Quantity *We want to keep paper towels on hand at all times. The cost of stockpiling towels is \$ h per case per week, including the cost of space, the cost of tying up capital in inventory, etc. The time and paperwork involved in ordering towels costs \$ K per order (this is in addition to the cost of the towels themselves, which is not relevant here because we are supposing that there are no price discount for large orders.) Towels are used at a rate of d cases per week. What quantity Q should we order at a time, to minimize cost?*

We must first calculate the holding cost (cost of stockpiling). If each order period begins with Q cases and ends with zero cases, and if usage is more or less constant, then the average stock level is $Q/2$. This means that the average holding cost is $hQ/2$.

Since each order of Q cases lasts Q/d weeks, the average ordering cost *per week* is $\frac{K}{Q/d} = \frac{Kd}{Q}$. Thus the average total cost per week is

$$\frac{hQ}{2} + \frac{Kd}{Q}.$$

To find Q that minimizes cost, we set to zero the derivative with respect to Q.

$$\frac{h}{2} - \frac{Kd}{Q^2} = 0.$$

This implies that the optimal order quantity is

$$Q = \sqrt{\frac{2Kd}{h}}.$$

This is the classical *economic order quantity (EOQ)* model for inventory management. (You will learn when the EOQ model is appropriate and when to use other inventory models in 45-765).

Exercise 20 Find the local maxima and minima, if any, of the following functions.

(a) $x^3 - 3x + 1$

(b) $x^3 - 3x^2 + 3x - 1$

(c) $(3 - 2x)e^{x^2}$

Exercise 21 We want to design a cone-shaped paper drinking cup of volume V that requires a minimum of paper to make. That is, we want to minimize its surface area $\pi r\sqrt{r^2 + h^2}$, where r is the radius of the open end and h the height. Since $V = \frac{\pi}{3}r^2 h$, we can solve for h to get $h = 3V/(\pi r^2)$. Thus the area in terms of r is

$$A(r) = \pi r\left(r^2 + 9\frac{V^2}{\pi^2 r^4}\right)^{1/2} = \left(\pi^2 r^4 + 9\frac{V^2}{r^2}\right)^{1/2}.$$

a) Compute the derivative $A'(r)$.

b) Find the radius r at which $A(r)$ is minimized (the solution you get is in fact a minimum).

c) What is the ratio h/r when r has the optimal value?

Exercise 22 An air courier company operates a single depot in a large city and wants to decentralize by establishing several depots, each covering the same delivery radius r. The total area to be covered is A, and the area covered by each depot is πr^2. Thus if n depots are installed, we must have $A = n\pi r^2$, so that $n = A/(\pi r^2)$. Expenses and revenues break down as follows.

- Trucking expenses. The number of trucks required at each depot is $\alpha\pi r^3$, so that the total number of trucks required is $n\alpha\pi r^3 = (A/\pi r^2)\alpha\pi r^3 = A\alpha r$. Each truck costs \$$K$ per day, so that total expense for trucks is \$$KA\alpha r$.

- Overhead. Each depot incurs an overhead (fixed) expense of \$$F$, for a total overhead of $nF = FA/(\pi r^2)$.

- Revenue from customers. Since the total area is fixed, the total revenue from customers is independent of the delivery radius.

Since the total revenue is fixed, we can simply minimize costs:

$$\min KA\alpha r + \frac{FA}{\pi r^2}.$$

Your problem. Find the delivery radius r that minimizes cost. Your solution should indicate that the optimal delivery radius is proportional to the cube root (the 1/3 power) of the fixed cost. Thus if the fixed cost increases eightfold (and other costs were constant), the optimal delivery radius would only double. This is fortunate, since it means that the rent and other fixed costs must rise substantially (relative to other costs) before they justify relocating depots.

2.3 Binary Search

Binary search is a very simple idea for solving numerically $f(x) = 0$.

It requires:

(i) the function f to be continuous,

(ii) values (a, b) such that $f(a) > 0$ and $f(b) < 0$.

Then a desired value x_0 such that $f(x_0) = 0$ can be found between the values a and b.

The method consists in computing $f(c)$ for $c = \frac{a+b}{2}$.

- If $f(c) > 0$, the procedure is repeated with the values (a, b) replaced by the values (c, b).
- If $f(c) < 0$, the procedure is repeated with the values (a, b) replaced by (a, c).
- If $f(c) = 0$, the procedure terminates as c is the value we are looking for.

This method is easy to program and converges rapidly (20 iterations will give you x_0 with 5 significant figures).

Example 2.3.1 *Binary search can be used to compute the internal rate of return of an investment. The* internal rate of return *of an investment is the interest rate r that makes the present value of the cash flows from the investment equal to the cost of the investment. Mathematically, r is the interest rate that satisfies the equation*

$$\frac{F_1}{1+r} + \frac{F_2}{(1+r)^2} + \frac{F_3}{(1+r)^3} + \ldots + \frac{F_N}{(1+r)^N} - C = 0$$

where

$$\begin{aligned} F_t &= \textit{cash flow in year } t \\ N &= \textit{number of years} \\ C &= \textit{cost of the investment} \end{aligned}$$

For a bond, the internal rate of return r is known as the yield to maturity *or simply* yield. *Let us consider the case of a noncallable bond, i.e. a bond that the issuer cannot retire prior to its stated maturity date. For such a bond, the cash flows consist of periodic interest payments to the maturity date and the par value paid at maturity. As an example, consider a 4-year noncallable bond with a 10% coupon rate paid annually and a par value of $1000. Such a bond has the following cash flows:*

t Years from now	F_t
1	*$ 100*
2	*100*
3	*100*
4	*1100*

Suppose this bond is now selling for $900. Compute the yield of this bond.

The yield r of the bond is given by the equation

$$\frac{100}{1+r}+\frac{100}{(1+r)^2}+\frac{100}{(1+r)^3}+\frac{1100}{(1+r)^4}-900=0$$

Let us denote by $f(r)$ the left-hand-side of this equation. We find r such that $f(r)=0$ using binary search.

We start by finding values (a,b) such that $f(a)>0$ and $f(b)<0$. In this case, we expect r to be between 0 and 1. Since $f(0)=500$ and $f(1)=-743.75$, we have our starting values.

Next, we let $c=0.5$ (the midpoint) and calculate $f(c)$. Since $f(0.5)=-541.975$, we replace our range with $a=0$ and $b=0.5$ and repeat. When we continue, we get the following table of values:

Iter.	a	c	b	$f(a)$	$f(c)$	$f(b)$
1	0	0.5	1	500	-541.975	-743.75
2	0	0.25	0.5	500	-254.24	-541.975
3	0	0.125	0.25	500	24.85902	-254.24
4	0.125	0.1875	0.25	24.85902	-131.989	-254.24
5	0.125	0.15625	0.1875	24.85902	-58.5833	-131.989
6	0.125	0.140625	0.15625	24.85902	-18.2181	-58.5833
7	0.125	0.132813	0.140625	24.85902	2.967767	-18.2181
8	0.132813	0.136719	0.140625	2.967767	-7.71156	-18.2181
9	0.132813	0.134766	0.136719	2.967767	-2.39372	-7.71156
10	0.132813	0.133789	0.134766	2.967767	0.281543	-2.39372
11	0.133789	0.134277	0.134766	0.281543	-1.05745	-2.39372
12	0.133789	0.134033	0.134277	0.281543	-0.3883	-1.05745

So the yield of the bond is $r=13.4\%$.

Of course, this routine sort of calculation is perfectly set up for calculation by computer. In particular, we can use Microsoft's Excel (or any spreadsheet program) to do these calculations. On the course home page is a spreadsheet that implements binary search for this problem. We will show this spreadsheet in class. Cells B3:G3 down to B20:G20 contain the body of the table. You can generate the same table by following these steps:

In cell B3, type 0 [comment: this is initial value of a.]

In cell D3, type 1 [comment: this is the initial value of b.]

In cell C3, type =(B3+D3)/2 [comment: this is the middle point c.]

In cell E3, type =100/(1+B3) +100/(1+B3)∧2 +100/(1+B3)∧3 +1100/(1+B3)∧4 -900 [comment: this is $f(a)$.]

Use the Copy Command to copy cell E3 into cells F3 and G3 [comment: cell F3 contains the same function as cell E3, except that B3 is replaced by C3. So cell F3 contains $f(c)$. Similarly, G3 contains $f(b)$.]

In cell B4, type =IF(F3> 0,C3,B3) [comment: If F3> 0, cell B4 contains C3 and otherwise, cell B4 contains B3.]

In cell D4, type =IF(F3> 0,D3,C3)

Use the Copy Command to copy cell B4 down column B

Similarly, copy C3 down column C, D4 down column D, E3 down column E, F3 down column F, and G3 down column G.

Now you should have the full table! Read down column C to find the value of r that makes $f(r)=0$.

Excel also has an addin called Solver that can search for any particular value for a function. We will also review this system in class (we will use Solver extensively in this course).

Exercise 23 (a) Show the cash flows for the three bonds below, each of which pays interest annually:

Bond	Coupon Rate %	Par Value	Years to Maturity	Price
X	14	\$ 1000	5	\$ 900
Y	16	2000	7	1900
Z	17	1000	4	950

(b) Calculate the yield to maturity for the three bonds.

Golden Section Search

Golden section search is similar in spirit to binary search. It is used to compute the maximum of a function $f(x)$ defined on an interval $[a, b]$, when the method presented in Section 2.2 is not applicable for one reason or another.

It assumes that

(i) f is continuous

(ii) f has a unique local maximum in the interval $[a, b]$.

The golden search method consists in computing $f(c)$ and f(d) for $a < d < c < b$.

- If $f(c) > f(d)$, the procedure is repeated with the interval (a, b) replaced by (d, b).
- If $f(c) < f(d)$, the procedure is repeated with the interval (a, b) replaced by (a, c).

Note: The name "golden section" comes from a certain choice of c and d that yields fast convergence, namely $c = a + r(b - a)$ and $d = b + r(a - b)$, where $r = \frac{\sqrt{5}-1}{2} = .618034\ldots$. This is the golden ratio, already known to the ancient greeks.

Example 2.3.2 *Find the maximum of the function* $x^5 - 10x^2 + 2x$ *in the interval* $[0, 1]$.

In this case, we begin with $a = 0$ and $b = 1$. Using golden section search, that gives $d = 0.382$ and $c = 0.618$. The function values are $f(a) = 0$, $f(d) = -0.687$, $f(c) = -2.493$, and $f(b) = -7$. Since $f(c) < f(d)$, our new range is $a = 0$, $b = .618$. Recalculating from the new range gives $d = .236$, $c = .382$ (note that our current c was our previous d: it is this reuse of calculated values that gives golden section search its speed). We repeat this process to get the following table:

Iter.	a	d	c	b	$f(a)$	$f(d)$	$f(c)$	$f(b)$
1	0	0.382	0.618	1	0	-0.6869	-2.4934	-7
2	0	0.2361	0.382	0.618	0	-0.0844	-0.6869	-2.4934
3	0	0.1459	0.2361	0.382	0	0.079	-0.0844	-0.6869
4	0	0.0902	0.1459	0.2361	0	0.099	0.079	-0.0844
5	0	0.0557	0.0902	0.1459	0	0.0804	0.099	0.079
6	0.0557	0.0902	0.1115	0.1459	0.0804	0.099	0.0987	0.079
7	0.0557	0.077	0.0902	0.1115	0.0804	0.0947	0.099	0.0987
8	0.077	0.0902	0.0983	0.1115	0.0947	0.099	0.1	0.0987
9	0.0902	0.0983	0.1033	0.1115	0.099	0.1	0.0999	0.0987
10	0.0902	0.0952	0.0983	0.1033	0.099	0.0998	0.1	0.0999
11	0.0952	0.0983	0.1002	0.1033	0.0998	0.1	0.1	0.0999
12	0.0983	0.1002	0.1014	0.1033	0.1	0.1	0.1	0.0999
13	0.0983	0.0995	0.1002	0.1014	0.1	0.1	0.1	0.1
14	0.0995	0.1002	0.1007	0.1014	0.1	0.1	0.1	0.1
15	0.0995	0.0999	0.1002	0.1007	0.1	0.1	0.1	0.1
16	0.0995	0.0998	0.0999	0.1002	0.1	0.1	0.1	0.1
17	0.0998	0.0999	0.1	0.1002	0.1	0.1	0.1	0.1
18	0.0999	0.1	0.1001	0.1002	0.1	0.1	0.1	0.1
19	0.0999	0.1	0.1	0.1001	0.1	0.1	0.1	0.1
20	0.0999	0.1	0.1	0.1	0.1	0.1	0.1	0.1
21	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1

Again we can use Excel to implement golden section search or use Solver to maximize this function directly.

2.4 Convexity

Finding global extrema and checking that we have actually found one is harder than finding local extrema, in general. There is one nice case: that of convex and concave functions. A convex function is one where the line segment connecting two points $(x, f(x))$ and $(y, f(y))$ lies above the function. Mathematically, a function f is *convex* if, for all x, y and all $0 < \alpha < 1$,

$$f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y).$$

See Figure 2.3. The function f is *concave* if $-f$ is convex.

There is an easy way to check for convexity when f is twice differentiable: the function f is convex on some domain $[a,b]$ if (and only if) $f''(x) \geq 0$ for all x in the domain. Similarly, f is concave on some domain $[a,b]$ if (and only if) $f''(x) \leq 0$ for all x in the domain.

If $f(x)$ is convex, then any local minimum is also a global minimum.

If $f(x)$ is concave, then any local maximum is also a global maximum.

Exercise 24 Find whether the following functions are concave, convex or neither.

(a) $x^4 - 4x^3 + 6x^2 + 3x + 1$

(b) $-e^{x^2}$

Exercise 25 Find a local extremum of the function $f(x) = xe^{-x}$. Indicate whether it is a local maximum, a local minimum or neither. Is it a global optimum on the domain [-2,2]?

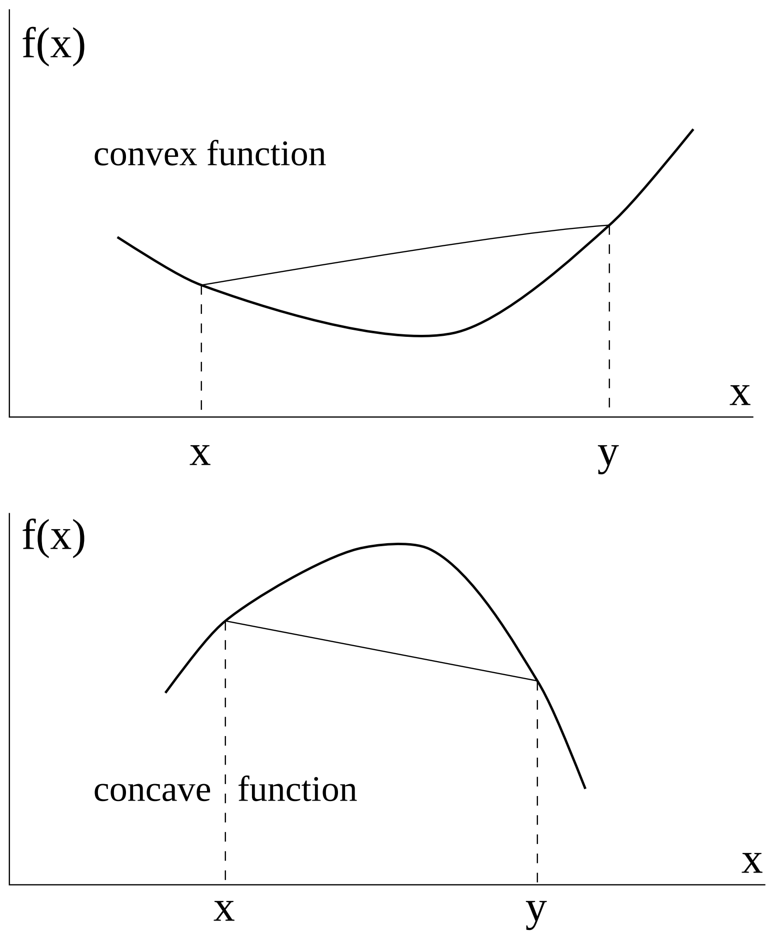


Figure 2.3: Convex function and concave function

Exercise 26 Consider the linear demand function $D(P) = K - aP$, where $P(\leq K/a)$ is the price and K and a are positive constants. Thus demand for a free good is K, and demand drops linearly until the price reaches K/a, at which point demand is zero.

a) Find the price P that maximizes revenue $PD(P)$ by setting the derivative of revenue equal to zero.

b) Use the second derivative to show that you found a local maximum.

c) Use the second derivative to show that the revenue function is concave. It follows that your local maximum is a global maximum.

Exercise 27 Oil is to be shipped in barrels having an undetermined height h and radius r. Each barrel must have a fixed volume V. We want to pick r and h so as to minimize the surface area of each barrel, given by $A = 2\pi r^2 + 2\pi rh$, and thereby minimize the cost of making the barrels. Since $V = \pi r^2 h$, we can solve for h in terms of r to get $h = V/\pi r^2$, and we can substitute this in the formula for A to get a formula for area solely in terms of r, namely,

$$A(r) = 2\pi r^2 + 2V/r.$$

a) Compute $A'(r)$, $A''(r)$ and show that the function A is convex on the set of positive reals.

b) Find the radius at which $A(r)$ has its global minimum. Explain how you know the minimum is global.

c) What is the ratio h/r when r has the optimal value you just derived?

Chapter 3

Unconstrained Optimization: Functions of Several Variables

Many of the concepts for functions of one variable can be extended to functions of several variables. For example, the gradient extends the notion of derivative. In this chapter, we review the notion of gradient, the formula for small changes, how to find extrema and the notion of convexity.

3.1 Gradient

Given a function f of n variables $x_1, x_2, \ldots, x_n$, we define the *partial derivative* relative to variable x_i, written as $\frac{\partial f}{\partial x_i}$, to be the derivative of f with respect to x_i treating all variables except x_i as constant.

Example 3.1.1 *Compute the partial derivatives of* $f(x_1, x_2) = (x_1 - 2)^2 + 2(x_2 - 1)^2$.

The answer is: $\frac{\partial f}{\partial x_1}(x_1, x_2) = 2(x_1 - 2), \qquad \frac{\partial f}{\partial x_2}(x_1, x_2) = 4(x_2 - 1)$.

Let x denote the vector $(x_1, x_2, \ldots, x_n)$. With this notation, $f(x) = f(x_1, x_2, \ldots, x_n)$, $\frac{\partial f}{\partial x_i}(x) = \frac{\partial f}{\partial x_i}(x_1, x_2, \ldots, x_n)$, etc. The *gradient* of f at x, written $\nabla f(x)$, is the vector $\begin{pmatrix} \frac{\partial f}{\partial x_1}(x) \\ \frac{\partial f}{\partial x_2}(x) \\ \vdots \\ \frac{\partial f}{\partial x_n}(x) \end{pmatrix}$. The gradient vector $\nabla f(x)$ gives the direction of steepest ascent of the function f at point x. The gradient acts like the derivative in that small changes around a given point x^* can be estimated using the gradient.

$$f(x^* + \Delta) \approx f(x^*) + \Delta \nabla f(x^*)$$

where $\Delta = (\Delta_1, \ldots, \Delta_n)$ denotes the vector of changes.

Example 3.1.2 *If* $f(x_1, x_2) = x_1^2 - 3x_1x_2 + x_2^2$, *then* $f(1,1) = -1$. *What about* $f(1.01, 1.01)$*?*

In this case, $x^* = (1,1)$ and $\Delta = (0.01, 0.01)$. Since $\frac{\partial f}{\partial x_1}(x_1, x_2) = 2x_1 - 3x_2$ and $\frac{\partial f}{\partial x_2}(x_1, x_2) = -3x_1 + 2x_2$, we get

$$\nabla f(1,1) = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$$

.

So $f(1.01, 1.01) = f((1,1)+(0.01, 0.01)) \approx f(1,1)+(0.01, 0.01)\nabla f(1,1) = -1+(0.01, 0.01) \begin{pmatrix} -1 \\ -1 \end{pmatrix} = -1.02$.

Example 3.1.3 Suppose that we want to put away a fishing pole in a closet having dimensions 3 by 5 by 1 feet. If the ends of the pole are placed at opposite corners, there is room for a pole of length,

$$f(x_1, x_2, x_3) = f(3,5,1) = \sqrt{x_1^2 + x_2^2 + x_3^2} = 5.9ft.$$

It turns out that the actual dimensions of the closet are $3 + \Delta_1$, $5 + \Delta_2$ and $1 + \Delta_3$ feet, where Δ_1, Δ_2 and Δ_3 are small correction terms. What is the change in pole length, taking into account these corrections?

By the formula for small changes, the change in pole length is

$$f(3 + \Delta_1, 5 + \Delta_2, 1 + \Delta_3) - f(3,5,1) \approx (\Delta_1, \Delta_2, \Delta_3)\nabla f(3,5,1).$$

So, we need to compute the partial derivatives of f. For $i = 1, 2, 3$

$$\frac{\partial}{\partial x_i} f(x_1, x_2, x_3) = \frac{x_i}{\sqrt{x_1^2 + x_2^2 + x_3^2}}.$$

Now we get

$$(\Delta_1, \Delta_2, \Delta_3)\nabla f(3,5,1) = (\Delta_1, \Delta_2, \Delta_3) \begin{pmatrix} 0.51 \\ 0.85 \\ 0.17 \end{pmatrix} = 0.51\Delta_1 + 0.85\Delta_2 + 0.17\Delta_3.$$

Exercise 28 Consider the function $f(x_1, x_2) = x_1 \ln x_2$.

(a) Compute the gradient of f.

(b) Give the value of the function f and give its gradient at the point $(3, 1)$.

(c) Use the formula for small changes to obtain an approximate value of the function at the point $(2.99, 1.05)$.

Exercise 29 Consider a conical drinking cup with height h and radius r at the open end. The volume of the cup is $V(r, h) = \frac{\pi}{3} r^2 h$.

a) Suppose the cone is now 5 cm high with radius 2 cm. Compute its volume.

b) Compute the partial derivatives $\partial V/\partial r$ and $\partial V/\partial h$ at the current height and radius.

c) By about what *fraction* (i.e., percentage) would the volume change if the cone were lengthened 10%? (Use the partial derivatives.)

d) If the radius were increased 5%?

e) If both were done simultaneously?

Hessian matrix

Second partials $\frac{\partial^2 f}{\partial x_i \partial x_j}(x)$ are obtained from $f(x)$ by taking the derivative relative to x_i (this yields the first partial $\frac{\partial f}{\partial x_i}(x)$) and then by taking the derivative of $\frac{\partial f}{\partial x_i}(x)$ relative to x_j. So we can compute $\frac{\partial^2 f}{\partial x_1 \partial x_1}(x)$, $\frac{\partial^2 f}{\partial x_1 \partial x_2}(x)$ and so on. These values are arranged into the *Hessian* matrix

$$H(x) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1 \partial x_1}(x) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(x) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(x) & \frac{\partial^2 f}{\partial x_2 \partial x_2}(x) & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(x) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x) & \frac{\partial^2 f}{\partial x_n \partial x_2}(x) & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_n}(x) \end{pmatrix}$$

The Hessian matrix is a symmetric matrix, that is $\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x)$.

Example 3.1.1 (continued): *Find the Hessian matrix of* $f(x_1, x_2) = (x_1 - 2)^2 + 2(x_2 - 1)^2$.

The answer is $H(x) = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$.

3.2 Maximum and Minimum

Optima can occur in three places:

1. at the boundary of the domain,
2. at a nondifferentiable point, or
3. at a point x^* with $\nabla f(x^*) = 0$.

We will identify the first type of point with Kuhn–Tucker conditions (see next chapter). The second type is found only by ad hoc methods. The third type of point can be found by solving the gradient equations.

In the remainder of this chapter, we discuss the important case where $\nabla f(x^*) = 0$. To identify if a point x^* with zero gradient is a local maximum or local minimum, check the Hessian.

- If $H(x^*)$ is positive definite then x^* is a local minimum.
- If $H(x^*)$ is negative definite, then x^* is a local maximum.

Remember (Section 1.6) that these properties can be checked by computing the determinants of the principal minors.

Example 3.2.1 *Find the local extrema of* $f(x_1, x_2) = x_1^3 + x_2^3 - 3x_1x_2$.

This function is everywhere differentiable, so extrema can only occur at points x^* such that $\nabla f(x^*) = 0$.

$$\nabla f(x) = \begin{pmatrix} 3x_1^2 - 3x_2 \\ 3x_2^2 - 3x_1 \end{pmatrix}$$

This equals 0 iff $(x_1, x_2) = (0, 0)$ or $(1, 1)$. The Hessian is

$$H(x) = \begin{pmatrix} 6x_1 & -3 \\ -3 & 6x_2 \end{pmatrix}$$

So,

$$H(0, 0) = \begin{pmatrix} 0 & -3 \\ -3 & 0 \end{pmatrix}$$

Let H_1 denote the first principal minor of $H(0, 0)$ and let H_2 denote its second principal minor (see Section 1.6). Then $\det(H_1) = 0$ and $\det(H_2) = -9$. Therefore $H(0, 0)$ is neither positive nor negative definite.

$$H(1, 1) = \begin{pmatrix} 6 & -3 \\ -3 & 6 \end{pmatrix}$$

Its first principal minor has $\det(H_1) = 6 > 0$ and its second principal minor has $\det(H_2) = 36 - 9 = 25 > 0$. Therefore $H(1, 1)$ is positive definite, which implies that $(1, 1)$ is a local minimum.

Example 3.2.2 *Jane and Jim invested \$20,000 in the design and development of a new product. They can manufacture it for \$2 per unit. For the next step, they hired marketing consultants XYZ. In a nutshell, XYZ's conclusions are the following: if Jane and Jim spend \$a on advertizing and sell the product at price p (per unit), they will sell*

$$2,000 + 4\sqrt{a} - 20p \quad \textit{units.}$$

Using this figure, express the profit that Jane and Jim will make as a function of a and p. What price and level of advertising will maximize their profits?

The revenue from sales is $(2,000 + 4\sqrt{a} - 20p)p$.

The production costs are $(2,000 + 4\sqrt{a} - 20p)2$, the development cost is \$20,000 and the cost of advertizing is a.

Therefore, Jane and Jim's profit is

$$f(p, a) = (2,000 + 4\sqrt{a} - 20p)(p - 2) - a - 20,000$$

To find the maximum profit, we compute the partial derivatives of f and set them to 0:

$\frac{\partial f}{\partial p}(p, a) = 2,040 + 4\sqrt{a} - 40p = 0.$

$\frac{\partial f}{\partial a}(p, a) = 2(p - 2)/\sqrt{a} - 1 = 0.$

Solving this system of two equations yields

$$p = 63.25 \qquad a = 15,006.25$$

We verify that this is a maximum by computing the Hessian.

$$H(x) = \begin{pmatrix} -40 & 2/\sqrt{a} \\ 2/\sqrt{a} & -(p-2)/a\sqrt{a} \end{pmatrix}$$

$\det(H_1) = -40 < 0$ and $\det(H_2) = 40(p-2)/a\sqrt{a} - 4/a > 0$ at the point $p = 63.25$, $a = 15,006.25$. So, indeed, this solution maximizes profit.

Example 3.2.3 *Find the local extrema of* $f(x_1, x_2, x_3) = x_1^2 + (x_1 + x_2)^2 + (x_1 + x_3)^2$.

$\frac{\partial f}{\partial x_1}(x) = 2x_1 + 2(x_1 + x_2) + 2(x_1 + x_3)$
$\frac{\partial f}{\partial x_2}(x) = 2(x_1 + x_2)$
$\frac{\partial f}{\partial x_3}(x) = 2(x_1 + x_3)$

Setting these partial derivatives to 0 yields the unique solution $x_1 = x_2 = x_3 = 0$. The Hessian matrix is

$$H(0,0,0) = \begin{pmatrix} 6 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix}$$

The determinants of the principal minors are $\det(H_1) = 6 > 0$, $\det(H_2) = 12 - 4 = 8 > 0$ and $\det(H_3) = 24 - 8 - 8 = 8 > 0$. So $H(0,0,0)$ is positive definite and the solution $x_1 = x_2 = x_3 = 0$ is a minimum.

Exercise 30 Find maxima or minima of the following functions when possible.

(a) $f(x_1, x_2, x_3) = -x_1^2 - 3x_2^2 - 10x_3^2 + 4x_1 + 24x_2 + 20x_3$

(b) $f(x_1, x_2, x_3) = x_1x_2 + x_2x_3 + x_3x_1 - 2x_1 - 2x_2 - 2x_3$

Exercise 31 Consider the function of three variables given by

$$f(x_1, x_2, x_3) = x_1^2 - x_1 - x_1x_2 + x_2^2 - x_2 + x_3^4 - 4x_3.$$

(a) Compute the gradient $\nabla f(x_1, x_2, x_3)$.

(b) Compute the Hessian matrix $H(x_1, x_2, x_3)$.

(c) Use the gradient to find a local extremum of f.
Hint: if $x_3^3 = 1$, *then* $x_3 = 1$.

(d) Compute the three principal minors of the Hessian matrix and use them to identify this extremum as a local minimum or a local maximum.

3.3 Global Optima

Finding global maxima and minima is harder. There is one case that is of interest.

We say that a domain is *convex* if every line drawn between two points in the domain lies within the domain.

We say that a function f is *convex* if the line connecting any two points lies above the function. That is, for all x, y in the domain and $0 < \alpha < 1$, we have $f(\alpha x + (1-\alpha)y) \leq \alpha f(x) + (1-\alpha) f(y)$, as before (see Chapter 2).

- If a function is convex on a convex domain, then any local minimum is a global minimum.
- If a function is concave on a convex domain, then any local maximum is a global maximum.

To check that a function is convex on a domain, check that its Hessian matrix $H(x)$ is positive semidefinite for every point x in the domain. To check that a function is concave, check that its Hessian is negative semidefinite for every point in the domain.

Example 3.3.1 *Show that the function* $f(x_1, x_2, x_3) = x_1^4 + (x_1 + x_2)^2 + (x_1 + x_3)^2$ *is convex over* $\Re^3$.

$\frac{\partial f}{\partial x_1}(x) = 4x_1^3 + 2(x_1 + x_2) + 2(x_1 + x_3)$
$\frac{\partial f}{\partial x_2}(x) = 2(x_1 + x_2)$
$\frac{\partial f}{\partial x_3}(x) = 2(x_1 + x_3)$

$$H(x_1, x_2, x_3) = \begin{pmatrix} 12x_1^2 + 4 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix}$$

The determinants of the principal minors are $\det(H_1) = 12x_1^2 + 4 > 0$, $\det(H_2) = 12x_1^2 \geq 0$ and $\det(H_3) = 48x_1^2 \geq 0$. So $H(x_1, x_2, x_3)$ is positive semidefinite for all (x_1, x_2, x_3) in $\Re^3$. This implies that f is convex over $\Re^3$.

Exercise 32 For each of the following, determine whether the function is convex, concave, or neither over $\Re^2$.
(a) $f(x) = x_1x_2 - x_1^2 - x_2^2$
(b) $f(x) = 10x_1 + 20x_2$
(c) $f(x) = x_1^4 + x_1x_2$
(d) $f(x) = -x_1^2 - x_1x_2 - 2x_2^2$

Exercise 33 Let the following function be defined for all points (x, y) in the plane.

$$f(x, y) = 2xy - x^4 - x^2 - y^2.$$

(a) Write the gradient of the function f.

(b) Write the Hessian matrix of f.

(c) Is the function f convex, concave or neither?

(d) Use the gradient to find a local extremum of f.

(e) Identify this extremum as a minimum, a maximum or neither.

Chapter 4

Constrained Optimization

4.1 Equality Constraints (Lagrangians)

Suppose we have a problem:

Maximize $5 - (x_1 - 2)^2 - 2(x_2 - 1)^2$
subject to
$x_1 + 4x_2 = 3$

If we ignore the constraint, we get the solution $x_1 = 2, x_2 = 1$, which is too large for the constraint. Let us penalize ourselves λ for making the constraint too big. We end up with a function

$$L(x_1, x_2, \lambda) = 5 - (x_1 - 2)^2 - 2(x_2 - 1)^2 + \lambda(3 - x_1 - 4x_2)$$

This function is called the *Lagrangian* of the problem. The main idea is to adjust λ so that we use exactly the right amount of the resource.

$\lambda = 0$ leads to $(2, 1)$.

$\lambda = 1$ leads to $(3/2, 0)$ which uses too little of the resource.

$\lambda = 2/3$ gives $(5/3, 1/3)$ and the constraint is satisfied exactly.

We now explore this idea more formally. Given a nonlinear program (P) with equality constraints:

Minimize (or maximize) $f(x)$
subject to
$g_1(x) = b_1$
$g_2(x) = b_2$
$\vdots$
$g_m(x) = b_m$

a solution can be found using the *Lagrangian*:

$$L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i (b_i - g_i(x))$$

(Note: this can also be written $f(x) - \sum_{i=1}^{m} \lambda_i (g_i(x) - b_i)$).
Each λ_i gives the price associated with constraint i.

The reason L is of interest is the following:

> Assume $x^* = (x_1^*, x_2^*, \ldots, x_n^*)$ maximizes or minimizes $f(x)$ subject to the constraints $g_i(x) = b_i$, for $i = 1, 2, \ldots, m$. Then either
>
> (i) the vectors $\nabla g_1(x^*), \nabla g_2(x^*), \ldots, \nabla g_m(x^*)$ are linearly dependent, or
>
> (ii) there exists a vector $\lambda^* = (\lambda_1^*, \lambda_2^*, \ldots, \lambda_m^*)$ such that $\nabla L(x^*, \lambda^*) = 0$. I.e.
>
> $$\frac{\partial L}{\partial x_1}(x^*, \lambda^*) = \frac{\partial L}{\partial x_2}(x^*, \lambda^*) = \cdots = \frac{\partial L}{\partial x_n}(x^*, \lambda^*) = 0$$
>
> $$\text{and} \qquad \frac{\partial L}{\partial \lambda_1}(x^*, \lambda^*) = \cdots = \frac{\partial L}{\partial \lambda_m}(x^*, \lambda^*) = 0$$

Of course, Case (i) above cannot occur when there is only one constraint. The following example shows how it might occur.

Example 4.1.1
Minimize $x_1 + x_2 + x_3^2$
subject to
$x_1 = 1$
$x_1^2 + x_2^2 = 1.$

It is easy to check directly that the minimum is acheived at $(x_1, x_2, x_3) = (1, 0, 0)$. The associated Lagrangian is

$$L(x_1, x_2, x_3, \lambda_1, \lambda_2) = x_1 + x_2 + x_3^2 + \lambda_1(1 - x_1) + \lambda_2(1 - x_1^2 - x_2^2).$$

Observe that

$$\frac{\partial L}{\partial x_2}(1, 0, 0, \lambda_1, \lambda_2) = 1 \qquad \text{for all} \;\; \lambda_1, \lambda_2,$$

and consequently $\frac{\partial L}{\partial x_2}$ *does not* vanish at the optimal solution. The reason for this is the following. Let $g_1(x_1, x_2, x_3) = x_1$ and $g_2(x_1, x_2, x_3) = x_1^2 + x_2^2$ denote the left hand sides of the constraints. Then $\nabla g_1(1, 0, 0) = (1, 0, 0)$ and $\nabla g_2(1, 0, 0) = (2, 0, 0)$ are linearly dependent vectors. So Case (i) occurs here!

Nevertheless, Case (i) will not concern us in this course. When solving optimization problems with equality constraints, we will only look for solutions x^* that satisfy Case (ii).

Note that the equation

$$\frac{\partial L}{\partial \lambda_i}(x^*, \lambda^*) = 0$$

is nothing more than

$$b_i - g_i(x^*) = 0 \qquad \text{or} \qquad g_i(x^*) = b_i.$$

In other words, taking the partials with respect to λ does nothing more than return the original constraints.

Once we have found candidate solutions x^*, it is not always easy to figure out whether they correspond to a minimum, a maximum or neither. The following situation is one when we can conclude. If $f(x)$ is concave and all of the $g_i(x)$ are linear, then any feasible x^* with a corresponding λ^* making $\nabla L(x^*, \lambda^*) = 0$ maximizes $f(x)$ subject to the constraints. Similarly, if $f(x)$ is convex and each $g_i(x)$ is linear, then any x^* with a λ^* making $\nabla L(x^*, \lambda^*) = 0$ minimizes $f(x)$ subject to the constraints.

Example 4.1.2
Minimize $2x_1^2 + x_2^2$
subject to
$x_1 + x_2 = 1$

$L(x_1, x_2, \lambda) = 2x_1^2 + x_2^2 + \lambda_1(1 - x_1 - x_2)$

$$\frac{\partial L}{\partial x_1}(x_1^*, x_2^*, \lambda^*) = 4x_1^* - \lambda_1^* = 0$$

$$\frac{\partial L}{\partial x_2}(x_1^*, x_2^*, \lambda^*) = 2x_2^* - \lambda_1^* = 0$$

$$\frac{\partial L}{\partial \lambda}(x_1^*, x_2^*, \lambda^*) = 1 - x_1^* - x_2^* = 0$$

Now, the first two equations imply $2x_1^* = x_2^*$. Substituting into the final equation gives the solution $x_1^* = 1/3$, $x_2^* = 2/3$ and $\lambda^* = 4/3$, with function value 2/3.

Since $f(x_1, x_2)$ is convex (its Hessian matrix $H(x_1, x_2) = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}$ is positive definite) and $g(x_1, x_2) = x_1 + x_2$ is a linear function, the above solution minimizes $f(x_1, x_2)$ subject to the constraint.

4.1.1 Geometric Interpretation

There is a geometric interpretation of the conditions an optimal solution must satisfy. If we graph Example 4.1.2, we get a picture like that in Figure 4.1.

Now, examine the gradients of f and g at the optimum point. They must point in the same direction, though they may have different lengths. This implies:

$$\nabla f(x^*) = \lambda^* \nabla g(x^*)$$

which, along with the feasibility of x^*, is exactly the condition $\nabla L(x^*, \lambda^*) = 0$ of Case (ii).

4.1.2 Economic Interpretation

The values λ_i^* have an important economic interpretation: If the right hand side b_i of Constraint i is increased by Δ, then the optimum objective value increases by approximately $\lambda_i^* \Delta$.

In particular, consider the problem

Maximize $p(x)$
subject to
$g(x) = b,$

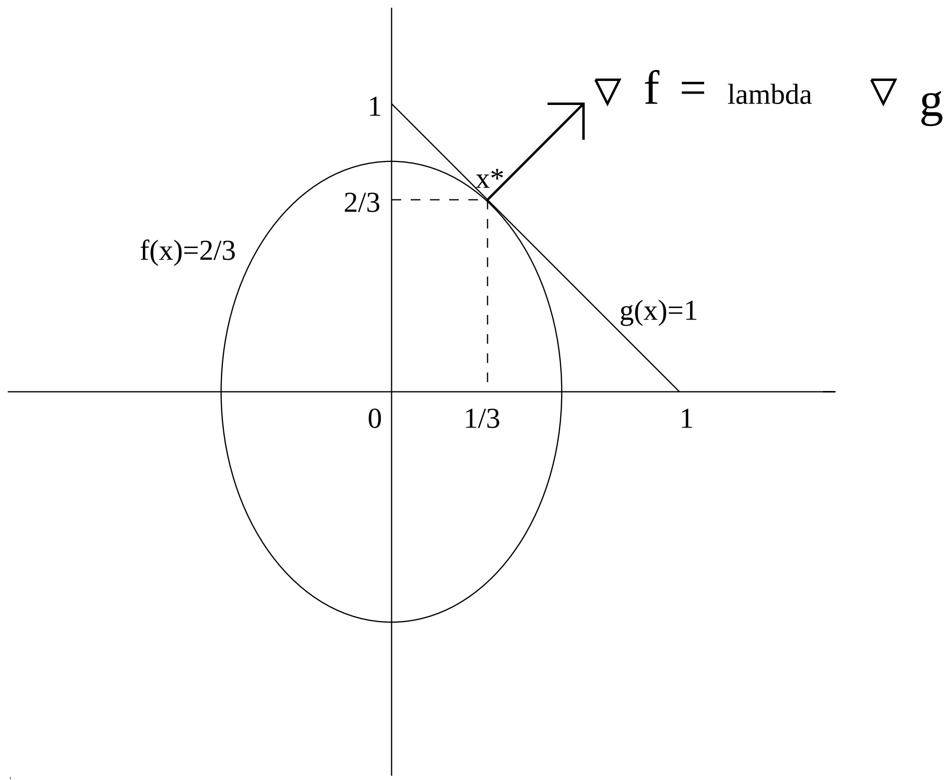


Figure 4.1: Geometric interpretation

where $p(x)$ is a profit to maximize and b is a limited amount of resource. Then, the optimum Lagrange multiplier λ^* is the marginal value of the resource. Equivalently, if b were increased by Δ, profit would increase by $\lambda^*\Delta$. This is an important result to remember. It will be used repeatedly in your Managerial Economics course.

Similarly, if

Minimize $c(x)$
subject to
$d(x) = b$,

represents the minimum cost $c(x)$ of meeting some demand b, the optimum Lagrange multiplier λ^* is the marginal cost of meeting the demand.

In Example 4.1.2

Minimize $2x_1^2 + x_2^2$
subject to
$x_1 + x_2 = 1$,

if we change the right hand side from 1 to 1.05 (i.e. $\Delta = 0.05$), then the optimum objective function value goes from $\frac{2}{3}$ to roughly

$$\frac{2}{3} + \frac{4}{3}(0.05) = \frac{2.2}{3}.$$

If instead the right hand side became 0.98, our estimate of the optimum objective function value would be

$$\frac{2}{3} + \frac{4}{3}(-0.02) = \frac{1.92}{3}$$

Example 4.1.3 *Suppose we have a refinery that must ship finished goods to some storage tanks. Suppose further that there are two pipelines, A and B, to do the shipping. The cost of shipping x units on A is ax^2; the cost of shipping y units on B is by^2, where $a > 0$ and $b > 0$ are given. How can we ship Q units while minimizing cost? What happens to the cost if Q increases by $r\%$?*

Minimize $ax^2 + by^2$
Subject to
$x + y = Q$

$$L(x, y, \lambda) = ax^2 + by^2 + \lambda(Q - x - y)$$
$$\frac{\partial L}{\partial x}(x^*, y^*, \lambda^*) = 2ax^* - \lambda^* = 0$$
$$\frac{\partial L}{\partial y}(x^*, y^*, \lambda^*) = 2by^* - \lambda^* = 0$$
$$\frac{\partial L}{\partial \lambda}(x^*, y^*, \lambda^*) = Q - x^* - y^* = 0$$

The first two constraints give $x^* = \frac{b}{a}y^*$, which leads to

$$x^* = \frac{bQ}{a+b}, \qquad y^* = \frac{aQ}{a+b}, \qquad \lambda^* = \frac{2abQ}{a+b}$$

and cost of $\frac{abQ^2}{a+b}$. The Hessian matrix $H(x_1, x_2) = \begin{pmatrix} 2a & 0 \\ 0 & 2b \end{pmatrix}$ is positive definite since $a > 0$ and $b > 0$. So this solution minimizes cost, given a, b, Q.

If Q increases by $r\%$, then the RHS of the constraint increases by $\Delta = rQ$ and the minimum cost increases by $\lambda^*\Delta = \frac{2abrQ^2}{a+b}$. That is, the minimum cost increases by $2r\%$.

Example 4.1.4 How should one divide his/her savings between three mutual funds with expected returns 10%, 10% and 15% repectively, so as to minimize risk while achieving an expected return of 12%. We measure risk as the variance of the return on the investment (you will learn more about measuring risk in 45-733): when a fraction x of the savings is invested in Fund 1, y in Fund 2 and z in Fund 3, where $x + y + z = 1$, the variance of the return has been calculated to be

$$400x^2 + 800y^2 + 200xy + 1600z^2 + 400yz.$$

So your problem is

$$\begin{array}{lrcrcrcrcr} \min & 400x^2 & + & 800y^2 & + & 200xy & + & 1600z^2 & + & 400yz \\ s.t. & x & + & y & + & 1.5z & = & 1.2 & & \\ & x & + & y & + & z & = & 1 & & \end{array}$$

Using the Lagrangian method, the following optimal solution was obtained

$$x = 0.5 \quad y = 0.1 \quad z = 0.4 \quad \lambda_1 = 1800 \quad \lambda_2 = -1380$$

where λ_1 is the Lagrange multiplier associated with the first constraint and λ_2 with the second constraint. The corresponding objective function value (i.e. the variance on the return) is 390. If an expected return of 12.5% was desired (instead of 12%), what would be (approximately) the correcponding variance of the return?

Since $\Delta = 0.05$, the variance would increase by

$$\Delta\lambda_1 = 0.05 \times 1800 = 90.$$

So the answer is 390+90=480.

Exercise 34 Record'm Records needs to produce 100 gold records at one or more of its three studios. The cost of producing x records at studio 1 is $10x$; the cost of producing y records at studio 2 is $2y^2$; the cost of producing z records at studio 3 is $z^2 + 8z$.

(a) Formulate the nonlinear program of producing the 100 records at minimum cost.

(b) What is the Lagrangian associated with your formulation in (a)?

(c) Solve this Lagrangian. What is the optimal production plan?

(d) What is the marginal cost of producing one extra gold record?

(e) Union regulations require that exactly 60 hours of work be done at studios 2 and 3 combined. Each gold record requires 4 hours at studio 2 and 2 hours at studio 3. Formulate the problem of finding the optimal production plan, give the Lagrangian, and give the set of equations that must be solved to find the optimal production plan. It is not necessary to actually solve the equations.

Exercise 35 (a) Solve the problem

$$\begin{array}{ll} \max & 2x + y \\ \text{subject to} & 4x^2 + y^2 = 8 \end{array}$$

(b) Estimate the change in the optimal objective function value when the right hand side increases by 5%, i.e. the right hand side increases from 8 to 8.4.

Exercise 36

(a) Solve the following constrained optimization problem using the method of Lagrange multipliers.

$$\begin{array}{ll} \max & \ln x + 2\ln y + 3\ln z \\ \text{subject to} & x + y + z = 60 \end{array}$$

(b) Estimate the change in the optimal objective function value if the right hand side of the constraint increases from 60 to 65.

4.2 Equality and Inequality Constraints

How do we handle both equality and inequality constraints in (P)? Let (P) be:

Maximize $f(x)$
Subject to
$g_1(x) = b_1$
$\vdots$
$g_m(x) = b_m$
$h_1(x) \leq d_1$
$\vdots$
$h_p(x) \leq d_p$

If you have a program with $\geq$ constraints, convert it into $\leq$ by multiplying by -1. Also convert a minimization to a maximization.

The Lagrangian is

$$L(x, \lambda, \mu) = f(x) + \sum_{i=1}^{m} \lambda_i (b_i - g_i(x)) + \sum_{j=1}^{p} \mu_j (d_j - h_j(x))$$

The fundamental result is the following:

> Assume $x^* = (x_1^*, x_2^*, \ldots, x_n^*)$ maximizes $f(x)$ subject to the constraints $g_i(x) = b_i$, for $i = 1, 2, \ldots, m$ and $h_j(x) \leq d_j$, for $j = 1, 2, \ldots, p$. Then either
>
> (i) the vectors $\nabla g_1(x^*), \ldots, \nabla g_m(x^*), \ \nabla h_1(x^*), \ldots, \nabla h_p(x^*)$ are linearly dependent, or
>
> (ii) there exists vectors $\lambda^* = (\lambda_1^*, \ldots, \lambda_m^*)$ and $\mu^* = (\mu_1^*, \ldots, \mu_p^*)$ such that
>
> $$\nabla f(x^*) - \sum_{i=1}^{m} \lambda_i^* \nabla g_i(x^*) - \sum_{j=1}^{p} \mu_j^* \nabla h_j(x^*) = 0$$
>
> $$\mu_j^*(h_j(x^*) - d_j) = 0 \qquad \text{(Complementarity)}$$
>
> $$\mu_j^* \geq 0$$

In this course, we will not concern ourselves with Case (i). We will only look for candidate solutions x^* for which we can find λ^* and μ^* satisfying the equations in Case (ii) above.

In general, to solve these equations, you begin with complementarity and note that either μ_j^* must be zero or $h_j(x^*) - d_j = 0$. Based on the various possibilities, you come up with one or more candidate solutions. If there is an optimal solution, then one of your candidates will be it.

The above conditions are called the *Kuhn–Tucker (or Karush–Kuhn–Tucker)* conditions. Why do they make sense?

For x^* optimal, some of the inequalities will be tight and some not. Those not tight can be ignored (and will have corresponding price $\mu_j^* = 0$). Those that are tight can be treated as equalities which leads to the previous Lagrangian stuff. So

$$\mu_j^*(h_j(x^*) - d_j) = 0 \qquad \text{(Complementarity)}$$

forces either the price μ_j^* to be 0 or the constraint to be tight.

Example 4.2.1
Maximize $x^3 - 3x$
Subject to
$x \leq 2$

The Lagrangian is

$$L = x^3 - 3x + \mu(2 - x)$$

So we need

$$3x^2 - 3 - \mu = 0$$
$$x \leq 2$$
$$\mu(2 - x) = 0$$
$$\mu \geq 0$$

Typically, at this point we must break the analysis into cases depending on the complementarity conditions.

If $\mu = 0$ then $3x^2 - 3 = 0$ so $x = 1$ or $x = -1$. Both are feasible. $f(1) = -2$, $f(-1) = 2$.

If $x = 2$ then $\mu = 9$ which again is feasible. Since $f(2) = 2$, we have two solutions: $x = -1$ and $x = 2$.

Example 4.2.2 *Minimize* $(x - 2)^2 + 2(y - 1)^2$
Subject to
$x + 4y \leq 3$
$x \geq y$

First we convert to standard form, to get

Maximize $-(x - 2)^2 - 2(y - 1)^2$
Subject to
$x + 4y \leq 3$
$-x + y \leq 0$

$$L(x, y, \mu_1, \mu_2) = -(x - 2)^2 - 2(y - 1)^2 + \mu_1(3 - x - 4y) + \mu_2(0 + x - y)$$

which gives optimality conditions

$$-2(x - 2) - \mu_1 + \mu_2 = 0$$
$$-4(y - 1) - 4\mu_1 - \mu_2 = 0$$
$$\mu_1(3 - x - 4y) = 0$$

$$\mu_2(x-y)=0$$
$$x+4y\le 3$$
$$-x+y\le 0$$
$$\mu_1,\mu_2\ge 0$$

Since there are two complementarity conditions, there are four cases to check:
$\mu_1=0,\mu_2=0$: gives $x=2$, $y=1$ which is not feasible.
$\mu_1=0, x-y=0$: gives $x=4/3, y=4/3, \mu_2=-4/3$ which is not feasible.
$\mu_2=0, 3-x-4y=0$ gives $x=5/3, y=1/3, \mu_1=2/3$ which is O.K.
$3-x-4y=0, x-y=0$ gives $x=3/5, y=3/5, \mu_1=22/25, \mu_2=-48/25$ which is not feasible.
Since it is clear that there is an optimal solution, $x=5/3, y=1/3$ is it!

Economic Interpretation

The economic interpretation is essentially the same as the equality case. If the right hand side of a constraint is changed by a small amount Δ, then the optimal objective changes by $\mu^*\Delta$, where μ^* is the optimal Lagrange multiplier corresponding to that constraint. Note that if the constraint is not tight then the objective does not change (since then $\mu^*=0$).

Handling Nonnegativity

A special type of constraint is nonnegativity. If you have a constraint $x_k\ge 0$, you can write this as $-x_k\le 0$ and use the above result. This constraint would get a Lagrange multiplier of its own, and would be treated just like every other constraint.

An alternative is to treat nonnegativity implicitly. If x_k must be nonnegative:

1. Change the equality associated with its partial to a less than or equal to zero:

$$\frac{\partial f(x)}{\partial x_k}-\sum_i \lambda_i\frac{\partial g_i(x)}{\partial x_k}-\sum_j \mu_j\frac{\partial h_j(x)}{\partial x_k}\le 0$$

2. Add a new complementarity constraint:

$$\left(\frac{\partial f(x)}{\partial x_k}-\sum_i \lambda_i\frac{\partial g_i(x)}{\partial x_k}-\sum_j \mu_j\frac{\partial h_j(x)}{\partial x_k}\right)x_k=0$$

3. Don't forget that $x_k\ge 0$ for x to be feasible.

Sufficiency of conditions

The Karush–Kuhn–Tucker conditions give us candidate optimal solutions x^*. When are these conditions sufficient for optimality? That is, given x^* with λ^* and μ^* satisfying the KKT conditions, when can we be certain that it is an optimal solution?

The most general condition available is:

1. $f(x)$ is concave, and
2. the feasible region forms a convex region.

While it is straightforward to determine if the objective is concave by computing its Hessian matrix, it is not so easy to tell if the feasible region is convex. A useful condition is as follows:

The feasible region is convex if all of the $g_i(x)$ are linear and all of the $h_j(x)$ are convex. If this condition is satisfied, then any point that satisfies the KKT conditions gives a point that maximizes $f(x)$ subject to the constraints.

Example 4.2.3 *Suppose we can buy a chemical for \$10 per ounce. There are only 17.25 oz available. We can transform this chemical into two products: A and B. Transforming to A costs \$3 per oz, while transforming to B costs \$5 per oz. If x_1 oz of A are produced, the price we command for A is $\$30 - x_1$; if x_2 oz of B are produced, the price we get for B is $\$50 - x_2$. How much chemical should we buy, and what should we transform it to?*

There are many ways to model this. Let's let x_3 be the amount of chemical we purchase. Here is one model:

Maximize $x_1(30 - x_1) + x_2(50 - 2x_2) - 3x_1 - 5x_2 - 10x_3$
Subject to
$x_1 + x_2 - x_3 \leq 0$
$x_3 \leq 17.25$

The KKT conditions are the above feasibility constraints along with:
$30 - 2x_1 - 3 - \mu_1 = 0$
$50 - 4x_2 - 5 - \mu_1 = 0$
$-10 + \mu_1 - \mu_2 = 0$
$\mu_1(-x_1 - x_2 + x_3) = 0$
$\mu_2(17.25 - x_3) = 0$
$\mu_1, \mu_2 \geq 0$

There are four cases to check:
$\mu_1 = 0, \mu_2 = 0$. This gives us $-10 = 0$ in the third constraint, so is infeasible.
$\mu_1 = 0, x_3 = 17.25$. This gives $\mu_2 = -10$ so is infeasible.
$-x_1 - x_2 + x_3 = 0$, $\mu_2 = 0$. This gives $\mu_1 = 10, x_1 = 8.5, x_2 = 8.75, x_3 = 17.25$, which is feasible. Since the objective is concave and the constraints are linear, this must be an optimal solution. So there is no point in going through the last case ($-x_1 - x_2 + x_3 = 0$, $x_3 = 17.25$). We are done with $x_1 = 8.5, x_2 = 8.75$, and $x_3 = 17.25$.

What is the value of being able to purchase 1 more unit of chemical?

This question is equivalent to increasing the right hand side of the constraint $x_3 \leq 17.25$ by 1 unit. Since the corresponding lagrangian value is 0, there is no value in increasing the right hand side.

Review of Optimality Conditions.

The following reviews what we have learned so far:

Single Variable (Unconstrained)
Solve $f'(x) = 0$ to get candidate x^*.
If $f''(x^*) > 0$ then x^* is a local min.
$f''(x^*) < 0$ then x^* is a local max.
If $f(x)$ is convex then a local min is a global min.
$f(x)$ is concave then a local max is a global max.

Multiple Variable (Unconstrained)
Solve $\nabla f(x) = 0$ to get candidate x^*.
If $H(x^*)$ is positive definite then x^* is a local min.
$H(x^*)$ is negative definite x^* is a local max.
If $f(x)$ is convex then a local min is a global min.
$f(x)$ is concave then a local max is a global max.

Multiple Variable (Equality constrained) Form Lagrangian $L(x, \lambda) = f(x) + \sum_i \lambda_i(b_i - g_i(x))$
Solve $\nabla L(x, \lambda) = 0$ to get candidate x^* (and λ^*).
Best x^* is optimum if optimum exists.

Multiple Variable (Equality and Inequality constrained)
Put into standard form (maximize and $\leq$ constraints)
Form Lagrangian $L(x, \lambda) = f(x) + \sum_i \lambda_i(b_i - g_i(x)) + \sum_j \mu_j(d_j - h_j(x))$
Solve
$\nabla f(x) - \sum_i \lambda_i \nabla g_i(x) - \sum_j \mu_j \nabla h_j(x) = 0$
$g_i(x) = b_i$
$h_j(x) \leq d_j$
$\mu_j(d_j - h_j(x)) = 0$
$\mu_j \geq 0$
to get candidates x^* (and λ^*, μ^*).
Best x^* is optimum if optimum exists.
Any x^* is optimum if $f(x)$ concave, $g_i(x)$ convex, $h_j(x)$ linear.

4.3 Exercises

Exercise 37 Solve the following constrained optimization problem using the method of Lagrange multipliers.

$$\begin{array}{ll} \max & 2\ln x_1 + 3\ln x_2 + 3\ln x_3 \\ \text{s.t.} & x_1 + 2x_2 + 2x_3 = 10 \end{array}$$

Exercise 38 Find the two points on the ellipse given by $x_1^2+4x_2^2=4$ that are at minimum distance of the point $(1,0)$. Formulate the problem as a minimization problem and solve it by solving the Lagrangian equations. [*Hint:* To minimize the distance d between two points, one can also minimize d^2. The formula for the distance between points (x_1,x_2) and (y_1,y_2) is $d^2=(x_1-y_1)^2+(x_2-y_2)^2$.]

Exercise 39 Solve using Lagrange multipliers.

a) $\min x_1^2+x_2^2+x_3^2$ subject to $x_1+x_2+x_3=b$, where b is given.

b) $\max \sqrt{x_1}+\sqrt{x_2}+\sqrt{x_3}$ subject to $x_1+x_2+x_3=b$, where b is given.

c) $\min c_1x_1^2+c_2x_2^2+c_3x_3^2$ subject to $x_1+x_2+x_3=b$, where $c_1>0$, $c_2>0$, $c_3>0$ and b are given.

d) $\min x_1^2+x_2^2+x_3^2$ subject to $x_1+x_2=b_1$ and $x_2+x_3=b_2$, where b_1 and b_2 are given.

Exercise 40 Let a, b and c be given positive scalars. What is the change in the optimum value of the following constrained optimization problem when the right hand side of the constraint is increased by 5%, i.e. a is changed to $a+\frac{5}{100}a$.

$$\begin{array}{ll} \max & by-x^4 \\ \text{s.t.} & x^2+cy=a \end{array}$$

Give your answer in terms of a, b and c.

Exercise 41 You want to invest in two mutual funds so as to maximize your expected earnings while limiting the variance of your earnings to a given figure s^2. The expected yield rates of Mutual Funds 1 and 2 are r_1 and r_2 respectively, and the variance of earnings for the portfolio (x_1,x_2) is $\sigma^2x_1^2+\rho x_1x_2+\sigma^2x_2^2$. Thus the problem is

$$\begin{array}{ll} \max & r_1x_1+r_2x_2 \\ \text{s.t.} & \sigma^2x_1^2+\rho x_1x_2+\sigma^2x_2^2=s^2 \end{array}$$

(a) Use the method of Lagrange multipliers to compute the optimal investments x_1 and x_2 in Mutual Funds 1 and 2 respectively. Your expressions for x_1 and x_2 should not contain the Lagrange multiplier λ.

(b) Suppose both mutual funds have the same yield r. How much should you invest in each?

Exercise 42 You want to minimize the surface area of a cone-shaped drinking cup having fixed volume V_0. Solve the problem as a constrained optimization problem. To simplify the algebra, minimize the square of the area. The area is $\pi r\sqrt{r^2+h^2}$. The problem is,

$$\begin{array}{ll} \min & \pi^2r^4+\pi^2r^2h^2 \\ \text{s.t.} & \frac{1}{3}\pi r^2h=V_0. \end{array}$$

Solve the problem using Lagrange multipliers.

Hint. You can assume that $r\neq 0$ in the optimal solution.

Exercise 43 A company manufactures two types of products: a standard product, say A, and a more sophisticated product B. If management charges a price of p_A for one unit of product A and a price of p_B for one unit of product B, the company can sell q_A units of A and q_B units of B, where

$$q_A = 400 - 2p_A + p_B, \qquad q_B = 200 + p_A - p_B.$$

Manufacturing one unit of product A requires 2 hours of labor and one unit of raw material. For one unit of B, 3 hours of labor and 2 units of raw material are needed. At present, 1000 hours of labor and 200 units of raw material are available. Substituting the expressions for q_A and q_B, the problem of maximizing the company's total revenue can be formulated as:

$$\begin{array}{ll} \max & 400p_A + 200p_B - 2p_A^2 - p_B^2 + 2p_Ap_B \\ \text{s.t.} & -p_A - p_B \leq -400 \\ & -p_B \leq -600 \end{array}$$

(a) Use the Khun-Tucker conditions to find the company's optimal pricing policy.

(b) What is the maximum the company would be willing to pay for

- another hour of labor,
- another unit of raw material?

Chapter 5

Modeling with Linear Programming

5.1 Introductory Example

SilComputers makes quarterly decisions about their product mix. While their full product line includes hundreds of products, we will consider a simpler problem with just two products: notebook computers and desktop computers. SilComputers would like to know how many of each product to produce in order to maximize profit for the quarter.

There are a number of limits on what SilComputers can produce. The major constraints are as follows:

1. Each computer (either notebook or desktop) requires a Processing Chip. Due to tightness in the market, our supplier has allocated 10,000 such chips to us.

2. Each computer requires memory. Memory comes in 16MB chip sets. A notebook computer has 16MB memory installed (so needs 1 chip set) while a desktop computer has 32MB (so requires 2 chip sets). We received a great deal on chip sets, so have a stock of 15,000 chip sets to use over the next quarter.

3. Each computer requires assembly time. Due to tight tolerances, a notebook computer takes more time to assemble: 4 minutes versus 3 minutes for a desktop. There are 25,000 minutes of assembly time available in the next quarter.

Given current market conditions, material cost, and our production system, each notebook computer produced generates \$750 profit, and each desktop produces \$1000 profit.

There are many questions SilComputer might ask. The most obvious are such things as "How many of each type computer should SilComputer produce in the next quarter?" "What is the maximum profit SilComputer can make?" Less obvious, but perhaps of more managerial interest are "How much should SilComputer be willing to pay for an extra memory chip set?" "What is the effect of losing 1,000 minutes of assembly time due to an unexpected machine failure?" "How much profit would we need to make on a 32MB notebook computer to justify its production?"

Linear programming gives us a mechanism for answering all of these questions quickly and easily. There are three steps in applying linear programming: modeling, solving, and interpreting.

5.1.1 Modeling

We begin by modeling this problem. Modeling a problem using linear programming involves writing it in the language of linear programming. There are rules about what you can and cannot do within linear programming. These rules are in place to make certain that the remaining steps of the process (solving and interpreting) can be successful.

Key to a linear program are the *decision variables*, *objective*, and *constraints*.

Decision Variables. The decision variables represent (unknown) decisions to be made. This is in contrast to *problem data*, which are values that are either given or can be simply calculated from what is given. For this problem, the decision variables are the number of notebooks to produce and the number of desktops to produce. We will represent these unknown values by x_1 and x_2 respectively. To make the numbers more manageable, we will let x_1 be the number of 1000 notebooks produced (so $x_1 = 5$ means a decision to produce 5000 notebooks) and x_2 be the number of 1000 desktops. Note that a value like the quarterly profit is not (in this model) a decision variable: it is an outcome of decisions x_1 and x_2.

Objective. Every linear program has an objective. This objective is to be either minimized or maximized. This objective has to be *linear* in the decision variables, which means it must be the sum of constants times decision variables. $3x_1 - 10x_2$ is a linear function. x_1x_2 is not a linear function. In this case, our objective is to maximize the function $750x_1 + 1000x_2$ (what units is this in?).

Constraints. Every linear program also has constraints limiting feasible decisions. Here we have four types of constraints: Processing Chips, Memory Sets, Assembly, and Nonnegativity.

In order to satisfy the limit on the number of chips available, it is necessary that $x_1 + x_2 \leq 10$. If this were not the case (say $x_1 = x_2 = 6$), the decisions would not be implementable (12,000 chips would be required, though we only have 10,000). Linear programming cannot handle arbitrary restrictions: once again, the restrictions have to be *linear*. This means that a linear function of the decision variables must be *related* to a constant, where *related* can mean less than or equal to, greater than or equal to, or equal to. So $3x_1 - 2x_2 \geq 10$ is a linear constraint, as is $-x_1 + x_3 = 6$. $x_1x_2 \leq 10$ is not a linear constraint, nor is $x_1 + 3x_2 < 3$. Our constraint for Processing Chips $x_1 + x_2 \leq 10$ is a linear constraint.

The constraint for memory chip sets is $x_1 + 2x_2 \leq 15$, a linear constraint.

Our constraint on assembly can be written $4x_1 + 3x_2 \leq 25$, again a linear constraint.

Finally, we do not want to consider decisions like $x_1 = -5$, where production is negative. We add the linear constraints $x_1 \geq 0$, $x_2 \geq 0$ to enforce nonnegativity of production.

Final Model. This gives us the complete model of this problem:

$$
\begin{array}{lrcl}
\text{Maximize} & 750x_1 + 1000x_2 & & \\
\text{Subject to} & & & \\
 & x_1 + x_2 & \leq & 10 \\
 & x_1 + 2x_2 & \leq & 15 \\
 & 4x_1 + 3x_2 & \leq & 25 \\
 & x_1 & \geq & 0 \\
 & x_2 & \geq & 0
\end{array}
$$

Formulating a problem as a linear program means going through the above process to clearly define the decision variables, objective, and constraints.

5.1.2 Solving the Model

Models are useful in their own right: they allow for a formal definition of a problem and can be useful in thinking about a problem. Linear programming models are particularly useful, however, because it is easy to have a computer solve for the optimal decision values.

For a model with only two variables, it is possible to solve the problem without a computer by drawing the feasible region and determining how the objective is optimized on that region. We go through that process here. The purpose of this exercise is to give you intuition and understanding of linear programming models and their solution. In any real application, you would use a computer to solve even two variable problems (we outline how to use Excel's Solver routine to find solutions in the next section).

We can represent a model with two variables by labeling the axes of a graph with each of the variables. The entire graph then represents possible decisions. Constraints are represented by lines on the graph, with the feasible region lying on one side of the line. The following figure illustrates this with the constraint $x_1 + x_2 \leq 10$.

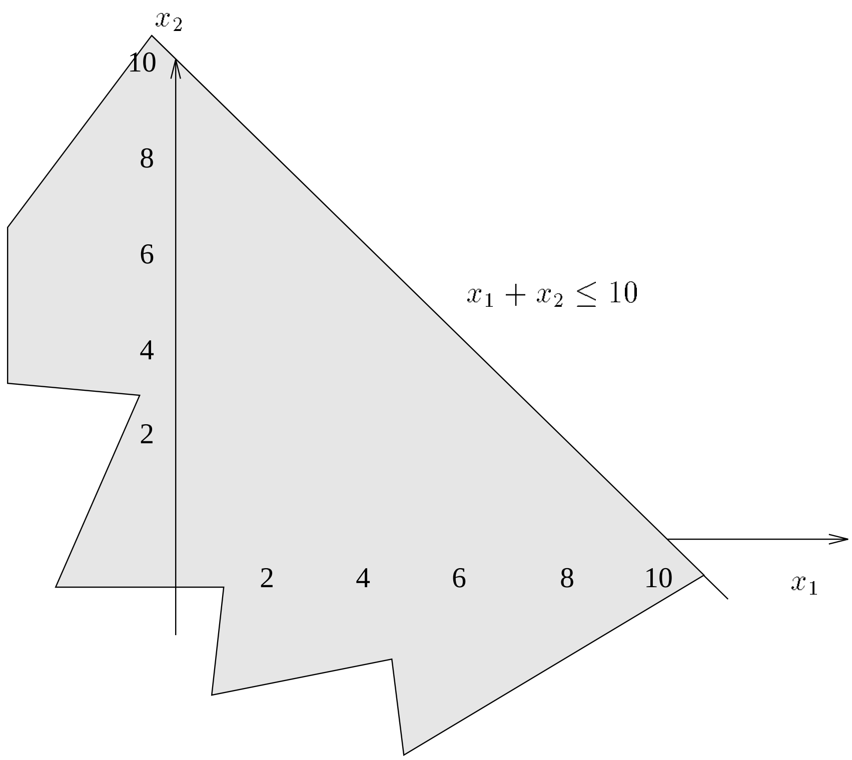


Figure 5.1: Single Constraint

We can continue this process and add in all of the constraints. Since every constraint must be satisfied, the resulting feasible region is the intersection of the feasible region for each constraint. This is shown in the following figure.

Note that just graphing the model gives us information we did not have before. It seems that the Chip constraint ($x_1 + x_2 \leq 10$) plays little role in this model. This constraint is dominated by other constraints.

Now, how can we find the optimal solution? We can include the objective function on this diagram by drawing *iso-profit* lines: lines along which the profit is the same. Since our goal is to maximize the profit, we can push the isoprofit line out until moving it any further would result in no feasible point (see diagram, z represents profit). Clearly the optimal profit occurs at point X.

What are the variable values at point X. Note that X is the intersection of the constraints:

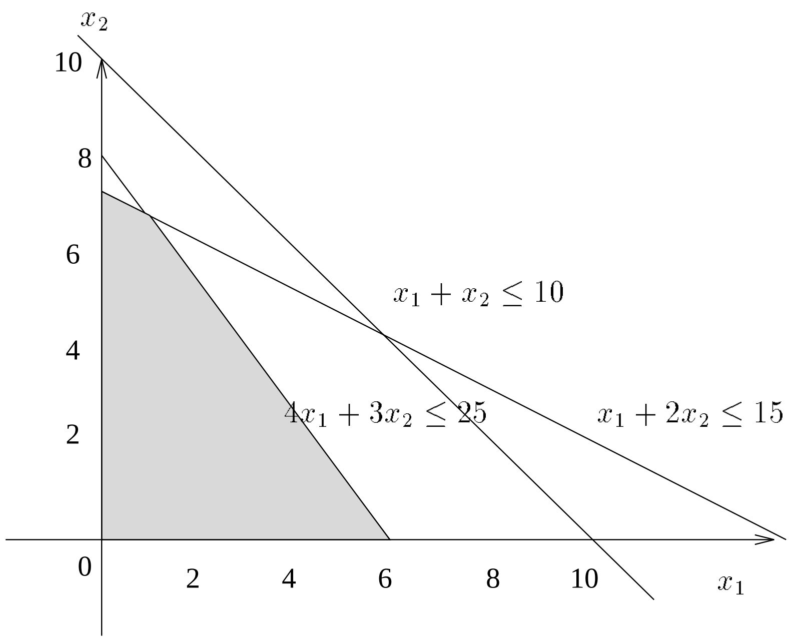


Figure 5.2: All Constraints

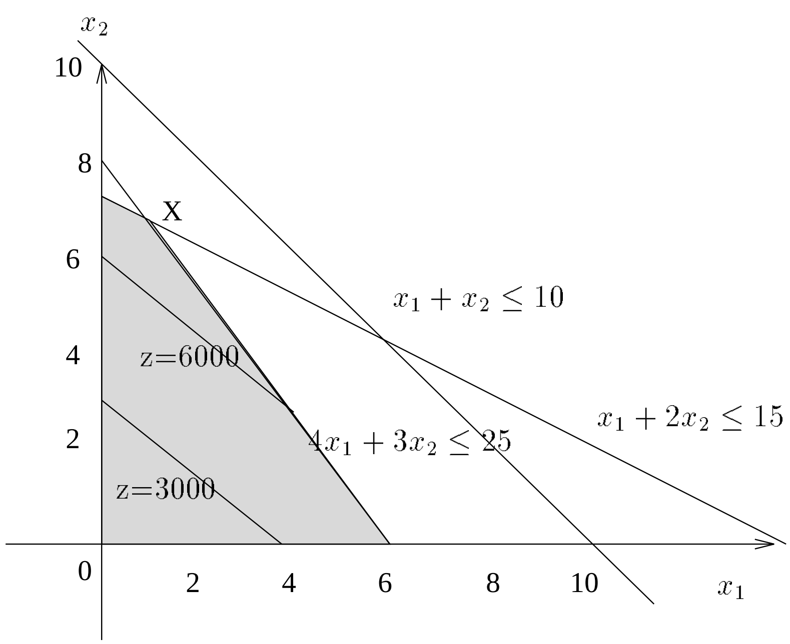


Figure 5.3: Finding Optimal Profit

$$\begin{aligned} x_1 + 2x_2 &= 15 \\ 4x_1 + 3x_2 &= 25 \end{aligned}$$

We discussed in the previous section how to solve such equations: the solution here is $x_1 = 1$ and $x_2 = 7$. The optimal decision is to produce 1,000 notebooks and 7,000 desktops, for a profit of \$7,750,000.

5.1.3 Using Solver

Rather than the somewhat tedious and error-prone graphical method (which is limited to linear programs with 2 variables), special computer programs can be used to find solutions to linear programming models. The most widespread program is undoubted Solver, included in all recent versions of the Excel spreadsheet program. Solver, while not a state of the art code (which can cost upwards of \$15,000 per copy) is a reasonably robust, easy-to-use tool for linear programming. Solver uses standard spreadsheets together with an interface to define variables, objective, and constraints to define a linear program.

It is difficult to describe in words how to create a Solver spreadsheet, so we will do one or two in class. Here is a brief outline and some hints and shortcuts:

- We will start with a spreadsheet that has all of the data entered in some reasonably neat way.
- We will create the model in a separate part of the spreadsheet. We will have one cell for each variable. Solver will eventually put the optimal values in each cell.
- We will have a single cell to represent the objective. We will enter a formula that represents the objective. This formula must be a linear formula, so it must be of the form: cell1*cell1'+cell2*cell2'+..., where cell1, cell2 and so on contain constant values and cell1', cell2' and so are the variable cells.

  Helpful Hint: Excel has a function `sumproduct()` that is designed for linear programs. `sumproduct(a1..a10,b1..b10)` is identical to `a1*b1+a2*b2+a3*b3+...+a10*b10`. This function will save much time and aggravation. All that is needed is that the length of the first range is the same as the length of the second range (so one can be horizontal and the other vertical).

  Helpful Hint: It is possible to assign names to cells and ranges (under the Insert-Name menu). Rather than use `a1..a10` as the variables, you can name that range `var` (for example) and then use `var` wherever `a1..a10` would have been used.

- We then have a cell to represent the left hand side of each constraint (again a linear function) and another cell to represent the right hand side (a constant).
- We then select Solver under the Tools menu. This gives a form to fill out to define the linear program.
- In the "Set Cell" box, select the objective cell. Choose Maximize or Minimize.
- In the "By Changing Cells", put in the range containing the variable cells.

- We next add the constraints. Press the "Add..." button to add constraints. The dialog box has three parts for the left hand side, the type of constraint, and the right hand side. Put the cell references for a constraint in the form, choose the right type, and press "Add". Continue until all constraints are added. On the final constraint, press "OK".

- We need to explicitly include nonnegativity constraints.

  Helpful Hint: It is possible to include ranges of constraints, as long as they all have the same type. `c1..e1 <= c3..e3` means `c1 <= c3`, `d1 <= d3`, `e1 <= e3`. `a1..a10 >= 0` means each individual cell must be greater than or equal to 0.

- Push the options button and toggle the "Assume Linear Model" in the resulting dialog box. This tells Excel to call a linear rather than a nonlinear pgrogramming routine so as to solve the problem more efficiently. This also gives you sensitivity ranges, which are not available for nonlinear models.

- Push the Solve button. In the resulting dialog box, select "Answer" and "Sensitivity". This will put the answer and sensitivity analysis in two new sheets. Ask Excel to "Keep Solver values", and your worksheet will be updated so that the optimal values are in the variable cells.

5.2 Second Example: Market Allocations

MegaMarketing is planning a concentrated one week advertising campaign for their new CutsEverything SuperKnife. The ads have been designed and produced and now they wish to determine how much money to spend in each advertising outlet. In reality, they have hundreds of possible outlets to choose from. We will illustrate their problem with two outlets: Prime-time TV, and newsmagazines.

The problem of optimally spending advertising dollars can be formulated in many ways. For instance, given a fixed budget, the goal might be to maximize the number of target customers reached (a target customer is a customer with a reasonable chance of purchasing the product).

An alternative approach, which we adopt here, is to define targets for reaching each market segment and to minimize the money spent to reach those targets. For this product, the target segments are Teenage Boys, Affluent Women (ages 40-49), and Retired Men. Each minute of primetime TV and page of newsmagazine advertisement reaches the following number of people (in millions):

Outlet	Boys	Women	Men	Cost
TV	5	1	3	600
Mag	2	6	3	500
Target	24	18	24	

Again, MegaMarketing is interested in straightforward answers like how many units of each outlet to purchase to meet the segment goals. They are also interested in such questions as "How much will it cost to reach an extra million retired men?", "One radio spot reaches 1 million boys, 1 million women, and 1 million men: how much are we willing to pay for such a spot?", and similar questions.

5.2.1 Modeling

The process of modeling using linear programming is fairly consistent: define the decision variables, objective, and constraints.

Decision Variables. In this case, the decision variables are the number of units of each outlet to purchase. Let x_1 be the number of TV spots, and x_2 the number of magazine pages.

Objective. Our objective is to minimize the amount we spend: $600x_1 + 500x_2$.

Constraints. We have one constraint for each market segment: we must be certain that we reach sufficient people of each segment. For boys, this corresponds to the constraint:

$$5x_1 + 2x_2 \geq 24$$

Similar constraints for the other segments gives us the full formulation as:

$$\begin{array}{rrcl} \text{Minimize} & 600x_1 + 500x_2 & & \\ \text{Subject to} & & & \\ & 5x_1 + 2x_2 & \geq & 24 \\ & x_1 + 6x_2 & \geq & 18 \\ & 3x_1 + 3x_2 & \geq & 24 \\ & x_1 & \geq & 0 \\ & x_2 & \geq & 0 \end{array}$$

5.2.2 Solving the Model

Since this model has only two variables, we can solve the problem in two ways: graphically and using SOLVER (with SOLVER being much easier). We start by solving this graphically.

The first step is to graph the feasible region, as given in the following figure:

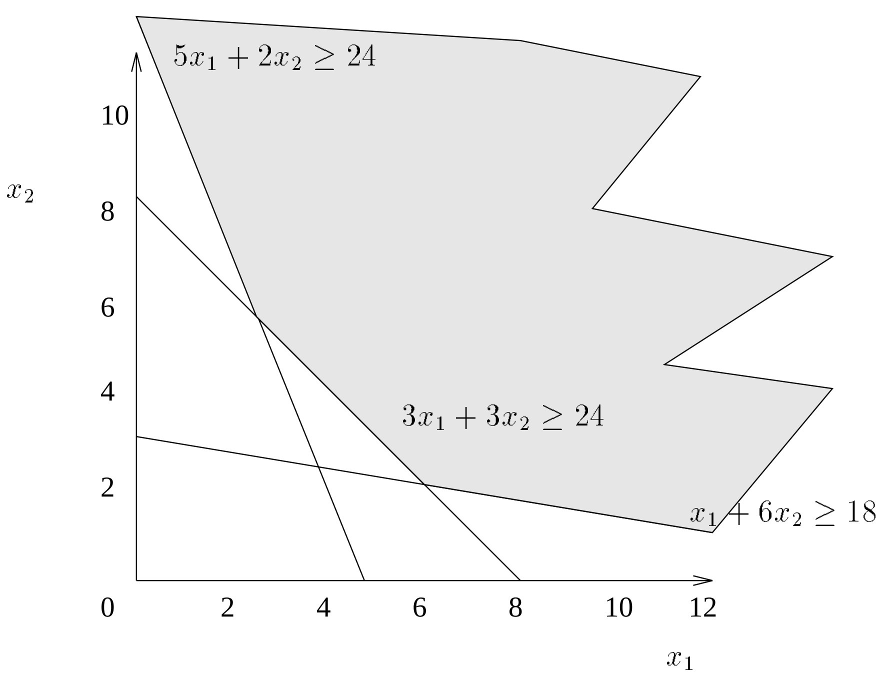


Figure 5.4: Marketing example

The next step is to put some iso-cost lines on the diagram: lines representing points that have the same cost. We mark the optimal point with an X.

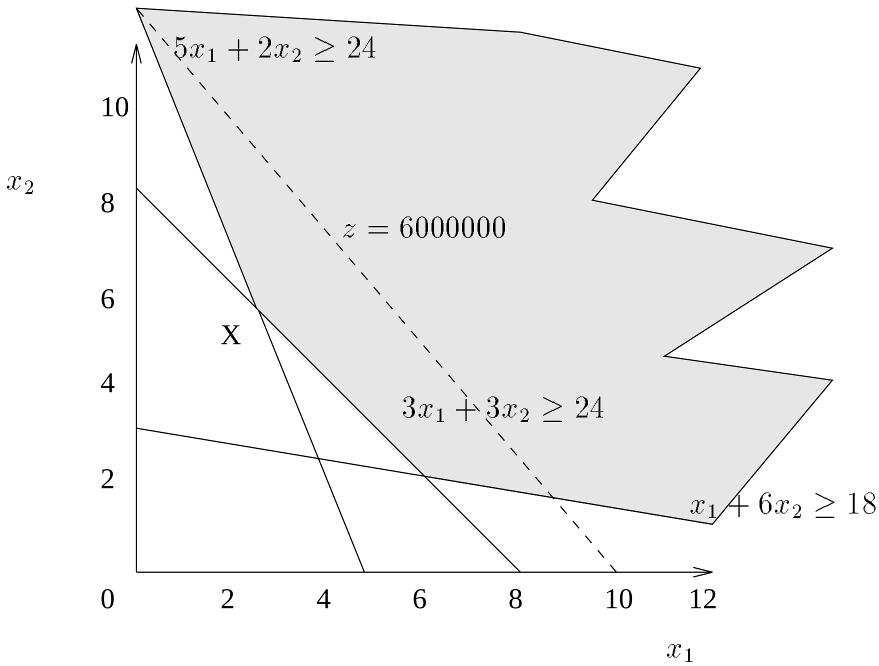


Figure 5.5: Marketing example with Isocost Line

X is the intersection of the constraints:

$$\begin{aligned} 5x_1 + 2x_2 &= 24 \\ 3x_1 + 3x_2 &= 24 \end{aligned}$$

The optimal solution is $x_1 = 2.67$ and $x_2 = 5.33$

5.2.3 Using Solver

The SOLVER model for this problem is on the course web pages.

5.3 Common Features

Hidden in our models of these problems are a number of assumptions. The usefulness of a model is directly related to how close reality matches up with these assumptions.

The first assumption is due to the linear form of our functions. Since the objective is linear, the contribution to the objective of any decision variable is proportional to the value of the decision variable. Producing twice as much of a product produces twice as much profit; buying twice as many pages of ads costs twice as much. This is the *Proportionality Assumption.*

Furthermore, the contribution of a variable to the objective is independent of the values of the other variables. One notebook computer is worth \$750, independent of how many desktop computers we produce. This is the *Additivity Assumption.*

Similarly, since each constraint is linear, the contribution of each variable to the left hand side of each constraint is proportional to the value of the variable and independent of the values of any other variable.

These assumptions are quite restrictive. We will see, however, that clever modeling can handle situations that may appear to violate these assumptions.

The next assumption is the *Divisibility Assumption*: it is possible to take any fraction of any variable. Rethinking the Marketing example, what does it mean to purchase 2.67 television ads? It may be that the divisibility assumption is violated in this example. Or, it may be that the units are such that 2.67 "ads" actually corresponds to 2666.7 minutes of ads, in which case we can "round off" our solution to 2667 minutes with little doubt that we are getting an optimal or nearly optimal solution. Similarly, a fractional production quantity may be worisome if we are producing a small number of battleships or be innocuous if we are producting millions of paperclips. If the Divisibility Assumption is important and does not hold, then a technique called *integer programming* rather than linear programming is required. This technique takes orders of magnitude more time to find solutions but may be necessary to create realistic solutions. You will learn more about this in 45-761.

The final assumption is the *Certainty Assumption*: linear programming allows for no uncertainty about the numbers. An ad will reach the given number of people; the number of assembly hours we give will certainly be available.

It is very rare that a problem will meet all of the assumptions exactly. That does not negate the usefulness of a model. A model can still give useful managerial insight even if reality differs slightly from the rigorous requirements of the model. For instance, the knowledge that our chip inventory is more than sufficient holds in our first model even if the proportionality assumptions are not satisfied completely.

5.4 Linear Programming Models

Linear programming models are found in almost every field of business (and beyond!). The next sections go through a number of problems, showing how to model them by the appropriate choice of decision variables, objective, and constraints. In all cases, we will describe the problem and give a model. Solver worksheets for each of these is available on the web page.

5.4.1 Diet Problem

Problem Definition

What is the perfect diet? An ideal diet would meet or exceed basic nutritional requirements, be inexpensive, have variety, and be "pleasing to the palate". How can we find such a diet?

Suppose the only foods in the world are as follows:

Food	Serving Size	Energy (kcal)	Protein (g)	Calcuim (mg)	Price (cents/serving)	Limit (servings/day)
Oatmeal	28g	110	4	2	3	4
Chicken	100g	205	32	12	24	3
Eggs	2 Large	160	13	54	13	2
Whole Milk	237cc	160	8	285	9	8
Cherry Pie	170g	420	4	22	20	2
Pork&Beans	260g	260	14	80	19	2

After consulting with nutitionists, we decree that a satisfactory diet has at least 2000 kcal of energy, 55 g of protein, and 800 mg of calcium (vitamins and iron are supplied by pills). While some of us would be happy to subsist on 10 servings of pork and beans, we have decided to impose variety by having a limit on the number of servings/day for each of our six foods. What is the least cost satisfactory diet?

Problem Modeling

First we decide on decision variables. Let us label the foods 1, 2, ...6, and let x_i represent the number of servings of food i in the diet.

Our objective is to minimize cost, which can be written

$$3x_1 + 24x_2 + 13x_3 + 9x_4 + 20x_5 + 19x_6$$

We have constraints for energy, protein, calcuim, and for each serving/day limit. This gives the formulation:

$$\begin{array}{rrcl}
\text{Minimize} & 3x_1 + 24x_2 + 13x_3 + 9x_4 + 20x_5 + 19x_6 & & \\
\text{Subject to} & & & \\
& 110x_1 + 205x_2 + 160x_3 + 160x_4 + 420x_5 + 260x_6 & \geq & 2000 \\
& 4x_1 + 32x_2 + 13x_3 + 8x_4 + 4x_5 + 14x_6 & \geq & 55 \\
& 2x_1 + 12x_2 + 54x_3 + 285x_4 + 22x_5 + 80x_6 & \geq & 800 \\
& x_1 & \leq & 4 \\
& x_2 & \leq & 3 \\
& x_3 & \leq & 2 \\
& x_4 & \leq & 8 \\
& x_5 & \leq & 2 \\
& x_6 & \leq & 2 \\
& x_i & \geq & 0 \text{ (for all } i)
\end{array}$$

Discussion

The creation of optimal diets was one of the first uses of linear programming. Some difficulties with linear programming include difficulties in formulating "palatability" requirements, and issues of divisibility (no one wants to eat half a green bean). These linear programming models do give an idea on how much these palatability requirements are costing.

5.4.2 Workforce Planning

Problem Definition.

Consider a restaurant that is open seven days a week. Based on past experience, the number of workers needed on a particular day is given as follows:

Day	Mon	Tue	Wed	Thu	Fri	Sat	Sun
Number	14	13	15	16	19	18	11

Every worker works five consecutive days, and then takes two days off, repeating this pattern indefinitely. How can we minimize the number of workers that staff the restaurant?

Model.

A natural (and wrong!) first attempt at this problem is to let x_i be the number of people working on day i. Note that such a variable definition does not match up with what we need to find. It does us no good to know that 15 people work Monday, 13 people Tuesday, and so on because it does not tell us how many workers are needed. Some workers will work both Monday and Tuesday, some only one day, some neither of those days. Instead, let the days be numbers 1 through 7 and let x_i be the number of workers who begin their five day shift on day i. Our objective is clearly:

$$x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7$$

Consider the constraint for Monday's staffing level of 14. Who works on Mondays? Clearly those who start their shift on Monday (x_1). Those who start on Tuesday (x_2) do not work on Monday, nor do those who start on Wednesday (x_3). Those who start on Thursday (x_4) do work on Monday, as do those who start on Friday, Saturday, and Sunday. This gives the constraint:

$$x_1 + x_4 + x_5 + x_6 + x_7 \geq 14$$

Similar arguments give a total formulation of:

$$\begin{array}{lrcl}
\text{Minimize} & \sum_i x_i & & \\
\text{Subject to} & & & \\
& x_1 + x_4 + x_5 + x_6 + x_7 & \geq & 14 \\
& x_1 + x_2 + x_5 + x_6 + x_7 & \geq & 13 \\
& x_1 + x_2 + x_3 + x_6 + x_7 & \geq & 15 \\
& x_1 + x_2 + x_3 + x_4 + x_7 & \geq & 16 \\
& x_1 + x_2 + x_3 + x_4 + x_5 & \geq & 19 \\
& x_2 + x_3 + x_4 + x_5 + x_6 & \geq & 18 \\
& x_3 + x_4 + x_5 + x_6 + x_7 & \geq & 11 \\
& x_i & \geq & 0 \text{ (for all } i\text{)}
\end{array}$$

Discussion.

Workforce modeling is a well developed area. Note that our model has only one type of shift, but the model is easily extended to other types of shifts, with differing shift costs.

5.4.3 Financial Portfolio

Problem Definition.

In your finance courses, you will learn a number of techniques for creating "optimal" portfolios. The optimality of a portfolio depends heavily on the model used for defining risk and other aspects of financial instruments. Here is a particularly simple model that is amenable to linear programming techniques.

Consider a mortgage team with \$100,000,000 to finance various investments. There are five categories of loans, each with an associated return and risk (1-10, 1 best):

Loan/investment	Return (%)	Risk
First Mortgages	9	3
Second Mortgages	12	6
Personal Loans	15	8
Commercial Loans	8	2
Government Securities	6	1

Any uninvested money goes into a savings account with no risk and 3% return. The goal for the mortgage team is to allocate the money to the categories so as to:

(a) Maximize the average return per dollar

(b) Have an average risk of no more than 5 (all averages and fractions taken over the invested money (not over the saving account)).

(c) Invest at least 20% in commercial loans

(d) The amount in second mortgages and personal loans combined should be no higher than the amount in first mortgages.

Model

Let the investments be numbered 1...5, and let x_i be the amount invested in investment i. Let x_s be the amount in the savings account. The objective is to maximize

$$9x_1 + 12x_2 + 15x_3 + 8x_4 + 6x_5 + 3x_s$$

subject to

$$x_1 + x_2 + x_3 + x_4 + x_5 + x_s = 100,000,000.$$

Now, let's look at the average risk. Since we want to take the average over only the invested amount, a direct translation of this constraint is

$$\frac{3x_1 + 6x_2 + 8x_3 + 2x_4 + x_5}{x_1 + x_2 + x_3 + x_4 + x_5} \leq 5$$

This constraint is not linear, but we can cross multiply, and simplify to get the equivalent linear constraint:

$$-2x_1 + x_2 + 3x_3 - 3x_4 - 4x_5 \leq 0$$

Similarly we need

$$x_4 \geq 0.2(x_1 + x_2 + x_3 + x_4 + x_5)$$

or

$$-0.2x_1 - 0.2x_2 - 0.2x_3 + 0.8x_4 - 0.2x_5 \geq 0$$

The final constraint is

$$x_2 + x_3 - x_1 \leq 0$$

Together with nonnegativity, this gives the entire formulation.

Discussion

Optimal portfolios do not just happen: they must be calculated, and there is a constant interplay between models and solvability. Linear programming models provide great modeling power with a great limit: the handling of risk must be done in a linear fashion (like our Risk factors here). Other models you will see in finance will look at the co-variance of returns between investments, a fundamentally nonlinear effect. This can give rise to nonlinear models like those that try to minimize variance subject to return requirements. It is very difficult to embed idiosyncratic constraints (like (c) and (d) here) in such models.

5.4.4 Production over Time

Problem Description

Many problems are multiperiod, where a series of decisions have to be made over time. For instance, in our very first SilComputer example, production decisions for computers have to be made with the future in mind. In this example, it would be easy to work with both desktop and notebook computers but let's work just with notebook computers.

SilComputer needs to meet the demand of its largest corporate and educational customers for notebook computers over the next four quarters (before its current model becomes obsolete). SilComputer currently has 5,000 notebook computers in inventory. Expected demand over the next four quarters for its notebook is 7,000; 15,000; 10,000; and 8,000. SilComputer has sufficient capacity and material to produce up to 10,000 computers in each quarter at a cost of $2000 per notebook. By using overtime, up to an additional 2,500 computers can be produced at a cost of $2200 each. Computers produced in a quarter can be used either to meet that quarter's demand, or be held in inventory for use later. Each computer in inventory is charged $100 to reflect carying costs.

How should SilComputer meet its demand for notebooks at minimum cost?

Modeling.

The decisions seem to be how many of each computer to produce in each period at regular time, how many to produce at overtime, and how much inventory to carry in each period. Let's denote our time periods $t = 1, 2, 3, 4$. Let x_t be the number of notebooks produced in period t at regular time and y_t be the number of notebooks produced in period t at overtime. Finally, let i_t be the inventory at the end of period t.

Look at the first quarter: how are these variables restricted and related?

Clearly we need $x_1 \leq 10000$ and $y_1 \leq 2500$. Now, anything that starts as inventory or is produced in the period must either be used to meet demand or ends up as inventory at the end of period 1. This means:

$$5000 + x_1 + y_1 = 7000 + i_1$$

For period 2, in addition to the upper bounds, we get the constraint

$$i_1 + x_2 + y_2 = 15000 + i_2$$

For period 3, we get

$$i_2 + x_3 + y_3 = 10000 + i_3$$

and for period 4 (assuming no inventory at the end):

$$i_3 + x_4 + y_4 = 8000$$

Our objective charges 2000 for each x, 2200 for each y, and 100 for each i:

$$2000x_1 + 2000x_2 + 2000x_3 + 2000x_4 + 2200y_1 + 2200y_2 + 2200y_3 + 2200y_4 + 100i_1 + 100i_2 + 100i_3$$

Discussion.

The constraints tracking where production goes are known as *conservation* constraints, and model the requirement that production never disappears. Whenever you have multiple types of variables in a formulation, be certain to think about whether there are necessary constraints that link them.

5.4.5 Investing over Time

We are going to manage an investment portfolio over a 6-year time horizon. We begin with \$1000, and at various times we can invest in one or more of the following:

Savings account X, annual yield 5%.
Security Y, 2-year maturity, total yield 12% if bought now, 11% thereafter.
Security Z, 3-year maturity, total yield 18%.
Security W, 4-year maturity, total yield 24%.

To keep things simple we will assume that each security can be bought in any denomination. (This assumption can be relaxed if one uses *integer* or *dynamic* programming .) We can make savings deposits or withdrawals anytime. We can buy Security Y any year but year 3. We can buy Security Z anytime after the first year. Security W, now available, is a one-time opportunity.

We let x_t be the amount of money invested in the savings account X at the beginning of year t, and similarly for y_t, z_t, and w_t. We will put any money not tied up in securities into the savings account. The situation is summed up in the figure.

The problem is really a kind of inventory problem. In a given year, the amount of money carried forward from the previous year (in savings), plus the yield from securities that mature that year, equals the amount invested in new securities plus the amount of money left over for next year. There is one twist: inventory grows while in storage. (In real inventory problems, stock often decreases over time due to spoilage, etc., and this can be reflected in negative "interest" rates.) Let v be the final yield when all securities are cashed in at the end of the sixth year. Then if the objective is to maximize final yield, the LP is,

$$\begin{aligned}
\max \quad & v \\
\text{s.t.} \quad & x_1 + y_1 + w_1 = 1000 \\
& x_2 + y_2 + z_2 = 1.05x_1 \\
& x_3 + z_3 = 1.05x_2 + 1.12y_1 \\
& x_4 + y_4 + z_4 = 1.05x_3 + 1.11y_2 \\
& x_5 + y_5 = 1.05x_4 + 1.18z_2 + 1.24w_1 \\
& x_6 = 1.05x_5 + 1.11y_4 + 1.18z_3 \\
& v = 1.05x_6 + 1.11y_5 + 1.18z_4, \\
& x_t, y_t, z_t, w_t \geq 0
\end{aligned}$$

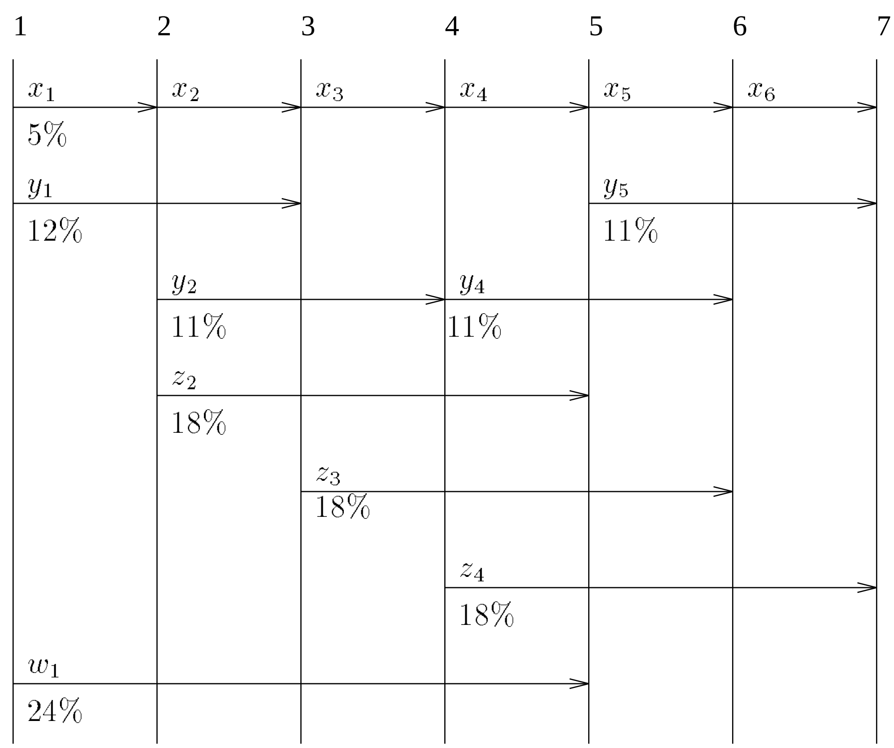


Figure 5.6: Possible investments over a 6-year horizon

The final assets come to \$1387.6, an average annual yield of 5.61%

5.4.6 Problems

Modeling is an art, and it takes practice. The following examples show the variety of problems that can be attacked by linear programming, and give you the opportunity to try your hand at some problems.

Exercise 44 A plant produces two types of refrigerators, A and B. There are two production lines, one dedicated to producing refrigerators of Type A, the other to producing refrigerators of Type B. The capacity of the production line for A is 60 units per day, the capacity of the production line for B is 50 units per day. A requires 20 minutes of labor whereas B requires 40 minutes of labor. Presently, there is a maximum of 40 hours of labor per day which can be assigned to either production line. Profit contributions are \$20 per refrigerator of Type A produced and \$30 per Type B produced. What should the daily production be? Solve graphically and by Solver.

Exercise 45 Albert, Bill, Charles, David, and Edward have gotten into a bind. After a series of financial transactions, they have ended up each owing some of the others huge amounts of money. In fact, near as the lawyers can make out, the debts are as follows

Debtor	Creditor	Amount (\$millions)
A	E	10
A	C	3
B	A	5
C	B	6
C	D	4
D	A	4
E	C	7
E	D	3

The question is, who is bankrupt? We will say that a person i is bankrupt if there is no possible transfer of funds among the people such that i completely pays off his obligations, and the transfer of funds satisfies the following condition: for every two persons j and k, the amount paid by person j to person k is no greater than the debt of j to k. For instance, Albert is bankrupt since he owes 10, and is only owed 9. Formulate the problem of determining whether Bill is bankrupt as a linear program. Then modify your formulation to determine if each of the others is bankrupt. *This example may look contrived, but it is inspired by a solution to the debts involved in a crash of Kuwait's al-Mankh stock market.*

Exercise 46 Due to an unexpected glut of orders, Blaster Steel has decided to hire temporary workers for a five day period to clear out the orders. Each temporary worker can work either a two day shift or a three day shift for this period (shifts must be consecutive days). At least 10 workers are needed on days 1, 3, 5, and at least 15 workers are needed on days 2 and 4. A worker on a two day shift gets paid \$125/day, while those on a three day shift gets paid \$100/day.

(a) Formulate the problem of hiring temporary workers to minimize cost while meeting the demand for workers.

(b) Due to a limited number of training personnel, no more than 10 workers can start their shift on any day. Update your formulation in (a) to take this into account.

(c) Union regulations require that at least half of all money spent on workers go to those who work three day shifts. Update your formulation in (a) to handle this requirement.

(d) There are four people who are willing to work a shift consisting of days 1, 2, and 5, for a payment of $110/day. Update your formulation in (a) to handle this possibility.

Exercise 47 You have $1000 to invest in Securities 1 and 2; uninvested funds are deposited in a savings account with annual yield of 5%. You will sell the securities when they reach maturity. Security 1 matures after 2 years with a total yield of 12%. Security 2 matures after 3 years with a total yield of 19%. Your planning horizon is 7 years, and you want to maximize total assets at the end of the seventh year. Suppose either security may be purchased in arbitrarily small denominations.

a) Let x_{it} be the amount invested in Security i at the beginning of year t; the savings account can be considered Security 0. Write the appropriate LP model.

b) Solve the problem by computer and indicate your optimal portfolio in each year 1,...,7.

Exercise 48 Chemco produces two chemicals: A and B. These chemicals are produced via two manufacturing processes. Process 1 requires 2 hours of labor and 1 lb. of raw material to produce 2 oz. of A and 1 oz. of B. Process 2 requires 3 hours of labor and 2 lb. of raw material to produce 3 oz. of A and 2 oz. of B. Sixty hours of labor and 40 lb. of raw material are available. Chemical A sells for $16 per oz. and B sells for $14 per oz. Formulate a linear program that maximizes Chemco's revenue.

Hint: Define the amounts of Process 1 and Process 2 used, as your decision variables.

Exercise 49 Gotham City National Bank is open Monday-Friday from 9 am to 5 pm. From past experience, the bank knows that it needs the following number of tellers.

Time Period	Tellers Required
9 – 10	4
10 – 11	3
11 – noon	4
noon – 1	6
1 – 2	5
2 – 3	6
3 – 4	8
4 – 5	8

The bank hires two types of tellers. Full-time tellers work 9-5 five days a week, except for 1 hour off for lunch, either between noon and 1 pm or between 1 pm and 2 pm. Full-time tellers are paid (including fringe benefits) $25/hour (this includes payment for lunch hour). The bank may also hire up to 3 part-time tellers. Each part-time teller must work exactly 4 consecutive hours each day. A part-time teller is paid $20/hour and receives no fringe benefits. Formulate a linear program to meet the teller requirements at minimum cost.

Exercise 50 Sales forecasts for the next four months are (in thousand of units):

October	10
November	16
December	10
January	12

September's production was set at 12,000 units. Varying production rate incurs some cost: production can be increased from one month to the next at a cost of $2 per unit and it can be decreased at a cost of \$0.50 per unit. In addition, inventory left at the end of a month can be stored at a cost of \$1 per unit per month. Given current demand, there will be no inventory at the end of September. No inventory is desired at the end of January. Formulate a linear program that minimizes the total cost (varying production rate + inventory costs) of meeting the above demand.

Exercise 51 An oil company blends gasoline from three ingredients: butane, heavy naphta and catalytic reformate. The characteristics of the ingredients as well as minimum requirements for regular gasoline are given below:

	Butane	Catalytic Reformate	Heavy Naphta	Gasoline
Octane	120	100	74	≥ 89
Vapor Pressure	60	2.5	4.0	≤ 11
Volatility	105	3	12	≥ 17

The cost (per gallon) of butane is \$0.58, it is \$1.55 for catalytic reformate and \$0.85 for heavy naphta. How many gallons of the three ingredients should be blended in order to produce 12,000 gallons of gasoline at minimum cost?

Exercise 52 An electric utility has six power plants on the drawing board. The anticipated useful life of these plants is 30 years for the coal-fired plants (plants 1, 2 and 3) and 40 years for the fuel-fired plants (plants 4, 5 and 6). Plants 1, 2 and 4 will be on line in year 5. Plants 3 and 5 in year 15. Plant 6 in year 25. The cost of installing generating capacity at Plant i, discounted to the present, is c_i per megawatt, for $i = 1, 2, \ldots, 6$. Projected power demand in year t is D_t megawatts, for $t = 10, 20, 30, 40$. Due to environmental regulations, the fraction of generating capacity at coal-fired plants in year t, relative to total generating capacity in year t, can be at most r_t for $t = 10, 20, 30, 40$.

The generating capacity of each plant has yet to be determined. Write a linear program that assigns capacities to the six plants so as to minimize total present cost of installing generating capacity, while meeting demand and satisfying the environmemtal constraints in years $t = 10, 20, 30, 40$.

Exercise 53 (Wagner) You must decide how many tons x_1 of pure steel and how many tons x_2 of scrap metal to use in manufacturing an alloy casting. Pure steel costs \$300 per ton and scrap \$600 per ton (larger because the impurities must be skimmed off). The customer wants at least 5 tons but will accept a larger order, and material loss in melting and casting is negligible. You have 4 tons of steel and 7 tons of scrap to work with, and the weight ratio of scrap to pure cannot exceed 7/8 in the alloy. You are allotted 18 hours melting and casting time in the mill; pure steel requires 3 hrs per ton and scrap 2 hrs per ton.

a) Write a linear programming model for the problem assuming that the objective is to minimize the total steel and scrap costs.

b) Graph the problem obtained in (a).

c) Solve the problem in (a) using Solver and indicate the optimal values of x_1, x_2. What is the weight ratio of scrap to pure steel?

Exercise 54 (Wagner) An airline must decide how many new flight attendants to hire and train over the next six months. The staff requirements in person-flight-hours are respectively 8000, 9000, 7000, 10,000, 9000, and 11,000 in the months January through June. A new flight attendant is trained for one month before being assigned to a regular flight and therefore must be hired a month before he or she is needed. Each trainee requires 100 hours of supervision by experienced flight attendants during the training month, so that 100 fewer hours are available for flight service by regular flight attendants.

Each experienced flight attendant can work up to 150 hours a month, and the airline has 60 regular flight attendants available at the beginning of January. If the maximum time available from experienced flight attendants exceeds a month's flying and training requirements, they work fewer hours, and none are laid off. At the end of each month, 10% of the experienced flight attendants quit their jobs.

An experienced flight attendant costs the company \$1700 and a trainee \$900 a month in salary and benefits.

a) Formulate the problem as a linear programming model. Let x_t be the number of flight attendants that begin training in month t. Hint. Let y_t be the number of experienced flight attendants available in month t; also let $x_0 = 0$, $y_1 = 60$. The problem is really an inventory problem, with two kinds of stock: trainees and experienced employees. The "holding costs" are the salaries, and the demands are the number of flight hours needed. "Spoilage" is attrition of experienced staff.

b) Solve the model with Solver and indicate the solution value for the x_t's and y_t's.

Exercise 55 (Wagner) A lumber company operates a sawmill that converts timber to lumber or plywood. A marketable mix of 1000 board feet of lumber products requires 1000 board feet of spruce and 4000 board feet of Douglas fir. Producing 1000 square feet of plywood requires 2000 board feet of spruce and 4000 board feet of fir. The company's timberland yields 32,000 board feet of spruce and 72,000 board feet of fir each season.

Sales commitments require that at least 5000 board feet of lumber and 12,000 board feet of plywood be produced during the season. The profit contributions are \$45 per 1000 board feet of lumber and \$60 per 1000 square feet of plywood.

a) Express the problem as a linear programming model.

b) Graph the problem and indicate the optimal solution on the graph.

Exercise 56 (Wagner) An electronics firm manufactures radio models A, B and C, which have profit contributions of 16, 30 and 50 respectively. Minimum production requirements are 20, 120 and 60 for the three models, respectively.

A dozen units of Model A requires 3 hours for manufacturing of components, 4 for assembling, and 1 for packaging. The corresponding figures for a dozen units of Model B are 3.5, 5, and 1.5, and for Model C are 5, 8, and 3. During the forthcoming week the company has available 120 plant-hours for manufacturing components, 160 for assembly, and 48 for packaging. Formulate this production planning problem as a linear programming problem.

Exercise 57 (Wagner) Process 1 in an oil refinery takes a unit feed of 1 bbl (barrel) of crude oil A and 3 bbl of crude B to make 50 gallons (gal) of gasoline X and 20 gal of gasoline Y. From a unit feed of 4 bbl crude A and 2 bbl crude B, process 2 makes 30 gal of X and 80 gal of Y. Let x_i be the number of units of feed type i; e.g., $x_1 = 1$ indicates that process 1 uses 1 bbl crude A and 3 bbl crude B.

There are 120 bbl of crude A available and 180 bbl of crude B. Sales commitments call for at least 2800 gal of gasoline X and 2200 gal of Y. The unit profits of process 1 and 2 are p_1 and p_2, respectively. Formulate an LP model.

Exercise 58 (Wagner) An air cargo firm has 8 aircraft of type 1, 15 of type 2 and 11 of type 3 available for today's flights. A type 1 craft can carry 45 tons, type 2, 7 tons and type 3, 5 tons. 20 tons of cargo are to be flown to city A and 28 tons to city B. Each plane makes at most one flight a day.

The costs of flying a plane from the terminal to each city are as follows.

	Type 1	Type 2	Type 3
City A	23	15	1.4
City B	58	20	3.8

Let x_i be the number of type i planes sent to A and y_i the number to B. Formulate an LP for this routing problem.

Exercise 59 (Wagner) A manufacturing firm produces widgets and distributes them to five wholesalers at a fixed delivered price of \$2.50 per unit. Sales forecasts indicate that monthly deliveries will be 2700, 2700, 9000, 4500 and 3600 widgets to wholesalers 1-5 respectively. The monthly production capacities are 4500, 9000 and 11,250 at plants 1, 2 and 3, respectively. The direct costs of producing each widget are \$2 at plant 1, \$1 at plant 2 and \$1.80 at plant 3.

The transport cost of shipping a widget from a plant to a wholesaler is given below.

Wholesaler	1	2	3	4	5
Plant 1	.05	.07	.11	.15	.16
Plant 2	.08	.06	.10	.12	.15
Plant 3	.10	.09	.09	.10	.16

Formulate an LP model for this production and distribution problem.

Exercise 60 (Wagner) Ft. Loudoun and Watts Bar are two large hydroelectric dams, the former upstream of the latter. The level of Watts Bar Lake must be kept within limits for recreational purposes, and the problem is to plan releases from Ft. Loudoun to do so. In reality this problem is solved simultaneously for numerous dams covering an entire watershed, but we focus on a single reservoir. There are also sophisticated models for predicting runoff into the reservoirs, but we will suppose that runoff is negligible. Thus any water entering Watts Bar Lake must be released through Ft. Loudoun Dam.

The planning period is 20 months. During month t, let x_t be the average water level of Watts Bar Lake before augmentation by water from Ft. Loudoun; $x_1 = 25$. Let y_t be the number of feet added to the average level in month t from Ft. Loudoun. L_t and U_t are the lower and upper bounds on the lake level in month t (more restrictive in summer). To model seepage, evaporation and hydroelectric release through Watts Bar Dam we suppose that Watts Bar Lake begins month $t+1$ at a level equal to .75 times the average level of the previous month (including the augmentation from Ft. Loudoun). The cost of water from Ft. Loudoun Lake is c_t for every foot added to the level of Watts Bar. Formulate the appropriate LP model. (In reality the model is a huge nonlinear program.)

Exercise 61 Red Dwarf Toasters needs to produce 1000 of their new "Talking Toaster". There are three ways this toaster can be produced: manually, semi-automatically, and robotically. Manual assembly requires 1 minute of skilled labor, 40 minutes of unskilled labor, and 3 minutes of assembly room time. The corresponding values for semiautomatic assembly are 4, 30, and 2; while those for robotic assembly are 8, 20, and 4. There are 4500 minutes of skilled labor, 36,000 minutes of unskilled labor, and 2700 minutes of assembly room time available for this product. The total cost for producing manually is \$7/toaster; semiautomatically is \$8/toaster; and robotically is \$8.50/toaster.

(a) Formulate the problem of producing 1000 toasters at minimum cost meeting the resource requirements. Clearly define your variables, objective and constraints.

(b) Our union contract states that the amount of skilled labor time used is at least 10% of the total labor (unskilled plus skilled) time used. Update your formulation in (a) to handle this requirement.

(c) Any unused assembly floor time can be rented out at a profit of \$0.50/minute. Update your formulation to include this possibility.

Answers to Exercise 61:

(a) *Let x_1 be the number of toasters produced manually, x_2 be the number produced semiautomatically, and x_3 be the number produced robotically.*

The objective is to Minimize $7x_1 + 8x_2 + 8.5x_3$.

The constraints are:

$x_1 + x_2 + x_3 = 1000$ *(produce enough toasters)*

$x_1 + 4x_2 + 8x_3 \leq 4500$ *(skilled labor used less than or equal to amount available).*

$40x_1 + 30x_2 + 20x_3 \leq 36000$ *(unskilled labor constraint)*

$3x_1 + 2x_2 + 4x_3 \leq 2700$ *(assembly time constraint)*

$x_1 \geq 0, x_2 \geq 0, x_3 \geq 0$ *(nonnegativity of production)*

(b) *Add a constraint* $x_1 + 4x_2 + 8x_3 \geq .1(41x_1 + 34x_2 + 28x_3)$

(c) *Add a variable s_a to represent the assembly time slack. Add $+0.5s_a$ to the objective. Change the assembly time constraint to*

$3x_1 + 2x_2 + 4x_3 + s_a = 2700$ *(assembly time constraint)*

$s_a \geq 0$

Chapter 6

Case Study

The Global Oil Company

You may complete the project individually or in groups. In the latter case, which is encouraged, members of the group submit a joint report and all receive the same grade. The ideal size of a group is three. Groups of up to four are allowed. There should be no collaboration among groups and/or students working individually.

The Global Oil Company is an international producer, refiner, transporter and distributor of oil, gasoline and petrochemicals. Global Oil is a holding company with subsidiary operating companies that are wholly or partially owned. A major problem for Global Oil is to coordinate the actions of these various subsidiaries into an overall corporate plan, while at the same time maintaining a reasonable amount of operating autonomy for the subsidiary companies.

To deal with this dilemma, the logistics department at Global Oil Headquarters develops an annual corporate-wide plan which details the pattern of shipments among the various subsidiaries. The plan is not rigid but provides general guidelines and the plan is revised periodically to reflect changing conditions. Within the framework of this plan, the operating companies can make their own decisions and plans. This corporate-wide plan is presently done on a trial and error basis. There are two problems with this approach. First, the management of the subsidiaries often complains that the plan does not reflect properly the operating conditions under which the subsidiary operates. The plan sometimes calls for operations or distribution plans that are impossible to accomplish. And secondly, corporate management is concerned that the plan does not optimize for the total company.

The technique of linear programming seems a possible approach to aid in the annual planning process, that will be able to answer at least in part, the two objections above. In addition the building of such a model will make it possible to make changes in plans quickly when the need arises. Before embarking on the development of a world-wide model, Global Oil asks you to build a model of the Far Eastern operations for the coming year.

Far Eastern Operations

The details of the 1998 planning model for the Far Eastern Operations are now described.

There are two sources of crude oil, Saudi Arabia and Borneo. The Saudi crude is relatively heavier (24 API), and the Far Eastern sector could obtain as much as 60,000 barrels per day at a cost of $18.50 per barrel during 1998. A second source of crude is from the Brunei fields in Borneo. This is a light crude oil (36 API). Under the terms of an agreement with the Netherlands Petroleum Company in Borneo, a fixed quantity of 40,000 b/d of Brunei crude, at a cost of $19.90 per barrel is to be supplied during 1998.

There are two subsidiaries that have refining operations. The first is in Australia, operating a refinery in Sydney with a capacity of 50,000 b/d throughput. The company markets its products throughout Australia, as well as having a surplus of refined products available for shipment to other subsidiaries.

The second subsidiary is in Japan, which operates a 30,000 b/d capacity refinery. Marketing operations are conducted in Japan, and excess production is available for shipment to other Far Eastern subsidiaries.

In addition, there are two marketing subsidiaries without refining capacity of their own. One of these is in New Zealand and the other is in the Philippines. Their needs can be supplied by shipments from Australia, Japan, or the Global Oil subsidiary in the United States. The latter is not a regular part of the Far Eastern Operations, but may be used as a source of refined products.

Finally, the company has a fleet of tankers that move the crude oil and refined products among the subsidiaries.

Refinery Operations

The operation of a refinery is a complex process. The characteristics of the crudes available, the desired output, the specific technology of the refinery, etc., make it difficult to use a simple model to describe the process. In fact, management at both Australia and Japan have complex linear programming models involving approximately 300 variables and 100 constraints for making detailed decisions on a daily or weekly basis.

For *annual* planning purposes the refinery model is greatly simplified. The two crudes (Saudi and Brunei) are input. Two general products are output – (a) gasoline products and (b) other products such as distillate, fuel oil, etc. In addition, although the refinery has processing flexibility that permits a wide range of yields, for planning purposes it was decided to include only the values at highest and lowest conversion rates (process intensity). Each refinery could use any combination of the two extreme intensities. These yields are shown in Table 6.1.

The incremental costs of operating the refinery depend somewhat upon the type of crude and process intensity. These costs are shown in Table 6.1. Also shown are the incremental transportation costs from either Borneo or Saudi Arabia.

Marketing Operations

Marketing is conducted in two home areas (Australia and Japan) as well as in the Philippines and New Zealand. Demand for gasoline and distillate in all areas has been estimated for 1998.

Area	**1998 Demand (thous. of b/d)**	
	Gasoline	Distillate
Australia	9.0	21.0
Japan	3.0	12.0
Philippines	5.0	8.0
New Zealand	5.4	8.7
TOTAL	**22.4**	**49.7**

Table 6.1: Refinery Operations

	Australia	Japan
Capacity (b/d of input)	50,000	30,000
Saudi Crude		
Transportation Cost ($/b)	0.65	0.70
High Process Intensity ($/b)	1.19	1.26
Yield of Gasoline	0.31	0.30
Yield of Distillate	0.61	0.62
Low Process Intensity ($/b)	0.89	0.88
Yield of Gasoline	0.19	0.18
Yield of Distillate	0.73	0.74
Brunei Crude		
Transportation Cost ($/b)	0.15	0.25
High Process Intensity ($/b)	0.93	0.91
Yield of Gasoline	0.36	0.35
Yield of Distillate	0.58	0.59
Low Process Intensity ($/b)	0.61	0.55
Yield of Gasoline	0.26	0.25
Yield of Distillate	0.69	0.70

Variable costs of supplying gasoline or distillate to New Zealand and the Philippines are:

Variable costs of shipment of gasoline/distillate in $/b

From:	To: New Zealand	Philippines
Australia	.20	.30
Japan	.25	.40

Tanker Operations

Tankers are used to bring crude from Saudi Arabia and Borneo to Australia and Japan and to transport refined products from Australia and Japan to the Philippines and New Zealand. The variable costs of these operations are included above.

However, there is a limited capacity of tankers available. The fleet has a capacity of 6.5 equivalent (standard sized) tankers.

The amount of capacity needed to deliver one barrel from one destination to another depends upon the distance traveled, port time, and other factors. The table below lists the fraction of one standard sized tanker needed to deliver 1,000 b/d over the indicated routes.

Tanker Usage Factors

(Fraction of Standard Sized Tanker Needed to Deliver 1,000 b/d)

Between	and Australia	Japan
Saudi Arabia	.12	.11
Borneo	.05	.05
Philippines	.02	.01
New Zealand	.01	.06

It is also possible to charter independent tankers. The rate for this is $5,400 per day for a standard sized tanker.

United States Supply

United States operations on the West Coast expect a surplus of 12,000 b/d of distillate during 1998. The cost of distillate at the loading port of Los Angeles is $20.70 per barrel. There is no excess gasoline capacity. The estimated variable shipping costs and tanker requirements of distillate shipments from the United States are:

	Variable costs of shipments	Tanker requirements
New Zealand	1.40	.18
Philippines	1.10	.15

Questions

Part I: Formulate a linear program which can be used to generate a comprehensive plan for the whole Far Eastern operations. Clearly define every variable used in your formulation.

Each group should be prepared to make a presentation of their linear programming model in class, on September 22. No written report is due on this date.

Part II: Written report due on October 6. Start your report with an executive summary (at most two pages) containing the most important results and recommendations. Attach supporting material in Appendices.

1. Solve your linear program using SOLVER.

   How many barrels of crude should Global Oil purchase from Saudi Arabia for its Far Eastern operations? How much crude should be refined in Australia? How much in Japan? Provide tables showing the quantities of gasoline and distillate shipped from each of the two refineries and from the U.S. to each of the four market areas.

2. Use sensitivity analysis to answer the following questions.

   - What is the marginal value of increasing supply from Brunei fields? Can this marginal value be used to estimate the total savings for Far Eastern operations if 41,000 b/d are supplied from Brunei fields instead of 40,000? Explain.
   - What is the marginal value of increasing the tanker fleet? Can this marginal value be used if we want to increase the size to 7 tankers (from the current size of 6.5)? Explain.
   - What is the additional cost to Far Eastern operations if demand for gasoline in the Philippines increases to 5,200 b/d? What is the minimum price of gasoline in the Philippines that would make it profitable for Global Oil to consider such an increase in distribution?
   - By how much should the production costs be reduced at the refinery in Japan when operating at high process intensity in order to make it cost effective to use Saudi crude?
   - Global Oil is planning a three day shutdown of its Australian refinery for maintenance purposes, in the coming year. It has storage facilities and at least two weeks of inventories, so a refinery shutdown for a few days will not disrupt distribution. What would be the cost of a planned shutdown of the refinery in Australia for three days per year? Same question for the refinery in Japan.
   - Currently, it is not economical to ship US distillate to the Philippines. What is the cost of US distillate at which Global Oil should consider starting such shipments?

3. Several opportunities present themselves to the Global Oil company (see the attached memos). Consider combinations of these options and prepare a recommendation. Document your report.

Memo to: Global Oil Headquarters

From: Australian Affiliate

Re: Supplements to Annual Plan

Since submitting data for annual planning purposes, two additional opportunities have arisen. We would like to include these in the plans.

A. Bid on Gasoline Contract with Australian Government

The government of Australia will submit to bid a contract for 1.5 thousand b/d of gasoline. We expect we could win this bid at a price of $26.40 per barrel. Estimated costs per barrel as follows:

Variable Costs (crude, refining, transportation)	$24.90
Allocated Overhead	.80
Total	**$25.70**

At these costs, the contract would have a profit contribution of $0.70 per barrel. We would like permission to bid on the contract. We hope that the linear programming wizards who are working for your logistics department will not contradict us!

B. Expansion of Australian Refinery

For the past two years, the Australian refinery has been operating at full capacity. We request authorization for capital expenditures to increase the refinery capacity to 55 thousand b/d. There are several reasons for the need for this expansion:

1. Australia can supply the current requirements in New Zealand and the Philippines more cheaply than can Japan.
2. The proposed bid on Australian government gasoline contract (above).
3. We understand the New Zealand affiliate is considering increasing its requirements by 4.5 thousand b/d. (See below.)

The cost of this expansion is 4.0 million dollars. To recover this investment[1], we need an annual savings of $702,000.

[1]This assumes a cost of capital rate of 20%. Depreciation tax effects are included. With these considerations, the $702,000 savings per year is equivalent to the $4 million investment

Memo to: Global Oil Headquarters

From: New Zealand Affiliate

Re: Supplement to Annual Plan

Negotiations have been begun with the NOZO Oil Company in New Zealand. This company is a distributor, with sales of 1.6 thousand b/d of gasoline and 3.2 thousand b/d of distillate. If negotiations are successfully completed, these requirements would be added to current requirements for New Zealand, making total requirements of:

Gasoline: 7.0 thousand b/d

Distillate: 11.9 thousand b/d

The anticipated revenue (after subtracting variable marketing costs) for this acquisition are $30.20 per barrel for gasoline and $24.60 per barrel for distillate. The purchase cost of NOZO oil is expected to be about $21.0 million. On an annual basis, this would require $3.5 million per year incremental profit to justify the purchase.

Memo to: Global Oil Headquarters

From: Tanker Affiliate

Re: Supplement to Annual Plan

We have been made aware that expansions in requirements are being considered in New Zealand and Australia. We are currently operating the tanker fleet at capacity. Additional requirements will increase the transport requirements both for crude and refined products. This will necessitate spot chartering unless additional tanker capacity is added.

We can lease on a long time basis additional tankers at a rate of $4.8 thousand dollars per day per 1 unit tanker equivalent. We propose a lease of 0.5 equivalent tanker units giving us a total capacity of 7.0 equivalent units. The cost of this would be $2.4 thousand per day or $876,000 per year. If you prefer other arrangements, we are willing to discuss them.

Memo to: Global Oil Headquarters

From: Borneo Office

Re: Supplement to Annual Plan

We have just been offered the opportunity to increase our contract with our supplier of Brunei crude. They are willing to supply us an additional 5 thousand b/d at a cost of $20.65 per barrel. Should we accept the offer? If not, is there a counter offer that we should propose?

Chapter 7

The Simplex Method

In this chapter, you will learn how to solve linear programs. This will give you insights into what SOLVER and other commercial linear programming software packages actually do. Such an understanding can be useful in several ways. For example, you will be able to identify when a problem has alternate optimal solutions (SOLVER never tells you this: it always give you only one optimal solution). You will also learn about degeneracy in linear programming and how this could lead to a very large number of iterations when trying to solve the problem.

7.1 Linear Programs in Standard Form

Before we start discussing the simplex method, we point out that every linear program can be converted into "standard" form

$$\begin{array}{rrcl} & \text{Max } c_1x_1 + c_2x_2 + \ldots + c_nx_n & & \\ \text{subject to} & a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n & = & b_1 \\ & \ldots \quad \ldots & & \ldots \\ & a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n & = & b_m \\ & x_1 \geq 0, \ldots x_n \geq 0 & & \end{array}$$

where the objective is maximized, the constraints are equalities and the variables are all nonnegative.

This is done as follows:

- If the problem is $min\ z$, convert it to $max\ -z$.
- If a constraint is $a_{i1}x_1+a_{i2}x_2+\ldots+a_{in}x_n \leq b_i$, convert it into an equality constraint by adding a nonnegative *slack* variable s_i. The resulting constraint is $a_{i1}x_1+a_{i2}x_2+\ldots+a_{in}x_n+s_i = b_i$, where $s_i \geq 0$.
- If a constraint is $a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n \geq b_i$, convert it into an equality constraint by subtracting a nonnegative *surplus* variable s_i. The resulting constraint is $a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n - s_i = b_i$, where $s_i \geq 0$.
- If some variable x_j is unrestricted in sign, replace it everywhere in the formulation by $x'_j - x''_j$, where $x'_j \geq 0$ and $x''_j \geq 0$.

Example 7.1.1 *Transform the following linear program into standard form.*

$$\begin{array}{lrrrcl} Min & -2x_1 & +3x_2 & & & \\ & x_1 & -3x_2 & +2x_3 & \leq & 3 \\ & -x_1 & +2x_2 & & \geq & 2 \\ & x_1 \ urs, & x_2 \geq 0, & x_3 \geq 0 & & \end{array}$$

Let us first turn the objective into a *max* and the constraints into equalities.

$$\begin{array}{lrrrrrcl} \text{Max} & 2x_1 & -3x_2 & & & & & \\ & x_1 & -3x_2 & +2x_3 & +s_1 & & = & 3 \\ & -x_1 & +2x_2 & & & -s_2 & = & 2 \\ & x_1 \text{ urs}, & x_2 \geq 0, & x_3 \geq 0 & s_1 \geq 0 & s_2 \geq 0 & & \end{array}$$

The last step is to convert the unrestricted variable x_1 into two nonnegative variables: $x_1 = x_1' - x_1''$.

$$\begin{array}{lrrrrrrcl} \text{Max} & 2x_1' & -2x_1'' & -3x_2 & & & & & \\ & x_1' & -x_1'' & -3x_2 & +2x_3 & +s_1 & & = & 3 \\ & -x_1' & +x_1'' & +2x_2 & & & -s_2 & = & 2 \\ & x_1' \geq 0, & x_1'' \geq 0, & x_2 \geq 0, & x_3 \geq 0 & s_1 \geq 0 & s_2 \geq 0 & & \end{array}$$

7.2 Solution of Linear Programs by the Simplex Method

For simplicity, in this course we solve "by hand" only the case where the constraints are of the form $\leq$ and the right-hand-sides are nonnegative. We will explain the steps of the simplex method while we progress through an example.

Example 7.2.1 *Solve the linear program*

$$\begin{array}{lrrcl} max & x_1 & +x_2 & & \\ & 2x_1 & +x_2 & \leq & 4 \\ & x_1 & +2x_2 & \leq & 3 \\ & x_1 \geq 0, & x_2 \geq 0 & & \end{array}$$

First, we convert the problem into standard form by adding slack variables $x_3 \geq 0$ and $x_4 \geq 0$.

$$\begin{array}{lrrrrcl} \text{max} & x_1 & +x_2 & & & & \\ & 2x_1 & +x_2 & +x_3 & & = & 4 \\ & x_1 & +2x_2 & & +x_4 & = & 3 \\ & x_1 \geq 0, & x_2 \geq 0 & x_3 \geq 0, & x_4 \geq 0 & & \end{array}$$

Let z denote the objective function value. Here, $z = x_1 + x_2$ or, equivalently,

$$z - x_1 - x_2 = 0.$$

Putting this equation together with the constraints, we get the following system of linear equations.

$$\begin{array}{crrrrcll} z & -x_1 & -x_2 & & & = & 0 & \text{Row 0} \\ & 2x_1 & +x_2 & +x_3 & & = & 4 & \text{Row 1} \\ & x_1 & +2x_2 & & +x_4 & = & 3 & \text{Row 2} \end{array} \tag{7.1}$$

Our goal is to maximize z, while satisfying these equations and, in addition, $x_1 \geq 0$, $x_2 \geq 0$, $x_3 \geq 0$, $x_4 \geq 0$.

Note that the equations are already in the form that we expect at the last step of the Gauss-Jordan procedure. Namely, the equations are solved in terms of the nonbasic variables x_1, x_2. The variables (other than the special variable z) which appear in only one equation are the *basic variables*. Here the basic variables are x_3 and x_4. A *basic solution* is obtained from the system of equations by setting the nonbasic variables to zero. Here this yields

$$x_1 = x_2 = 0 \quad x_3 = 4 \quad x_4 = 3 \quad z = 0.$$

Is this an optimal solution or can we increase z? (Our goal)

By looking at Row 0 above, we see that we can increase z by increasing x_1 or x_2. This is because these variables have a negative coefficient in Row 0. If all coefficients in Row 0 had been nonnegative, we could have concluded that the current basic solution is optimum, since there would be no way to increase z (remember that all variables x_i must remain ≥ 0). We have just discovered the first rule of the simplex method.

Rule 1 *If all variables have a nonnegative coefficient in Row 0, the current basic solution is optimal.*
Otherwise, pick a variable x_j with a negative coefficient in Row 0.

The variable chosen by Rule 1 is called the *entering* variable. Here let us choose, say, x_1 as our entering variable. It really does not matter which variable we choose as long as it has a negative coefficient in Row 0. The idea is to *pivot* in order to make the nonbasic variable x_1 become a basic variable. In the process, some basic variable will become nonbasic (the *leaving* variable). This *change of basis* is done using the Gauss-Jordan procedure. What is needed next is to choose the *pivot element*. It will be found using Rule 2 of the simplex method. In order to better understand the rationale behing this second rule, let us try both possible pivots and see why only one is acceptable.

First, try the pivot element in Row 1.

$$\begin{array}{llllllll} z & -x_1 & -x_2 & & & = & 0 & \text{Row 0} \\ & \mathbf{2x_1} & +x_2 & +x_3 & & = & 4 & \text{Row 1} \\ & x_1 & +2x_2 & & +x_4 & = & 3 & \text{Row 2} \end{array}$$

This yields

$$\begin{array}{llllllll} z & & -\frac{1}{2}x_2 & +\frac{1}{2}x_3 & & = & 2 & \text{Row 0} \\ & x_1 & +\frac{1}{2}x_2 & +\frac{1}{2}x_3 & & = & 2 & \text{Row 1} \\ & & \frac{3}{2}x_2 & -\frac{1}{2}x_3 & +x_4 & = & 1 & \text{Row 2} \end{array}$$

with basic solution $x_2 = x_3 = 0 \quad x_1 = 2 \quad x_4 = 1 \quad z = 2.$

Now, try the pivot element in Row 2.

$$\begin{array}{llllllll} z & -x_1 & -x_2 & & & = & 0 & \text{Row 0} \\ & 2x_1 & +x_2 & +x_3 & & = & 4 & \text{Row 1} \\ & \mathbf{x_1} & +2x_2 & & +x_4 & = & 3 & \text{Row 2} \end{array}$$

This yields

$$\begin{array}{lllllll} z & +x_2 & & +x_4 & = & 3 & \text{Row } 0 \\ & -3x_2 & +x_3 & -2x_4 & = & -2 & \text{Row } 1 \\ x_1 & +2x_2 & & +x_4 & = & 3 & \text{Row } 2 \end{array}$$

with basic solution $x_2 = x_4 = 0 \quad x_1 = 3 \quad x_3 = -2 \quad z = 3$.

Which pivot should we choose? The first one, of course, since the second yields an *infeasible* basic solution! Indeed, remember that we must keep all variables ≥ 0. With the second pivot, we get $x_3 = -2$ which is infeasible. How could we have known this ahead of time, before actually performing the pivots? The answer is, by comparing the ratios $\frac{\text{Right Hand Side}}{\text{Entering Variable Coefficient}}$ in Rows 1 and 2 of (7.1). Here we get $\frac{4}{2}$ in Row 1 and $\frac{3}{1}$ in Row 2. If you pivot in a row with *minimum* ratio, you will end up with a feasible basic solution (i.e. you will not introduce negative entries in the Right Hand Side), whereas if you pivot in a row with a ratio which is not minimum you will always end up with an infeasible basic solution. Just simple algebra! A negative pivot element would not be good either, for the same reason. We have just discovered the second rule of the simplex method.

Rule 2 *For each Row i, $i \geq 1$, where there is a strictly positive "entering variable coefficient", compute the ratio of the Right Hand Side to the "entering variable coefficient". Choose the pivot row as being the one with MINIMUM ratio.*

Once you have idendified the pivot element by Rule 2, you perform a Gauss-Jordan pivot. This gives you a new basic solution. Is it an optimal solution? This question is addressed by Rule 1, so we have closed the loop. The simplex method iterates between Rules 1, 2 and pivoting until Rule 1 guarantees that the current basic solution is optimal. That's all there is to the simplex method.

After our first pivot, we obtained the following system of equations.

$$\begin{array}{lllllll} z & -\frac{1}{2}x_2 & +\frac{1}{3}x_3 & & = & 2 & \text{Row } 0 \\ x_1 & +\frac{1}{2}x_2 & +\frac{1}{2}x_3 & & = & 2 & \text{Row } 1 \\ & \frac{3}{2}x_2 & -\frac{1}{2}x_3 & +x_4 & = & 1 & \text{Row } 2 \end{array}$$

with basic solution $x_2 = x_3 = 0 \quad x_1 = 2 \quad x_4 = 1 \quad z = 2$.

Is this solution optimal? No, Rule 1 tells us to choose x_2 as entering variable. Where should we pivot? Rule 2 tells us to pivot in Row 2, since the ratios are $\frac{2}{1/2} = 4$ for Row 1, and $\frac{1}{3/2} = \frac{2}{3}$ for Row 2, and the minimum occurs in Row 2. So we pivot on $\frac{3}{2}x_2$ in the above system of equations. This yields

$$\begin{array}{lllllll} z & & +\frac{1}{3}x_3 & +\frac{1}{3}x_4 & = & \frac{7}{3} & \text{Row } 0 \\ x_1 & & +\frac{2}{3}x_3 & -\frac{1}{3}x_4 & = & \frac{5}{3} & \text{Row } 1 \\ & x_2 & -\frac{1}{3}x_3 & +\frac{2}{3}x_4 & = & \frac{2}{3} & \text{Row } 2 \end{array}$$

with basic solution $x_3 = x_4 = 0 \quad x_1 = \frac{5}{3} \quad x_2 = \frac{2}{3} \quad z = \frac{7}{3}$.

Now Rule 1 tells us that this basic solution is optimal, since there are no more negative entries in Row 0.

All the above computations can be represented very compactly in *tableau form*.

z	x_1	x_2	x_3	x_4	RHS		Basic solution
1	-1	-1	0	0	0	basic	$x_3 = 4 \quad x_4 = 3$
0	2	1	1	0	4	nonbasic	$x_1 = x_2 = 0$
0	1	2	0	1	3		$z = 0$
1	0	$-\frac{1}{2}$	$\frac{1}{2}$	0	2	basic	$x_1 = 2 \quad x_4 = 1$
0	1	$\frac{1}{2}$	$\frac{1}{2}$	0	2	nonbasic	$x_2 = x_3 = 0$
0	0	$\frac{3}{2}$	$-\frac{1}{2}$	1	1		$z = 2$
1	0	0	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{7}{3}$	basic	$x_1 = \frac{5}{3} \quad x_2 = \frac{2}{3}$
0	1	0	$\frac{2}{3}$	$-\frac{1}{3}$	$\frac{5}{3}$	nonbasic	$x_3 = x_4 = 0$
0	0	1	$-\frac{1}{3}$	$\frac{2}{3}$	$\frac{2}{3}$		$z = \frac{7}{3}$

Since the above example has only two variables, it is interesting to interpret the steps of the simplex method graphically. See Figure 7.1. The simplex method starts in the corner point $(x_1 = 0, x_2 = 0)$ with $z = 0$. Then it discovers that z can increase by increasing, say, x_1. Since we keep $x_2 = 0$, this means we move along the x_1 axis. How far can we go? Only until we hit a constraint: if we went any further, the solution would become infeasible. That's exactly what Rule 2 of the simplex method does: the minimum ratio rule identifies the first constraint that will be encountered. And when the constraint is reached, its slack x_3 becomes zero. So, after the first pivot, we are at the point $(x_1 = 2, x_2 = 0)$. Rule 1 discovers that z can be increased by increasing x_2 while keeping $x_3 = 0$. This means that we move along the boundary of the feasible region $2x_1 + x_2 = 4$ until we reach another constraint! After pivoting, we reach the optimal point $(x_1 = \frac{5}{3}, x_2 = \frac{2}{3})$.

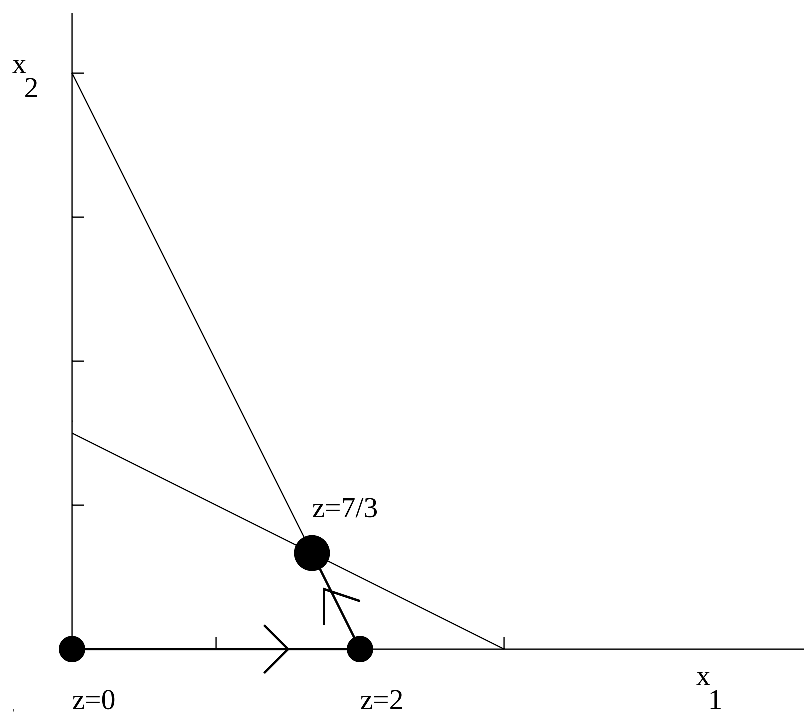


Figure 7.1: Graphical Interpretation

Exercise 62 Solve the following linear program by the simplex method.

$$
\begin{array}{lrrrrr}
\max & 4x_1 & +x_2 & -x_3 & & \\
 & x_1 & & +3x_3 & \leq & 6 \\
 & 3x_1 & +x_2 & +3x_3 & \leq & 9 \\
 & x_1 \geq 0, & x_2 \geq 0 & x_3 \geq 0 & &
\end{array}
$$

7.3 Alternate Optimal Solutions, Degeneracy, Unboudedness, Infeasibility

Alternate Optimal Solutions

Let us solve a small variation of the earlier example, with the same constraints but a slightly different objective:

$$
\begin{array}{lrrrr}
\max & x_1 & +\frac{1}{2}x_2 & & \\
 & 2x_1 & +x_2 & \leq & 4 \\
 & x_1 & +2x_2 & \leq & 3 \\
 & x_1 \geq 0, & x_2 \geq 0 & &
\end{array}
$$

As before, we add slacks x_3 and x_4, and we solve by the simplex method, using tableau representation.

z	x_1	x_2	x_3	x_4	RHS		Basic solution
1	-1	$-\frac{1}{2}$	0	0	0	basic	$x_3 = 4$ $x_4 = 3$
0	2	1	1	0	4	nonbasic	$x_1 = x_2 = 0$
0	1	2	0	1	3	$z = 0$	
1	0	0	$\frac{1}{2}$	0	2	basic	$x_1 = 2$ $x_4 = 1$
0	1	$\frac{1}{2}$	$\frac{1}{2}$	0	2	nonbasic	$x_2 = x_3 = 0$
0	0	$\frac{3}{2}$	$-\frac{1}{2}$	1	1	$z = 2$	

Now Rule 1 shows that this is an optimal solution. Interestingly, the coefficient of the nonbasic variable x_2 in Row 0 happens to be equal to 0. Going back to the rationale that allowed us to derive Rule 1, we observe that, if we increase x_2 (from its current value of 0), this will not effect the value of z. Increasing x_2 produces changes in the other variables, of course, through the equations in Rows 1 and 2. In fact, we can use Rule 2 and pivot to get a different basic solution with the same objective value $z = 2$.

z	x_1	x_2	x_3	x_4	RHS		Basic solution
1	0	0	$\frac{1}{2}$	0	2	basic	$x_1 = \frac{5}{3}$ $x_2 = \frac{2}{3}$
0	1	0	$\frac{2}{3}$	$-\frac{1}{3}$	$\frac{5}{3}$	nonbasic	$x_3 = x_4 = 0$
0	0	1	$-\frac{1}{3}$	$\frac{2}{3}$	$\frac{2}{3}$	$z = 2$	

Note that the coefficient of the nonbasic variable x_4 in Row 0 is equal to 0. Using x_4 as entering variable and pivoting, we would recover the previous solution!

Degeneracy

Example 7.3.1

$$\begin{array}{lrrcl} max & 2x_1 & +x_2 & & \\ & 3x_1 & +x_2 & \leq & 6 \\ & x_1 & -x_2 & \leq & 2 \\ & & x_2 & \leq & 3 \\ & x_1 \geq 0, & x_2 \geq 0 & & \end{array}$$

Let us solve this problem using the –by now familiar– simplex method. In the initial tableau, we can choose x_1 as the entering variable (Rule 1) and Row 2 as the pivot row (the minimum ratio in Rule 2 is a tie, and ties are broken arbitrarily). We pivot and this yields the second tableau below.

z	x_1	x_2	x_3	x_4	x_5	RHS	Basic solution
1	−2	−1	0	0	0	0	basic $x_3 = 6$ $x_4 = 2$ $x_5 = 3$
0	3	1	1	0	0	6	nonbasic $x_1 = x_2 = 0$
0	1	−1	0	1	0	2	$z = 0$
0	0	1	0	0	1	3	
1	0	−3	0	2	0	4	basic $x_1 = 2$ $x_3 = 0$ $x_5 = 3$
0	0	4	1	−3	0	0	nonbasic $x_2 = x_4 = 0$
0	1	−1	0	1	0	2	$z = 4$
0	0	1	0	0	1	3	

Note that this basic solution has a basic variable (namely x_3) which is equal to zero. When this occurs, we say that the basic solution is *degenerate*. Should this be of concern? Let us continue the steps of the simplex method. Rule 1 indicates that x_2 is the entering variable. Now let us apply Rule 2. The ratios to consider are $\frac{0}{4}$ in Row 1 and $\frac{3}{1}$ in Row 3. The minimum ratio occurs in Row 1, so let us perform the corresponding pivot.

z	x_1	x_2	x_3	x_4	x_5	RHS	Basic solution
1	0	0	$\frac{3}{4}$	$-\frac{1}{4}$	0	4	basic $x_1 = 2$ $x_2 = 0$ $x_5 = 3$
0	0	1	$\frac{1}{4}$	$-\frac{3}{4}$	0	0	nonbasic $x_3 = x_4 = 0$
0	1	0	$\frac{1}{4}$	$\frac{1}{4}$	0	2	$z = 4$
0	0	0	$-\frac{1}{4}$	$\frac{3}{4}$	1	3	

We get exactly the same solution! The only difference is that we have interchanged the names of a nonbasic variable with that of a degenerate basic variable (x_2 and x_3). Rule 1 tells us the solution is not optimal, so let us continue the steps of the simplex method. Variable x_4 is the entering variable and the last row wins the minimum ratio test. After pivoting, we get the tableau:

z	x_1	x_2	x_3	x_4	x_5	RHS	Basic solution
1	0	0	$\frac{2}{3}$	0	$\frac{1}{3}$	5	basic $x_1 = 1$ $x_2 = 3$ $x_4 = 4$
0	0	1	0	0	1	3	nonbasic $x_3 = x_5 = 0$
0	1	0	$\frac{1}{3}$	0	$-\frac{1}{3}$	1	$z = 5$
0	0	0	$-\frac{1}{3}$	1	$\frac{4}{3}$	4	

By Rule 1, this is the optimal solution. So, after all, degeneracy did not prevent the simplex method to find the optimal solution in this example. It just slowed things down a little. Unfortunately, on other examples, degeneracy may lead to *cycling*, i.e. a sequence of pivots that goes through the same tableaus and repeats itself indefinitely. In theory, cycling can be avoided by choosing the entering variable with smallest index in Rule 1, among all those with a negative coefficient in Row 0, and by breaking ties in the minimum ratio test by choosing the leaving variable with smallest index (this is known as Bland's rule). This rule, although it guaranties that cycling will never occur, turns out to be somewhat inefficient. Actually, in commercial codes, no effort is made to avoid cycling. This may come as a surprise, since degeneracy is a frequent occurence. But there are two reasons for this:

- Although degeneracy is frequent, cycling is extremely rare.
- The precision of computer arithmetic takes care of cycling by itself: round off errors accumulate and eventually gets the method out of cycling.

Our example of degeneracy is a 2-variable problem, so you might want to draw the constraint set in the plane and interpret degeneracy graphically.

Unbounded Optimum

Example 7.3.2

$$\begin{array}{lrrcl} max & 2x_1 & +x_2 & & \\ & -x_1 & +x_2 & \leq & 1 \\ & x_1 & -2x_2 & \leq & 2 \\ & x_1 \geq 0, & x_2 \geq 0 & & \end{array}$$

Solving by the simplex method, we get:

z	x_1	x_2	x_3	x_4	RHS		Basic solution
1	−2	−1	0	0	0	basic	$x_3 = 1$ $x_4 = 2$
0	−1	1	1	0	1	nonbasic	$x_1 = x_2 = 0$
0	1	−2	0	1	2	$z = 0$	
1	0	−5	0	2	4	basic	$x_1 = 2$ $x_3 = 3$
0	0	−1	1	1	3	nonbasic	$x_2 = x_4 = 0$
0	1	−2	0	1	2	$z = 4$	

At this stage, Rule 1 chooses x_2 as the entering variable, but there is no ratio to compute, since there is no positive entry in the column of x_2. As we start increasing x_2, the value of z increases (from Row 0) and the values of the basic variables increase as well (from Rows 1 and 2). There is nothing to stop them going off to infinity. So the problem is unbounded.

Interpret the steps of the simplex method graphically for this example.

Infeasible Linear Programs

It is easy to construct constraints that have no solution. The simplex method is able to identify such cases. We do not discuss it here. The interested reader is referred to Winston, for example.

Properties of Linear Programs

There are three possible outcomes for a linear program: it is infeasible, it has an unbounded optimum or it has an optimal solution.

If there is an optimal solution, there is a *basic* optimal solution. Remember that the number of basic variables in a basic solution is equal to the number of constraints of the problem, say m. So, even if the total number of variables, say n, is greater than m, at most m of these variables can have a positive value in an optimal basic solution.

Exercise 63 The following tableaus were obtained in the course of solving linear programs with 2 nonnegative variables x_1 and x_2 and 2 inequality constraints (the objective function z is maximized). Slack variables s_1 and s_2 were added. In each case, indicate whether the linear program

(i) is unbounded

(ii) has a unique optimum solution

(iii) has an alternate optimum solution

(iv) is degenerate (in this case, indicate whether any of the above holds).

(a)

z	x_1	x_2	s_1	s_2	RHS
1	0	3	2	0	20
0	1	−2	−1	0	4
0	0	−1	0	1	2

(b)

z	x_1	x_2	s_1	s_2	RHS
1	0	−1	0	2	20
0	0	0	1	−2	5
0	1	−2	0	3	6

(c)

z	x_1	x_2	s_1	s_2	RHS
1	2	0	0	1	8
0	3	1	0	−2	4
0	−2	0	1	1	0

(d)

z	x_1	x_2	s_1	s_2	RHS
1	0	0	2	0	5
0	0	−1	1	1	4
0	1	1	−1	0	4

Exercise 64 Suppose the following tableau was obtained in the course of solving a linear program with nonnegative variables x_1, x_2, x_3 and two inequalities. The objective function is maximized and slack variables s_1 and s_2 were added.

z	x_1	x_2	x_3	s_1	s_2	RHS
1	0	a	b	0	4	82
0	0	−2	2	1	3	c
0	1	−1	3	0	−5	3

Give conditions on a, b and c that are required for the following statements to be true:

(i) The current basic solution is a feasible basic solution.

Assume that the condition found in (i) holds in the rest of the exercise.

(ii) The current basic solution is optimal.

(iii) The linear program is unbounded (for this question, assume that $b > 0$).

(iv) The currrent basic solution is optimal and there are alternate optimal solutions (for this question, assume $a > 0$).

Exercise 65 A plant can manufacture five products P_1, P_2, P_3, P_4 and P_5. The plant consists of two work areas: the job shop area A_1 and the assembly area A_2. The time required to process one unit of product P_j in work area A_i is p_{ij} (in hours), for $i = 1, 2$ and $j = 1, \ldots, 5$. The weekly capacity of work area A_i is C_i (in hours). The company can sell all it produces of product P_j at a profit of s_j, for $i = 1, \ldots, 5$.

The plant manager thought of writing a linear program to maximize profits, but never actually did for the following reason: From past experience, he observed that the plant operates best when at most two products are manufactured at a time. He believes that if he uses linear programming, the optimal solution will consist of producing all five products. Do you agree with him? Explain, based on your knowledge of linear programming.

Exercise 66 Consider the linear program

Maximize $5x_1 + 3x_2 + x_3$

Subject to

$x_1 + x_2 + x_3 \leq 6$

$5x_1 + 3x_2 + 6x_3 \leq 15$

$x_1, x_2, x_3 \geq 0$

and an associated tableau

z	x_1	x_2	x_3	s_1	s_2	RHS
1	0	0	5	0	1	15
0	0	.4	-0.2	1	-.2	3
0	1	.6	1.2	0	.2	3

(a) What basic solution does this tableau represent? Is this solution optimal? Why or why not?

(b) Does this tableau represent a unique optimum. If not, find an alternative optimal solution.

Answer:

(a) *The solution is $x_1 = 3$, $x_2, x_3 = 0$, objective 15. It is optimal since cost row is all at least 0.*

(b) *It is not unique (since x_2 has reduced cost 0 but is not basic). Alternative found by pivoting in x_2 (question could have asked for details on such a pivot) for solution $x_2 = 5$, $x_1, x_3 = 0$, objective 15.*

Chapter 8

Sensitivity Analysis for Linear Programming

Finding the optimal solution to a linear programming model is important, but it is not the only information available. There is a tremendous amount of *sensitivity information*, or information about what happens when data values are changed.

Recall that in order to formulate a problem as a linear program, we had to invoke a *certainty assumption*: we had to know what value the data took on, and we made decisions based on that data. Often this assumption is somewhat dubious: the data might be unknown, or guessed at, or otherwise inaccurate. How can we determine the effect on the optimal decisions if the values change? Clearly some numbers in the data are more important than others. Can we find the "important" numbers? Can we determine the effect of misestimation?

Linear programming offers extensive capability for addressing these questions. We begin by showing how data changes show up in the optimal table. We then give two examples of how to interpret Solver's extensive output.

8.1 Tableau Sensitivity Analysis

Suppose we solve a linear program "by hand" ending up with an optimal table (or tableau to use the technical term). We know what an optimal tableau looks like: it has all non–negative values in Row 0 (which we will often refer to as the *cost row*), all non–negative right–hand–side values, and a basis (identity matrix) embedded. To determine the effect of a change in the data, we will try to determine how that change effected the final tableau, and try to reform the final tableau accordingly.

8.1.1 Cost Changes

The first change we will consider is changing a cost value by Δ in the original problem. We are given the original problem and an optimal tableau. If we had done exactly the same calculations beginning with the modified problem, we would have had the same final tableau except that the corresponding cost entry would be Δ lower (this is because we never do anything except add or subtract scalar multiples of Rows 1 through m to other rows; we never add or subtract Row 0 to other rows). For example, take the problem

```
Max 3x+2y
Subject to
   x+y <= 4
   2x+y <= 6
   x,y >= 0
```

The optimal tableau to this problem (after adding s_1 and s_2 as slacks to place in standard form) is:

z	x	y	s_1	s_2	RHS
1	0	0	1	1	10
0	0	1	2	-1	2
0	1	0	-1	1	2

Suppose the cost for x is changed to $3+\Delta$ in the original formulation, from its previous value 3. After doing the same operations as before, that is the same pivots, we would end up with the tableau:

z	x	y	s_1	s_2	RHS
1	$-\Delta$	0	1	1	10
0	0	1	2	-1	2
0	1	0	-1	1	2

Now this is not the optimal tableau: it does not have a correct basis (look at the column of x). But we can make it correct in form while keeping the same basic variables by adding Δ times the last row to the cost row. This gives the tableau:

z	x	y	s_1	s_2	RHS
1	0	0	$1-\Delta$	$1+\Delta$	$10+2\Delta$
0	0	1	2	-1	2
0	1	0	-1	1	2

Note that this tableau has the same basic variables and the same variable values (except for z) that our previous solution had. Does this represent an optimal solution? It does only if the cost row is all non–negative. This is true only if

$$1-\Delta \geq 0$$

$$1+\Delta \geq 0$$

which holds for $-1 \leq \Delta \leq 1$. For any Δ in that range, our previous basis (and variable values) is optimal. The objective changes to $10+2\Delta$.

In the previous example, we changed the cost of a basic variable. Let's go through another example. This example will show what happens when the cost of a nonbasic variable changes.

```
Max 3x+2y + 2.5w
Subject to
   x+y +2w <= 4
   2x+y +2w <= 6
   x,y,w >= 0
```

Here, the optimal tableau is :

z	x	y	w	s_1	s_2	RHS
1	0	0	1.5	1	1	10
0	0	1	2	2	-1	2
0	1	0	0	-1	1	2

Now suppose we change the cost on w from 2.5 to $2.5 + \Delta$ in the formulation. Doing the same calculations as before will result in the tableau:

z	x	y	w	s_1	s_2	RHS
1	0	0	$1.5 - \Delta$	1	1	10
0	0	1	2	2	-1	2
0	1	0	0	-1	1	2

In this case, we already have a valid tableau. This will represent an optimal solution if $1.5 - \Delta \geq 0$, so $\Delta \leq 1.5$. As long as the objective coefficient of w is no more than 2.5+1.5=4 in the original formulation, our solution of $x = 2, y = 2$ will remain optimal.

The value in the cost row in the simplex tableau is called the *reduced cost*. It is zero for a basic variable and, in an optimal tableau, it is non–negative for all other variables (for a maximization problem).

Summary: Changing objective function values in the original formulation will result in a changed cost row in the final tableau. It might be necessary to add a multiple of a row to the cost row to keep the form of the basis. The resulting analysis depends only on keeping the cost row non–negative.

8.1.2 Right Hand Side Changes

For these types of changes, we concentrate on maximization problems with all $\leq$ constraints. Other cases are handled similarly.

Take the following problem:

$$\begin{aligned} \text{Max} \quad & 4x + 5y \\ \text{Subject to} & \\ 2x + 3y &<= 12 \\ x + y &<= 5 \\ x, y &>= 0 \end{aligned} \tag{8.1}$$

The optimal tableau, after adding slacks s_1 and s_2 is

z	x	y	s_1	s_2	RHS
1	0	0	1	2	22
0	0	1	1	-2	2
0	1	0	-1	3	3

Now suppose instead of 12 units in the first constraint, we only had 11. This is *equivalent* to forcing s_1 to take on value 1. Writing the constraints in the optimal tableau long–hand, we get

$$z + s_1 + 2s_2 = 22$$

$$y + s_1 - 2s_2 = 2$$

$$x - s_1 + 3s_2 = 3$$

If we force s_1 to 1 and keep s_2 at zero (as a nonbasic variable should be), the new solution would be $z = 21$, $y = 1$, $x = 4$. Since all variables are nonnegative, this is the optimal solution.

In general, changing the amount of the right–hand–side from 12 to $12 + \Delta$ in the first constraint changes the tableau to:

z	x	y	s_1	s_2	RHS
1	0	0	1	2	$22 + \Delta$
0	0	1	1	-2	$2 + \Delta$
0	1	0	-1	3	$3 - \Delta$

This represents an optimal tableau as long as the righthand side is all non–negative. In other words, we need Δ between -2 and 3 in order for the basis not to change. For any Δ in that range, the optimal objective will be $22 + \Delta$. For example, with Δ equals 2, the new objective is 24 with $y = 4$ and $x = 1$.

Similarly, if we change the right–hand–side of the second constraint from 5 to $5 + \Delta$ in the original formulation, we get an objective of $22 + 2\Delta$ in the final tableau, as long as $-1 \leq \Delta \leq 1$.

Perhaps the most important concept in sensitivity analysis is the *shadow price* λ_i^* of a constraint: *If the RHS of Constraint i changes by Δ in the original formulation, the optimal objective value changes by $\lambda_i^*\Delta$.* The shadow price λ_i^* can be found in the optimal tableau. It is the reduced cost of the slack variable s_i. So it is found in the cost row (Row 0) in the column corresponding the slack for Constraint i. In this case, $\lambda_1^* = 1$ (found in Row 0 in the column of s_1) and $\lambda_2^* = 2$ (found in Row 0 in the column of s_2). The value λ_i^* is really the marginal value of the resource associated with Constraint i. For example, the optimal objective value (currently 22) would increase by 2 if we could increase the RHS of the second constraint by $\Delta = 1$. In other words, the marginal value of that resource is 2, i.e. we are willing to pay up to 2 to increase the right hand side of the second constraint by 1 unit. You may have noticed the similarity of interpretation between shadow prices in linear programming and Lagrange multipliers in constrained optimization. Is this just a coincidence? Of course not. This parallel should not be too surprising since, after all, linear programming is a special case of constrained optimization. To derive this equivalence (between shadow prices and optimal Lagrange multipiers), one could write the KKT conditions for the linear program...but we will skip this in this course!

In summary, changing the right–hand–side of a constraint is identical to setting the corresponding slack variable to some value. This gives us the shadow price (which equals the reduced cost for the corresponding slack) and the ranges.

8.1.3 New Variable

The shadow prices can be used to determine the effect of a new variable (like a new product in a production linear program). Suppose that, in formulation (8.1), a new variable w has coefficient 4 in the first constraint and 3 in the second. What objective coefficient must it have to be considered for adding to the basis?

If we look at making w positive, then this is equivalent to decreasing the right hand side of the first constraint by $4w$ and the right hand side of the second constraint by $3w$ in the original formulation. We obtain the same effect by making $s_1 = 4w$ and $s_2 = 3w$. The overall effect of this is to decrease the objective by $\lambda_1^*(4w) + \lambda_2^*(3w) = 1(4w) + 2(3w) = 10w$. The objective value must be sufficient to offset this, so the objective coefficient must be more than 10 (exactly 10 would lead to an alternative optimal solution with no change in objective).

Example 8.1.1 A factory can produce four products denoted by P_1, P_2, P_3 and P_4. Each product must be processed in each of two workshops. The processing times (in hours per unit produced) are given in the following table.

	P_1	P_2	P_3	P_4
Workshop 1	3	4	8	6
Workshop 2	6	2	5	8

400 hours of labour are available in each workshop. The profit margins are 4, 6, 10, and 9 dollars per unit of P_1, P_2, P_3 and P_4 produced, respectively. Everything that is produced can be sold. Thus, maximizing profits, the following linear program can be used.

```
MAX          4 X1 + 6 X2 + 10 X3 + 9 X4
SUBJECT TO
             3 X1 + 4 X2 +  8 X3 + 6 X4 <=  400      Row 1
             6 X1 + 2 X2 +  5 X3 + 8 X4 <=  400      Row 2
               X1, X2, X3, X4 >= 0
```

Introducing slack variables s_1 and s_2 in Rows 1 and 2, respectively, and applying the simplex method, we get the final tableau:

z	x_1	x_2	x_3	x_4	s_1	s_2	RHS
1	0.5	0	2	0	1.5	0	600
0	0.75	1	2	1.5	0.25	0	100
0	4.5	0	1	5	-0.5	1	200

(a) How many units of P_1, P_2, P_3 and P_4 should be produced in order to maximize profits?

(b) Assume that 20 units of P_3 have been produced by mistake. What is the resulting decrease in profit?

(c) In what range can the profit margin per unit of P_1 vary without changing the optimal basis?

(d) In what range can the profit margin per unit of P_2 vary without changing the optimal basis?

(e) What is the marginal value of increasing the production capacity of Workshop 1?

(f) In what range can the capacity of Workshop 1 vary without changing the optimal basis?

(g) Management is considering the production of a new product P_5 that would require 2 hours in Workshop 1 and ten hours in Workshop 2. What is the minimum profit margin needed on this new product to make it worth producing?

Answers:

(a) *From the final tableau, we read that* $x_2 = 100$ *is basic and* $x_1 = x_3 = x_4 = 0$ *are nonbasic. So 100 units of* P_2 *should be produced and none of* P_1, P_3 *and* P_4. *The resuting profit is \$ 600 and that is the maximum possible, given the constraints.*

(b) *The reduced cost for x_3 is 2 (found in Row 0 of the final tableau). Thus, the effect on profit of producing x_3 units of P_3 is $-2x_3$. If 20 units of P_3 have been produced by mistake, then the profit will be $2 \times 20 = \$40$ lower than the maximum stated in* (a).

(c) *Let $4 + \Delta$ be the profit margin on P_1. The reduced cost remains nonnegative in the final tableau if $0.5 - \Delta \geq 0$. That is $\Delta \leq 0.5$. Therefore, as long as the profit margin on P_1 is less than 4.5, the optimal basis remains unchanged.*

(d) *Let $6 + \Delta$ be the profit margin on P_2. Since x_2 is basic, we need to restore a correct basis. This is done by adding Δ times Row 1 to Row 0. This effects the reduced costs of the nonbasic variables, namely x_1, x_3, x_4 and s_1. All these reduced costs must be nonnegative. This implies:*

$0.5 + 0.75\Delta \geq 0$

$2 + 2\Delta \geq 0$

$0 + 1.5\Delta \geq 0$

$1.5 + 0.25\Delta \geq 0.$

Combining all these inequalities, we get $\Delta \geq 0$. So, as long as the profit margin on P_2 is 6 or greater, the optimal basis remains unchanged.

(e) *The marginal value of increasing capacity in Workshop 1 is $\lambda_1^* = 1.5$.*

(f) *Let $400 + \Delta$ be the capacity of Workshop 1. The resulting RHS in the final tableau will be:*

$100 + 0.25\Delta$ *in Row 1, and*

$200 - 0.5\Delta$ *in Row 2.*

The optimal basis remains unchanged as long as these two quantities are nonnegative. Namely, $-400 \leq \Delta \leq 400$. So, the optimal basis remains unchanged as long as the capacity of Workshop 1 is in the range 0 to 800.

(g) *The effect on the optimum profit of producing x_5 units of P_5 would be $\lambda_1^*(2x_5) + \lambda_2^*(10x_5) = 1.5(2x_5) + 0(10x_5) = 3x_5$. If the profit margin on P_5 is sufficient to offset this, then P_5 should be produced. That is, we should produce P_5 if its profit margin is at least 3.*

Exercise 67 A paper mill converts pulpwood to low, medium and high grade newsprint. The pulpwood requirements for each newsprint, availability of each pulpwood, and selling price (per ton) are shown below:

	Low grade	Medium grade	High grade	Available (tons)
Virginia pine	2	2	1	180
White pine	1	2	3	120
Loblolly pine	1	1	2	160
Price	\$900	\$1000	\$1200	

The associated linear program is

```
MAX      900 X1  + 1000 X2  + 1200 X3
SUBJECT TO
           2 X1  +    2 X2  +      X3  +  S1             =  180
             X1  +    2 X2  +    3 X3        +  S2       =  120
             X1  +      X2  +    2 X3              + S3 =  160
             X1, X2, X3, S1, S2, S3 >= 0
```

with the optimal tableau

z	x_1	x_2	x_3	s_1	s_2	s_3	RHS
1	0	200	0	300	300	0	90,000
0	1	0.8	0	0.6	−0.2	0	84
0	0	0.4	1	−0.2	0.4	0	12
0	0	−0.6	0	−0.2	−0.6	1	52

(a) In what range can the price of low grade paper vary without changing the optimal basis?

(b) What is the new optimal solution if the price of low grade paper changes to $800?

(c) At what price should medium grade paper be sold to make it profitable to produce?

(d) In what range can the availability of Virginia pine vary without changing the optimal basis?

(e) If 10 additional tons of Virginia pine are obtained, by how much will the optimal profit increase?

(f) If the pulpwood resources are increased as in (e), what is the new optimal solution (i.e. the optimal production levels of low, medium and high grade newsprint)?

(g) What would the plant manager be willing to pay for an additional ton of Loblolly pine?

Exercise 68 Snacks 'R Us is deciding what to produce in the upcoming month. Due to past purchases, they have 500 lb of walnuts, 1000 lb of peanuts, and 500 lb of chocolate. They currently sell three mixes: Trail Mix, which consists of 1 lb of each material; Nutty Crunch, which has two pounds of peanuts and 1 lb of walnut; and Choc-o-plenty, which has 2 pounds of chocolate and 1 lb of peanuts. They can sell an unlimited amount of these mixes, except they can sell no more than 100 units of Choc-o-plenty. The income on the three mixes is $2, $3, and $4, respectively. The problem of maximizing total income subject to these constraints is

Maximize $2x_1 + 3x_2 + 4x_3$

Subject to

$x_1 + x_2 \leq 500$

$x_1 + 2x_2 + x_3 \leq 1000$

$x_1 + 2x_3 \leq 500$

$x_3 \leq 100$

$x_i \geq 0$ for all i

The optimal tableau for this problem is (after adding slacks for the four constraints):

z	x_1	x_2	x_3	s_1	s_2	s_3	s_4	RHS
1	0	0	0	1	1	0	3	1800
0	1	0	0	2	−1	0	1	100
0	0	1	0	−1	1	0	−1	400
0	0	0	0	−2	1	1	−3	200
0	0	0	1	0	0	0	1	100

Using the above tableau, determine the following:

(a) What is the solution represented by this tableau (both production quantities and total income)? How do we know it is the optimal tableau?

Production quantities: ______________________________

Total income: ______________________

Optimal? ___________ Reason: ____

(b) Suppose the income for Trail Mix (x_1) is only an estimate. For what range of values would the basis given by the above tableau still be optimal? What would be the solution (both production quantities and total income), if the income was only \$2.75?

Range: Lower: ___________ Upper: ___________

Solution for \$2.75: Total Income ___________ Production: ________________________

(c) How much should Snacks 'R Us willing to spend to procure an extra pound of peanuts? How about a pound of walnuts? How about an extra pound of chocolate?

Peanuts: ___________ Walnuts: ___________ Chocolate: ___________

(d) For what range of peanut–stock values is the basis given by the above tableau still optimal? What would be the solution (both production quantities and total objective) if there were only 900 lbs of peanuts available?

Range: Lower: ___________ Upper: ___________

Solution for 900 lbs: Total Income ___________ Production: ______________________

(e) A new product, Extra-Walnutty Trail Booster, consists of 1 lb of each of peanuts and chocolate, and 2 lb of walnuts. What income do we need to make on this product in order to consider producing it?

Income: ___________ Reason: ____

8.2 Solver Output

A large amount (but not all) of the previous analysis is available from the output of Solver. To access the information, simply ask for the sensitivity report after optimizing. Rather than simply give rules for reading the report, here are two reports, each with a set of questions that can be answered from the output.

8.2.1 Tucker Automobiles

Tucker Inc. needs to produce 1000 Tucker automobiles. The company has four production plants. Due to differing workforces, technological advances, and so on, the plants differ in the cost of producing each car. They also use a different amount of labor and raw material at each. This is summarized in the following table:

Plant	Cost ('000)	Labor	Material
1	15	2	3
2	10	3	4
3	9	4	5
4	7	5	6

The labor contract signed requires at least 400 cars to be produced at plant 3; there are 3300 hours of labor and 4000 units of material that can be allocated to the four plants.

This leads to the following formulation:

```
MIN     15 X1 + 10 X2 + 9 X3 + 7 X4
SUBJECT TO
          2 X1 + 3 X2 + 4 X3 + 5 X4 <=   3300
```

```
3 X1 + 4 X2 + 5 X3 + 6 X4 <=   4000
X3 >=   400
X1 + X2 + X3 + X4 =    1000
X1, X2, X3, X4 >= 0
```

Attached is the Solver output for this problem.

Adjustable Cells

Cell	Name	Final Value	Reduced Cost	Objective Coefficient	Allowable Increase	Allowable Decrease
$B$11	X1	400	0	15	1E+30	3.5
$B$12	X2	200	0	10	2	1E+30
$B$13	X3	400	0	9	1E+30	4
$B$14	X4	0	7	7	1E+30	7

Constraints

Cell	Name	Final Value	Shadow Price	Constraint R.H. Side	Allowable Increase	Allowable Decrease
$B$18	Labor	3000	0	3300	1E+30	300
$B$19	Material	4000	-5	4000	300	200
$B$20	Plant3Min	400	4	400	100	400
$B$21	Total	1000	30	1000	66.6667	100

Here are some questions to be answered:

1. What are the current production quantities? What is the current cost of production?
2. How much will it cost to produce one more vehicle? How much will we save by producing one less?
3. How would our solution change if it cost only $8,000 to produce at plant 2? For what ranges of costs is our solution (except for the objective value) valid for plant 2?
4. How much are we willing to pay for a labor hour?
5. How much is our union contract costing us? What would be the value of reducing the 400 car limit down to 200 cars? To 0 cars? What would be the cost of increasing it by 100 cars? by 200 cars?
6. How much is our raw material worth (to get one more unit)? How many units are we willing to buy at that price? What will happen if we want more?

7. A new plant is being designed that will use only one unit of workers and 4 units of raw material. What is the maximum cost it can have in order for us to consider using it?

8. By how much can the costs at plant 1 increase before we would not produce there?

8.2.2 Carla's Maps

Carla Lee, a current MBA student, decides to spend her summer designing and marketing bicycling maps of Western Pennsylvania. She has designed 4 maps, corresponding to four quadrants around Pittsburgh. The maps differ in size, colors used, and complexity of the topographical relief (the maps are actually 3-dimensional, showing hills and valleys). She has retained a printer to produce the maps. Each map must be printed, cut, and folded. The time (in minutes) to do this for the four types of maps is:

	Print	Cut	Fold
A	1	2	3
B	2	4	2
C	3	1	5
D	3	3	3
Avail	15000	20000	20000

The printer has a limited amount of time in his schedule, as noted in the table.

The profit per map, based on the projected selling price minus printers cost and other variable cost, comes out to approximately $1 for A and B and $2 for C and D. In order to have a sufficiently nice display, at least 1000 of each type must be produced.

This gives the formulation:

```
MAX      A + B + 2 C + 2 D
SUBJECT TO
          A + 2 B + 3 C + 3 D <=   15000
          2 A + 4 B + C + 3 D <=   20000
          3 A + 2 B + 5 C + 3 D <=   20000
          A >=   1000
          B >=   1000
          C >=   1000
          D >=   1000
```

Attached is the Solver output.

Adjustable Cells

Cell	Name	Final Value	Reduced Cost	Objective Coefficient	Allowable Increase	Allowable Decrease
$B$11	XA	1500	0	1	1	0.333333
$B$12	XB	1000	0	1	0.333333	1E+30
$B$13	XC	1000	0	2	0.333333	1E+30
$B$14	XD	2833.333	0	2	1	0.5

Constraints

Cell	Name	Final Value	Shadow Price	Constraint R.H. Side	Allowable Increase	Allowable Decrease
$B$17	Print	15000	0.5	15000	100	366.6667
$B$18	Cut	16500	0	20000	1E+30	3500
$B$19	Fold	20000	0.166667	20000	7000	1000
$B$20	MinA	1500	0	1000	500	1E+30
$B$21	MinB	1000	-0.333333	1000	1750	1000
$B$22	MinC	1000	-0.333333	1000	500	1000
$B$23	MinD	2833.333	0	1000	1833.333	1E+30

Here are some questions to answer:

1. What are the production quantities and projected profit?
2. How much is Carla willing to pay for extra printing time? cutting time? folding time? For each, how many extra hours are we willing to buy at that price?
3. Suppose we reduced the 1000 limit on one item to 900. Which map should be decreased, and how much more would Carla make?
4. A fifth map is being thought about. It would take 2 minutes to print, 2 minutes to cut, and 3 minutes to fold. What is the least amount of profit necessary in order to consider producing this map? What is the effect of requiring 1000 of these also?
5. The marketing analysis on D is still incomplete, though it is known that the profit of $2 per item is within $.25 of the correct value. It will cost $500 to complete the analysis. Should Carla continue with the analysis?

Exercise 69 Recall the Red Dwarf toaster example (Exercise 61).

From Answer report: Final Cost: 7833.33

Adjustable Cells

Cell	Name	Final Value	Reduced Cost	Objective Coefficient	Allowable Increase	Allowable Decrease
$B$1	Manual	633.3333333	0	7	0.5	1E+30
$B$2	SemiAut	333.3333333	0	8	1E+30	0.25
$B$3	Robotic	33.33333333	0	8.5	0.5	2.5

Constraints

Cell	Name	Final Value	Shadow Price	Constraint R.H. Side	Allowable Increase	Allowable Decrease

$B$5	Total	1000	10.83333333	1000	170	100
$B$6	Skilled	2233.333333	0	4500	1E+30	2266.666667
$B$7	Unskill	36000	-0.083333333	36000	1000	6800
$B$8	Assembly	2700	-0.166666667	2700	500	100

(a) What is the optimal allocation of production? What is the average cost/toaster of production.

(b) By how much can the cost of robots increase before we will change that production plan.

(c) How much is Red Dwarf willing to pay for more assembly room time? How many units is Red Dwarf willing to purchase at that price?

(d) How much will we save if we decide to produce only 950 toasters?

(e) A new production process is available that uses only 2 minutes of skilled labor, 10 minutes of unskilled labor, and an undetermined amount of assembly floor time. Its production cost is determined to be $10. What is the maximum assembly floor time that the process can take before it is deemed too expensive to use?

Answer:

(a) *633.3 should be produced manually, 333.3 should be produced semiautomatically, and 33.3 produced robotically, for an average cost of $7.383/toaster.*

(b) *It can increase by $0.50.*

(c) *Value is $0.16/minute, willing to purchase 500 at that price.*

(d) *Objective will go down by 50(10.833).*

(e) *Cost of $10 versus marginal cost of $10.833, leave 0.83. Unskilled labor costs $0.0833/unit. Therefore, if the new process takes any time at all, it will be deemed too expensive.*

Exercise 70 Kennytrail Amusement park is trying to divide its new 50 acre park into one of three categories: rides, food, and shops. Each acre used for rides generates $150/hour profit; each acre used for food generates $200/hour profit. Shops generate $300/hour profit. There are a number of restrictions on how the space can be divided.

1. Only 10 acres of land is suitable for shops.
2. Zoning regulations require at least 1000 trees in the park. A food acre has 30 trees; a ride acre has 20 trees; while a shop acre has no trees.
3. No more than 200 people can work in the park. It takes 3 people to work an acre of rides, 6 to work an acre of food, and 5 to work an acre of shops.

The resulting linear program and Solver output is attached:

```
MAX     150 RIDE + 200 FOOD + 300 SHOP
SUBJECT TO
       RIDE + FOOD + SHOP <=   50
       SHOP <=   10
       20 RIDE + 30 FOOD >=   1000
       3 RIDE + 6 FOOD + 5 SHOP <=   200

-------------------------------------------------------------------
```

Answer report:

Target Cell (Max)

Cell	Name	Original Value	Final Value
$B$5	Cost	0	9062.5

Adjustable Cells

Cell	Name	Original Value	Final Value
$B$1	Ride	0	31.25
$B$2	Food	0	12.5
$B$3	Shop	0	6.25

Constraints

Cell	Name	Cell Value	Formula	Status	Slack
$B$6	Land	50	$B$6<=$D$6	Binding	0
$B$7	ShopLim	6.25	$B$7<=$D$7	Not Binding	3.75
$B$8	Trees	1000	$B$8>=$D$8	Binding	0
$B$9	Workers	200	$B$9<=$D$9	Binding	0

Sensitivity report:

Adjustable Cells

Cell	Name	Final Value	Reduced Cost	Objective Coefficient	Allowable Increase	Allowable Decrease
$B$1	Ride	31.25	0	150	83.33333333	76.6666667
$B$2	Food	12.5	0	200	115	125
$B$3	Shop	6.25	0	300	1E+30	116.666667

Constraints

Cell	Name	Final Value	Shadow Price	Constraint R.H. Side	Allowable Increase	Allowable Decrease
$B$6	Land	50	143.75	50	10	16.66666667
$B$7	ShopLim	6.25	0	10	1E+30	3.75
$B$8	Trees	1000	-4.375	1000	166.66667	100
$B$9	Workers	200	31.25	200	30	50

For each of the following changes, either find the answer or state that the information is not available from the Solver output. In the latter case, state why not.

(a) What is the optimal allocation of the space? What is the profit/hour of the park?

Optimal Allocation: ______________________ Profit/hour: __________

(b) Suppose Food only made a profit of $180/hour. What would be the optimal allocation of the park, and what would be the profit/hour of the park?

Optimal Allocation: ______________________ Profit/hour: __________

(c) City Council wants to increase our tree requirement to 1020. How much will that cost us (in \$/hour). What if they increased the tree requirement to 1200?

Increase to 1020: ____________ Increase to 1200: ____________

(d) A construction firm is willing to convert 5 acres of land to make it suitable for shops. How much should Kennytrail be willing to pay for this conversion (in \$/hour).

Maximum payment: ____________

(e) Kennytrail is considering putting in a waterslide. Each acre of waterslide can have 2 trees and requires 4 workers. What profit/hour will the waterslide have to generate for them to consider adding it?

Minimum Profit: ____________ Reason: ____

(f) An adjacent parcel of land has become available. It is five acres in size. The owner wants to share in our profits. How much \$/hour is Kennytrail willing to pay?

Maximum payment: ____________

Exercise 71 Attached is the sensitivity report for the Diet Problem (see Section 5.4.1 in Chapter 5).

Adjustable Cells

Cell	Name	Final Value	Reduced Cost	Objective Coefficient	Allowable Increase	Allowable Decrease
$B$14	oatmeal	4	-3.1875	3	3.1875	1E+30
$B$15	chicken	0	12.46875	24	1E+30	12.46875
$B$16	eggs	0	4	13	1E+30	4
$B$17	milk	4.5	0	9	2.6923	1.38095
$B$18	pie	2	-3.625	20	3.625	1E+30
$B$19	pork&beans	0	4.375	19	1E+30	4.375

Constraints

Cell	Name	Final Value	Shadow Price	Constraint R.H. Side	Allowable Increase	Allowable Decrease
$B$23	Energy	2000	0.05625	2000	560	100
$B$24	Protein	60	0	55	5	1E+30
$B$25	Calcium	1334.5	0	800	534.5	1E+30

Answer each of the following questions independently of the others.

1. What does the optimal diet consist of?
2. If the cost of oatmeal doubled to 6 cents/serving, should it be removed from the diet?
3. If the cost of chicken went down to half its current price, should it be added to the diet?
4. At what price would eggs start entering the diet?
5. In what range can the price of milk vary (rounding to the nearest tenth of a cent) while the current diet still remaining optimal?

6. During midterms, you need a daily diet with energy content increased from 2000 kcal to 2200 kcal. What is the resulting additional cost?

7. Your doctor recommends that you increase the calcium requirement in your diet from 800 mg to 1200 mg. What is the effect on total cost?

8. Potatoes cost 12 cents/serving and have energy content of 300 kcal per serving, but no protein nor calcium content. Should they be part of the diet?

Exercise 72 Attached is the sensitivity report for the Workforce Planning Problem (see Section 5.4.2 in Chapter 5).

Adjustable Cells

Cell	Name	Final Value	Reduced Cost	Objective Coefficient	Allowable Increase	Allowable Decrease
$B$14	Shift1	4	0	1	0.5	1
$B$15	Shift2	7	0	1	0	0.333333
$B$16	Shift3	1	0	1	0.5	0
$B$17	Shift4	4	0	1	0.5	0
$B$18	Shift5	3	0	1	0	0.333333
$B$19	Shift6	3	0	1	0.5	1
$B$20	Shift7	0	0.333333	1	1E+30	0.333333

Constraints

Cell	Name	Final Value	Shadow Price	Constraint R.H. Side	Allowable Increase	Allowable Decrease
$B$24	Monday	14	0.333333	14	1.5	6
$B$25	Tuesday	17	0	13	4	1E+30
$B$26	Wednesday	15	0.333333	15	6	3
$B$27	Thursday	16	0	16	3	4
$B$28	Friday	19	0.333333	19	4.5	3
$B$29	Saturday	18	0.333333	18	1.5	6
$B$30	Sunday	11	0	11	4	1

Answer each of the following questions independently of the others.

1. What is the current total number of workers needed to staff the restaurant?

2. Due to a special offer, demand on thurdays increases. As a result, 18 workers are needed instead of 16. What is the effect on the total number of workers needed to staff the restaurant?

3. Assume that demand on mondays decreases: 11 workers are needed instead of 14. What is the effect on the total number of workers needed to staff the restaurant?

4. Currently, 15 workers are needed on wednesdays. In what range can this number vary without changing the optimal basis?

5. Currently, every worker in the restaurant is paid $1000 per month. So the objective function in the formulation can be viewed as total wage expenses (in thousand dollars). Workers have complained that Shift 4 is the least desirable shift. Management is considering increasing the wages of workers on Shift 4 to $1100. Would this change the optimal solution? What would be the effect on total wage expenses?

6. Shift 1, on the other hand, is very desirable (sundays off while on duty fridays and saturdays, which are the best days for tips). Management is considering reducing the wages of workers on Shift 1 to $ 900 per month. Would this change the optimal solution? What would be the effect on total wage expenses?

7. Management is considering introducing a new shift with the days off on tuesdays and sundays. Because these days are not consecutive, the wages will be $ 1200 per month. Will this increase or reduce the total wage expenses?

Chapter 9

Decision Theory

Decision theory treats decisions against nature. This refers to a situation where the result (return) from a decision depends on action of another player (nature). For example, if the decision is to carry an umbrella or not, the return (get wet or not) depends on what action nature takes. It is important to note that, in this model, the returns accrue only to the decision maker. Nature does not care what the outcome is. This condition distinguishes decision theory from game theory. In game theory, both players have an interest in the outcome.

The fundamental piece of data for decision theory problems is a *payoff table*:

decision	state of nature			
	1	2	...	m
d_1	r_{11}	r_{12}	...	r_{1m}
d_2	r_{21}	r_{22}	...	r_{2m}
...	...	...	...	...
d_n	r_{n1}	r_{n2}	...	r_{nm}

The entries r_{ij} are the payoffs for each possible combination of decision and state of nature. The decision process is the following.

- The decision maker selects one of the possible decisions $d_1, d_2, \ldots, d_n$. Say d_i.
- After this decision is made, a state of nature occurs. Say state j.
- The return received by the decision maker is r_{ij}.

The question faced by the decision maker is: which decision to select? The decision will depend on the decision maker's belief concerning what nature will do, that is, which state of nature will occur. If we believe state j will occur, we select the decision d_i associated with the largest number r_{ij} in column j of the payoff table. Different assumptions about nature's behavior lead to different procedures for selecting "the best" decision.

If we know which state of nature will occur, we simply select the decision that yields the largest return for the known state of nature. In practice, there may be infinitely many possible decisions. If these possible decisions are represented by a vector d and the return by the real-valued function $r(d)$, the decision problem can then be formulated as

$$\max r(d) \text{ subject to feasibility constraints on } d.$$

The linear programming model studied earlier in this course fits in this category. In this case, $r(d)$ is a linear function and the feasibility constraints are linear as well. Many other deterministic models of operations research also fall in this category.

In the remainder of this chapter, we consider decisions under risk.

9.1 Decisions Under Risk

We make the assumption that there is more than one state of nature and that the decision maker knows the probability with which each state of nature will occur. Let p_j be the probability that state j will occur. If the decision maker makes decision d_i, her expected return ER_i is

$$ER_i = r_{i1}p_1 + r_{i2}p_2 + \ldots + r_{im}p_m.$$

She will make the decision d_{i^*} that maximizes ER_i, namely

$$ER_{i^*} = \text{ maximum over all } i \text{ of } ER_i.$$

Example 9.1.1 *Let us consider the example of the newsboy problem: a newsboy buys papers from the delivery truck at the beginning of the day. During the day, he sells papers. Leftover papers at the end of the day are worthless. Assume that each paper costs 15 cents and sells for 50 cents and that the following probability distribution is known.*

$$\begin{aligned}
p_0 = \mathit{Prob}\ \{\ \mathit{demand} = 0\ \} &= 2/10 \\
p_1 = \mathit{Prob}\ \{\ \mathit{demand} = 1\ \} &= 4/10 \\
p_2 = \mathit{Prob}\ \{\ \mathit{demand} = 2\ \} &= 3/10 \\
p_3 = \mathit{Prob}\ \{\ \mathit{demand} = 3\ \} &= 1/10
\end{aligned}$$

How many papers should the newsboy buy from the delivery truck?

To solve this exercise, we first construct the payoff table. Here r_{ij} is the reward acheived when i papers are bought and a demand j occurs.

decision	state of nature			
	0	1	2	3
0	0	0	0	0
1	−15	35	35	35
2	−30	20	70	70
3	−45	5	50	105

Next, we compute the expected returns for each possible decision:

$$\begin{aligned}
ER_0 &= 0(2/10) + 0(4/10) + 0(3/10) + 0(1/10) = 0 \\
ER_1 &= -15(2/10) + 35(4/10) + 35(3/10) + 35(1/10) = 25 \\
ER_2 &= -40(2/10) + 10(4/10) + 60(3/10) + 60(1/10) = 30 \\
ER_3 &= -45(2/10) + 5(4/10) + 50(3/10) + 105(1/10) = 18.50
\end{aligned}$$

The maximum occurs when the newsboy buys 2 papers from the delivery truck. His expected return is then 30 cents.

The fact that the newsboy must make his buying decision *before* demand is realized has a considerable impact on his revenues. If he could first see the demand being realized each day and then buy the corresponding number of newspapers for that day, his expected return would increase by an amount known as the *expected value of perfect information*. Millions of dollars are spent every year on market research projects, geological tests etc, to determine what state of nature will occur in a wide variety of applications. The expected value of perfect information indicates the expected gain from any such endeavor and thus places an upper bound on the amount that should be spent in gathering information.

Let us compute the expected value of perfect information $EVPI$ for the above newsboy example. If demand were known before the buying decision is made, the newsboy's expected return would be

$$0(2/10) + 35(4/10) + 70(3/10) + 105(1/10) = 45.5$$

$$\text{So, } EVPI = 45.5 - 30 = 15.5$$

It should be pointed out that the criterion of maximizing expected return can sometimes produce unacceptable results. This is because it ignores downside risk. Most people are risk averse, which means they would feel that the loss of x dollars is more painful than the benefit obtained from the gain of the same amount. Decision theory deals with this problem by introducing a function that measures the "attractiveness" of money. This function is called the *utility function*. You are referred to 45-749 (managerial economics) for more on this notion. Instead of working with a payoff table containing the dollar amounts r_{ij}, one would instead work with a payoff table containing the utilities, say u_{ij}. The optimal decision d_{i^*} is that which maximizes the expected utility

$$EU_i = u_{i1}p_1 + u_{i2}p_2 + \ldots + u_{im}p_m$$

over all i.

9.2 Decision Trees

Example 9.2.1 *Company ABC has developed a new line of products. Top management is attempting to decide on the appropriate marketing and production strategy. Three strategies are being considered, which we will simply refer to as A (aggressive), B (basic) and C (cautious). The market conditions under study are denoted by S (strong) or W (weak). Management's best estimate of the net profits (in millions of dollars) in each case are given in the following payoff table.*

decision	*state of nature*	
	S	*W*
A	30	−8
B	20	7
C	5	15

Management's best estimates of the probabilities of a strong or a weak market are 0.45 and 0.55 respectively. Which strategy should be chosen?

Using the approach introduced earlier, we can compute the expected return for each decision and select the best one, just as we did for the newsboy problem.

$$\begin{aligned} ER_A = 30(0.45) - 8(0.55) &= 9.10 \\ ER_B = 20(0.45) + 7(0.55) &= 12.85 \\ ER_C = 5(0.45) + 15(0.55) &= 10.50 \end{aligned}$$

The optimal decision is to select B.

A convenient way to represent this problem is through the use of *decision trees*, as in Figure 9.1. A *square node* will represent a point at which a decision must been made, and each line leading from a square will represent a possible decision. A *circular node* will represent situations where the outcome is uncertain, and each line leading from a circle will represent a possible outcome.

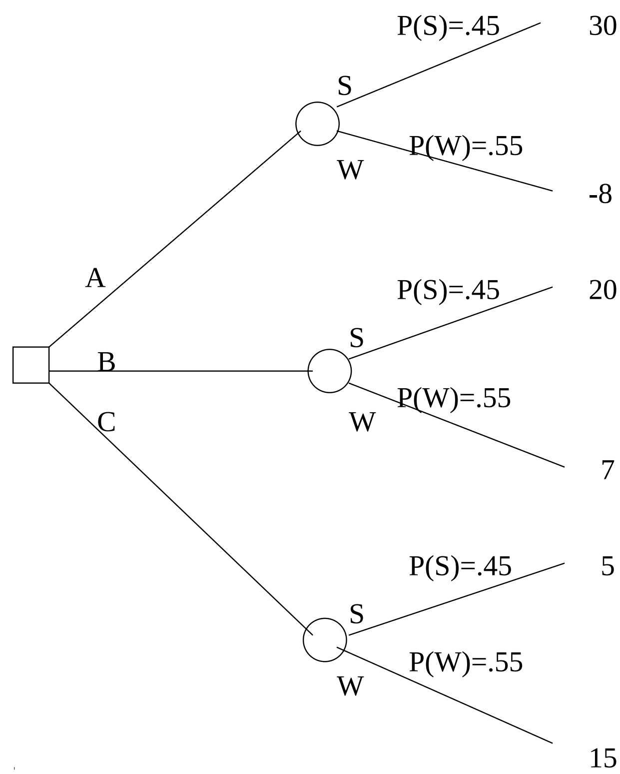


Figure 9.1: Decision Tree for the ABC Company

Using a decision tree to find the optimal decision is called solving the tree. To solve a decision tree, one works backwards. This is called *folding back* the tree. First, the terminal branches are folded back by calculating an expected value for each terminal node. See Figure 9.2.

Management now faces the simple problem of choosing the alternative that yields the highest expected terminal value. So, a decision tree provides another, more graphic, way of viewing the same problem. Exactly the same information is utilized, and the same calculations are made.

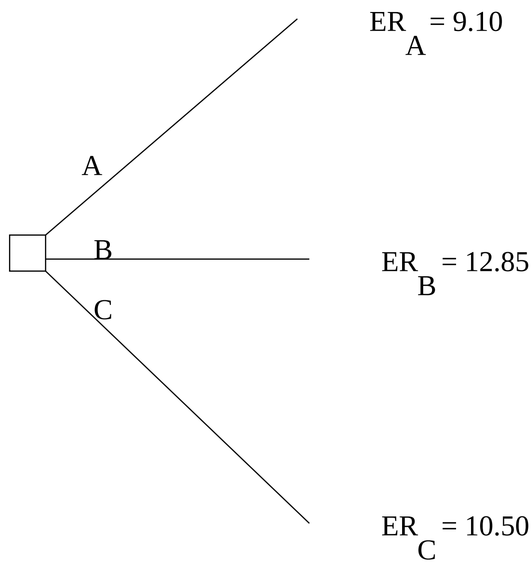


Figure 9.2: Reduced Decision Tree for the ABC Company

Sensitivity Analysis

The expected return of strategy A is

$$ER_A = 30P(S) - 8P(W)$$

or, equivalently,

$$ER_A = 30P(S) - 8(1 - P(W)) \quad = \quad -8 + 38P(S).$$

Thus, this expected return is a linear function of the probability that market conditions will be strong. Similarly

$$\begin{aligned} ER_B = 20P(S) + 7(1 - P(S)) &= 7 + 13P(S) \\ ER_C = 5P(S) + 15(1 - P(S)) &= 15 - 10P(S) \end{aligned}$$

We can plot these three linear functions on the same set of axes (see Figure 9.3).

This diagram shows that Company ABC should select the basic strategy (strategy B) as long as the probability of a strong market demand is between $P(S) = 0.348$ and $P(S) = 0.6$. This is reassuring, since the optimal decision in this case is not very sensitive to an accurate estimation of $P(S)$. However, if $P(S)$ falls below 0.348, it becomes optimal to choose the cautious strategy C, whereas if $P(S)$ is above 0.6, the agressive strategy A becomes optimal.

Sequential Decisions

Example 9.2.2 *Although the basic strategy B is appealing, ABC's management has the option of asking the marketing research group to perform a market research study. Within a month, this group can report on whether the study was encouraging (E) or discouraging (D). In the past, such studies have tended to be in the right direction: When market ended up being strong, such studies were encouraging 60% of the time and they were discouraging 40% of the time. Whereas, when market ended up being weak, these studies were discouraging 70% of the time and encouraging 30% of the time. Such a study would cost $500,000. Should management request the market research study or not?*

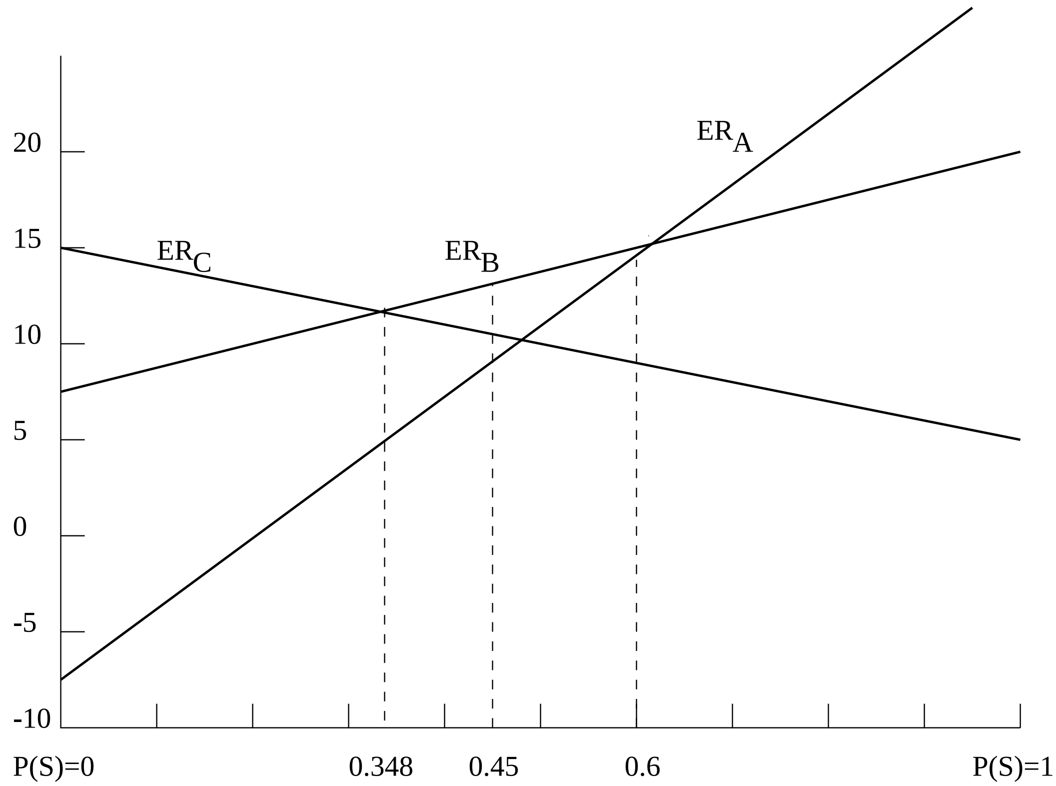


Figure 9.3: Expected Return as a function of P(S)

Let us construct the decision tree for this sequential decision problem. See Figure 9.4. It is important to note that the tree is created in the chronological order in which information becomes available. Here, the sequence of events is

- Test decision
- Test result (if any)
- Make decision
- Market condition.

The leftmost node correspond to the decision to test or not to test. Moving along the "Test" branch, the next node to the right is circular, since it corresponds to an uncertain event. There are two possible results. Either the test is encouraging (E), or it is discouraging (D). The probabilities of these two outcomes are $P(E)$ and $P(D)$ respectively. How does one compute these probabilities?

We need some fundamental results about probabilities. Refer to 45-733 for additional material. The information we are given is *conditional*. Given S, the probability of E is 60% and the probability of D is 40%. Similarly, we are told that, given W, the probability of E is 30% and the probability of D is 70%. We denote these *conditional probabilities* as follows

$$P(E|S) = 0.6 \quad P(E|W) = 0.3 \quad P(D|S) = 0.4 \quad P(D|W) = 0.7$$

In addition, we know $P(S) = 0.45$ and $P(W) = 0.55$. This is all the information we need to compute $P(E)$ and $P(D)$. Indeed, for events $S_1, S_2, \ldots, S_n$ that partition the space of possible

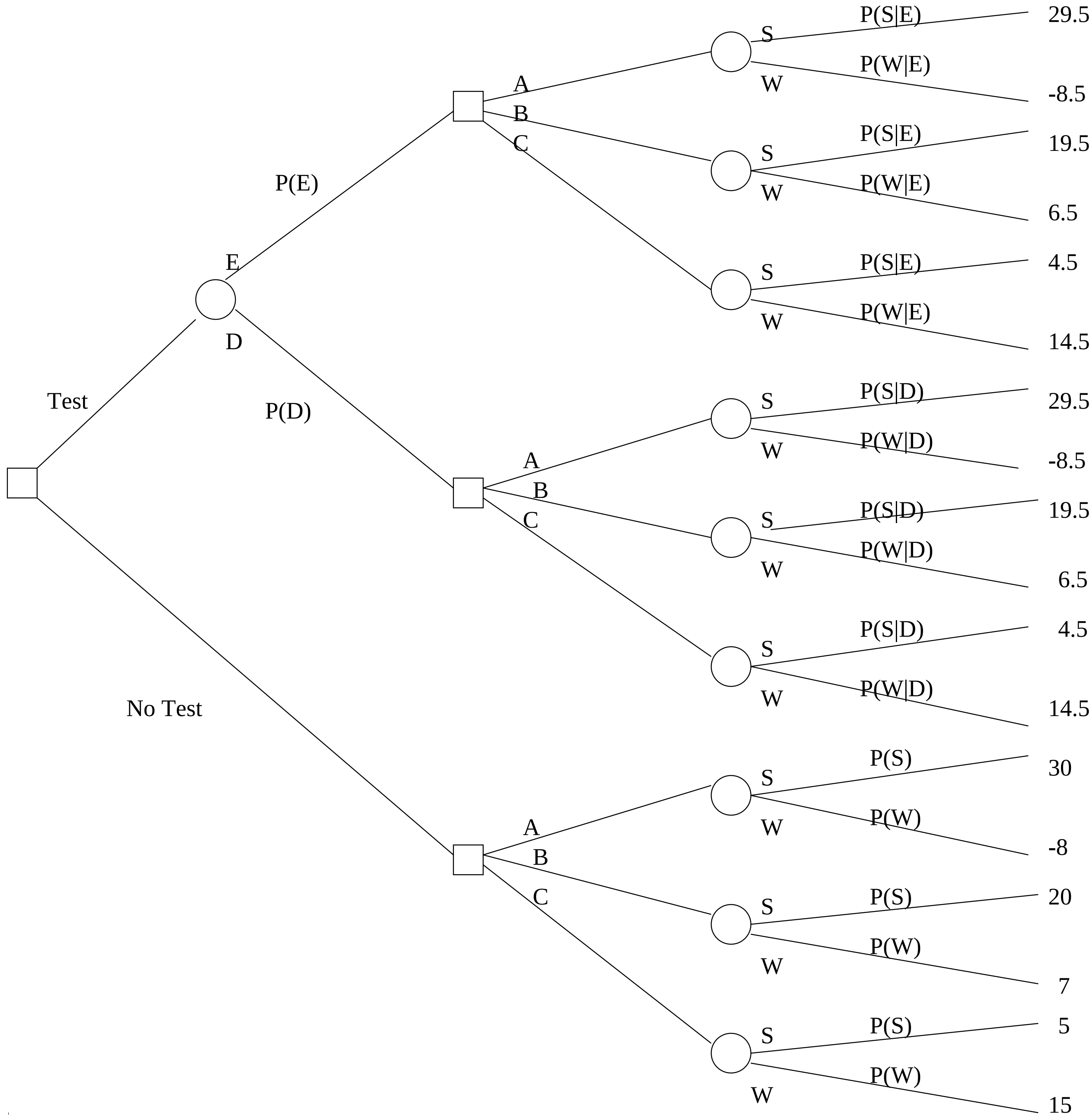


Figure 9.4: Test versus No-Test Decision Tree

outcomes and an event T, one has

$$P(T) = P(T|S_1)P(S_1) + P(T|S_2)P(S_2) + \ldots + P(T|S_n)P(S_n).$$

Here, this gives

$$\begin{aligned} P(E) &= P(E|S)P(S) + P(E|W)P(W) \\ &= (0.6)(0.45) + (0.3)(0.55) \\ &= 0.435 \end{aligned}$$

and

$$\begin{aligned} P(D) &= P(D|S)P(S) + P(D|W)P(W) \\ &= (0.4)(0.45) + (0.7)(0.55) \\ &= 0.565 \end{aligned}$$

As we continue to move to the right of the decision tree, the next nodes are square, corresponding to the three marketing and production strategies. Still further to the right are circular nodes corresponding to the uncertain market conditions: either weak or strong. The probability of these two events is now *conditional* on the outcome of earlier uncertain events, namely the result of the market reseach study, when such a study was performed. This means that we need to compute the following conditional probabilities: $P(S|E)$, $P(W|E)$, $P(S|D)$ and $P(W|D)$. These quantities are computed using the formula

$$P(R|T) = \frac{P(T|R)P(R)}{P(T)},$$

which is valid for any two events R and T. Here, we get

$$P(S|E) = \frac{P(E|S)P(S)}{P(E)} = \frac{(0.6)(0.45)}{0.435} = 0.621$$

Similarly,

$$P(S|E) = 0.379 \qquad P(S|D) = 0.318 \qquad P(W|D) = 0.682$$

Now, we are ready to solve the decision tree. As earlier, this is done by folding back. See Figures 9.5, 9.6 and 9.7. You fold back a circular node by calculating the expected returns. You fold back a square node by selecting the decision that yields the highest expected return. The expected return when the market research study is performed is 12.96 million dollars, which is greater than the expected return when no study is performed. So the study should be undertaken.

As a final note, let us compare the expected value of the study (denoted by $EVSI$, which stands for *expected value of sample information*) to the expected value of perfect information $EVPI$.

EVSI is computed without incorporating the cost of the study. So

$$EVSI = (12.96 + 0.5) - 12.85 = 0.61$$

whereas

$$EVPI = (30)(0.45) + (15)(0.55) - 12.85 = 21.75 - 12.85 = 8.90$$

We see that the market research study is not very effective. If it were, the value of $EVSI$ would be much closer to $EVPI$. Yet, its value is greater than its cost, so it is worth performing.

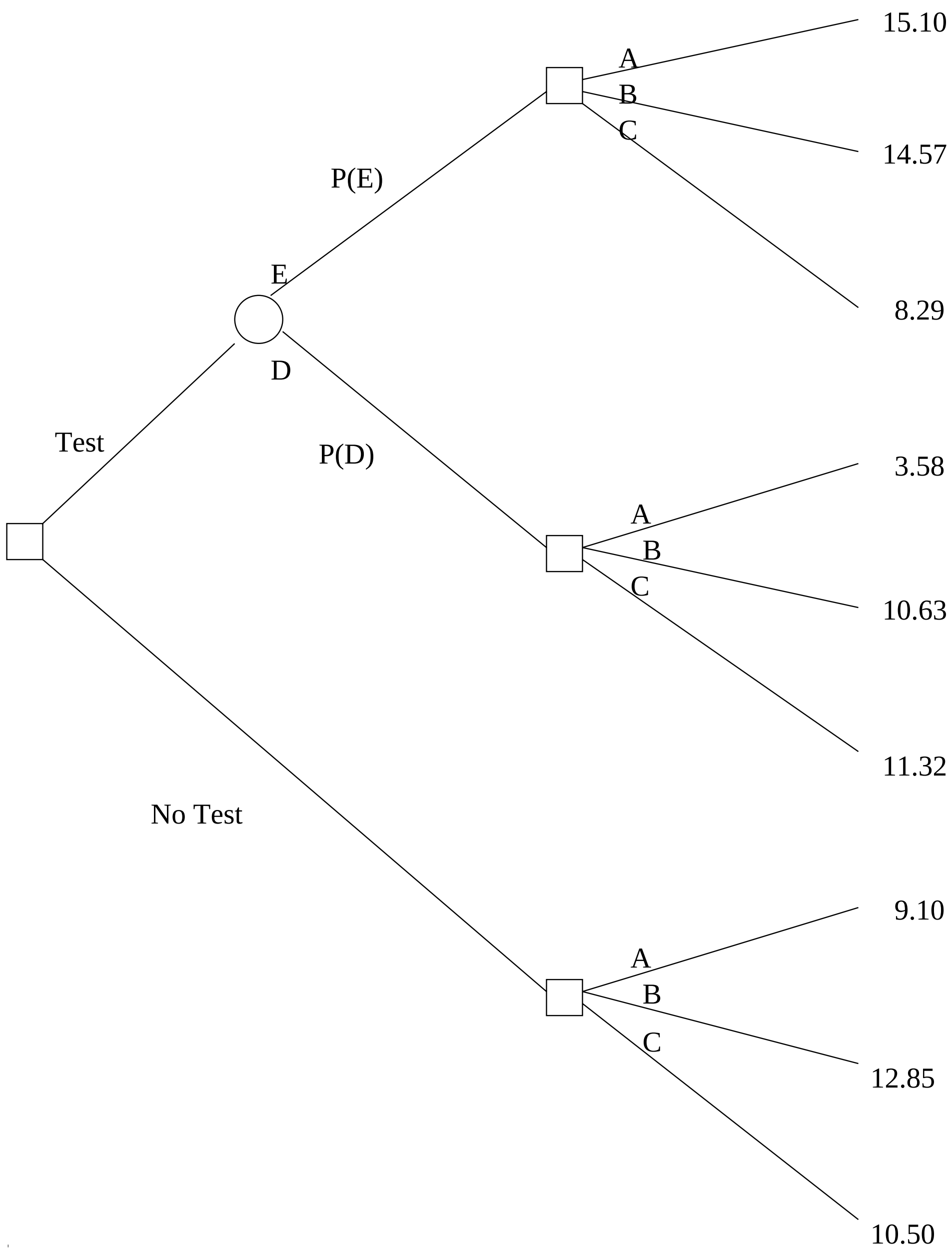


Figure 9.5: Solving the Tree

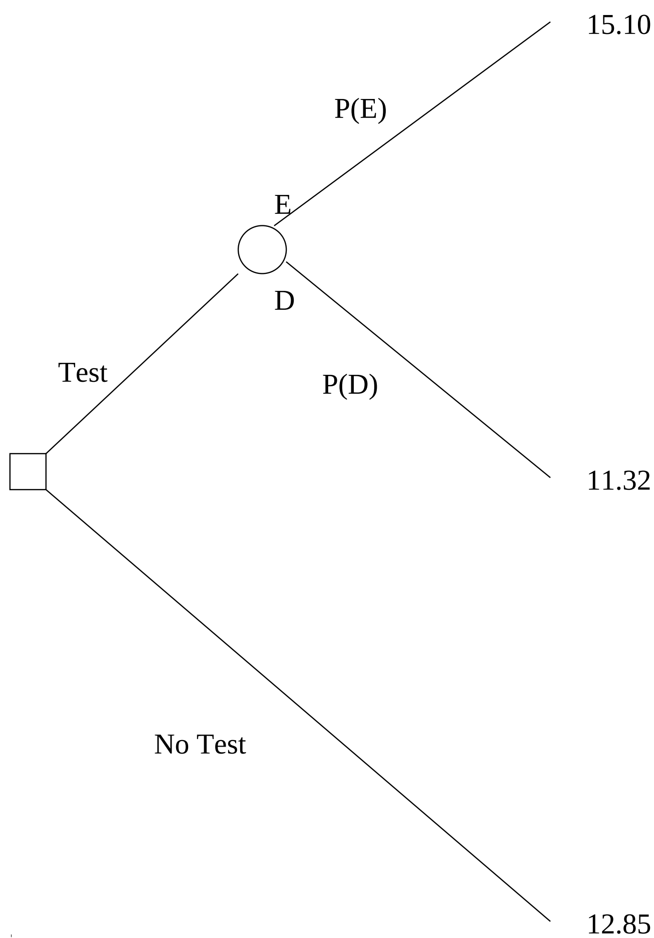


Figure 9.6: Solving the Tree

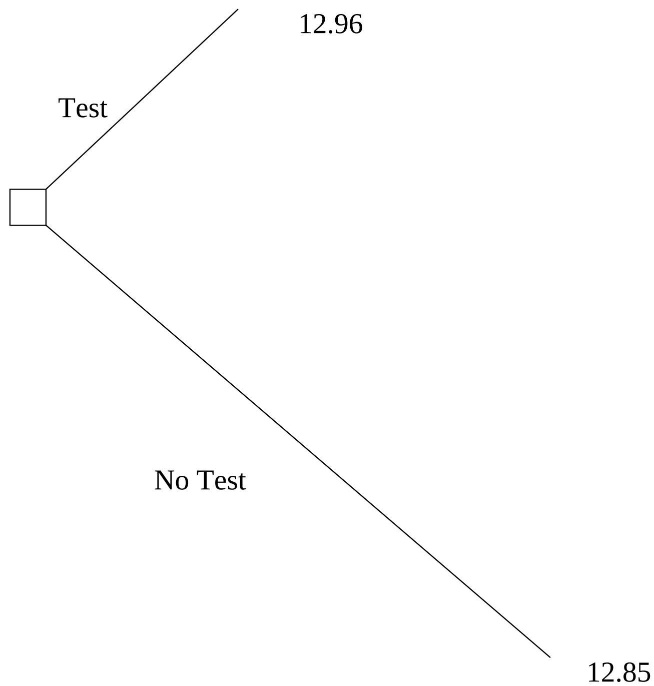


Figure 9.7: Solving the Tree

Example 9.2.3 An art dealer has a client who will buy the masterpiece *Rain Delay* for $50,000. The dealer can buy the painting now for $40,000 (making a profit of $10,000). Alternatively, he can wait one day, when the price will go down to $30,000. The dealer can also wait another day when the price will be $25,000. If the dealer does not buy by that day, then the painting will no longer be available. On each day, there is a 2/3 chance that the painting will be sold elsewhere and will no longer be available.

(a) Draw a decision tree representing the dealers decision making process.

(b) Solve the tree. What is the dealers expected profit? When should he buy the painting?

(c) What is the Expected Value of Perfect Information (value the dealer would place on knowing when the item will be sold)?

Answer:

(a)

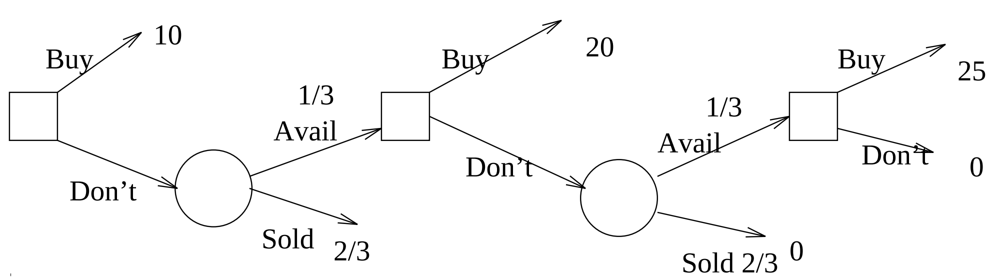


Figure 9.8: Solution

(b) *The value is 10, for an expected profit of $10,000. He should buy the painting immediately.*

(c) *With probability 2/3, the painting will be sold on the first day, so should be bought immediately. With probability 1/3(2/3) it will be sold on the second day, so should be bought after one day. Finally, with probability 1/3(1/3) it will not be sold on the first two days, so should be bought after two days. The value of this is 2/3(10)+1/3(2/3)20+1/3(1/3)25 = 13.89. The EVPI is therefore $3,889.*

Exercise 73 The Scrub Professional Cleaning Service receives preliminary sales contracts from two sources: its own agent and building managers. Historically, $\frac{3}{8}$ of the contracts have come from the Scrub agent and $\frac{5}{8}$ from building managers. Unfortunately, not all preliminary contracts result in actual sales contracts. Actually, only $\frac{1}{2}$ of those preliminary contracts received from building managers result in a sale, whereas $\frac{3}{4}$ of those received from the Scrub agent result in a sale. The net return to Scrub from a sale is $6400. The cost of processing and following up on a preliminary contract that does not result in a sale is $320. What is the expected return associated with a preliminary sales contract?

Exercise 74 Walter's Dog and Pony show is scheduled to appear in Cedar Rapids on July 4. The profits obtained are heavily dependent on the weather. In particular, if the weather is rainy, the show loses $28,000 and if sunny, the show makes a profit of $12,000. (We assume that all days are either rainy or sunny.) Walter can decide to cancel the show, but if he does, he forfeits a $1,000 deposit he put down when he accepted the date. The historical record shows that on July 4, it has rained $\frac{1}{4}$ of the time for the last 100 years.

(a) What decision should Walter make to maximize his expected net dollar return?

(b) What is the expected value of perfect information?

Walter has the option to purchase a forecast from Victor's Weather Wonder. Victor's accuracy varies. On those occasions when it has rained, Victor has been correct (i.e. predicted rain) 90% of the time. On the other hand, when it has been sunny, he has been right (i.e. he predicted sun) only 80% of the time.

(c) If Walter had the forecast, what strategy should he follow to maximize his expected net dollar return? (d) How much should Walter be willing to pay to have the forecast?

Exercise 75 Wildcat Oil is considering spending $100,000 to drill at a particular spot. The result of such a drilling is either a "Dry Well" (of no value), a "Wet Well" (providing $150,000 in revenues) or a "Gusher" (providing $250,000 in revenues). The probabilities for these three possibilities are .5, .3 and .2 respectively.

(a) Draw a decision tree for the problem of deciding whether to drill or not.

(b) Solve the decision tree assuming the goal is to maximize the expected net revenue. Should the company drill?

The following two problems should be done independently.

(c) SureFire Consultants is able to determine for certain the type of well before drilling. They offer to tell Wildcat the type for $50,000. Should Wildcat accept their offer?

(d) Close-enough Consultants offer to use their specialized seismic hammer. This hammer returns either encouraging or discouraging results. In the past, when applied to a Gusher, the hammer always returned encouraging results. When applied to a Wet Well, it was encouraging 75% of the time and discouraging 25% of the time. When applied to a Dry Well, it was encouraging one-third of the time and discouraging two-thirds of the time. What is the maximum Wildcat should pay Close-enough Consultants for use of their hammer?

Chapter 10

Game Theory

In previous chapters, we have considered situations where a *single* decision maker chooses an optimal decision without reference to its effect on other decision makers. In many business situations, however, the outcome for a given firm depends not just on what strategy it chooses, but also on what strategies its competitors choose. Game theory is an attempt to formalize such competitive situations. We present some of the main ideas in the context of Two-Person games and then discuss briefly n-Person Games.

10.1 Two-Person Games

Each player has several strategies. If the first player chooses Strategy i while the second player chooses Strategy j, then Player 1 gains a_{ij} while Player 2 gains b_{ij}. This outcome is represented by $(a_{ij}, b_{ij}$. 2-person games where the players' interests are completely opposed are called *zero-sum* or *constant-sum* games: one player's gain is the other player's loss. Games where the players' interests are not completely opposed are called *variable-sum* games. Such games arise in business on an everyday basis, and solving them is not an easy task. Certain 2-person games admit *pure strategies* whereas others require *mixed strategies*. A pure strategy is one where, each time the players play the game, they choose the same strategy. A mixed strategy is one where the players introduce a random element in their choice of a strategy, thus leaving the opponent guessing.

10.1.1 Pure Strategies

We start with a constant-sum game: for every possible outcome of the game, the utility u_1 of Player 1 plus the utility u_2 of Player 2, adds to a constant. For example, if two firms are competing for market shares, then $u_1 + u_2 = 100\%$.

Example 10.1.1 *(Battle of the Networks)*
Two television networks are battling for viewer shares. Viewer share is important because, the higher it is, the more money the network can make from selling advertising time during that program. Consider the following situation: the networks make their programming decisions independently and simultaneously. Each network can show either sports or a sitcom. Network 1 has a programming advantage in sitcoms and Network 2 has one in sports: If both networks show sitcoms, then Network 1 gets a 56% viewer share. If both networks show sports, then Network 2 gets a 54% viewer share. If Network 1 shows a sitcom and Network 2 shows sports, then Network 1 gets a 51% viewer share and Network 2 gets 49%. Finally, if Network 1 shows sports and Network 2 shows a sitcom, then each gets a 50% viewer share.

The possible outcomes are best represented in a table.

		Network 2	
		Sitcom	Sports
Network 1	Sitcom	(56%, 44%)	(51%, 49%)
	Sports	(50%, 50%)	(46%, 54%)

We can solve Battle of the Networks as follows: from Network 1's point of view, it is better to show a sitcom, whether Network 2 shows a sitcom or sports. The strategy "Show a Sitcom" is said to *dominate* the strategy "Show Sports" for Network 1. So Network 1 will show a sitcom. Similarly, Network 2 is better off showing sports whether Network 1 shows a sitcom or sports. In other words, Network 2 also has a dominating strategy. So Network 2 shows sports. The resulting outcome, namely 51% viewer share to Network 1 and 49% to Network 2, is an *equilibrium*, since neither of the two players in this game can unilaterally improve their outcome. If Network 1 were to switch from sitcom to sports, its viewer share would drop 5%, from 51% to 46%. If Network 2 were to switch from sports to sitcom, its viewer share would also drop, from 49% to 44%. Each network is getting the best viewer share it can, given the competition it is up against.

In a general 2-person game, Strategy i for Player 1 *dominates* Strategy k for Player 1 if $a_{ij} \geq a_{kj}$ for every j. Similarly, Strategy j for Player 2 *dominates* Strategy ℓ for Player 2 if $b_{ij} \geq b_{i\ell}$ for every i.

In a general 2-person game, Strategy i for Player 1 and Strategy j for Player 2 is an *equilibrium* if the corresponding outcome (a_{ij}, b_{ij}) has the following property: a_{ij} is the largest element a_{kj} in column j and b_{ij} is the largest element $b_{i\ell}$ in row i. When such a pair i, j exists, a pure strategy for each player provides an optimal solution to the game. When no such a pair exists, the players must play mixed strategies to maximize their gains (see Section 10.1.2).

There is an easy way to relate constant-sum games and zero-sum games. Let u_1 and u_2 be the payoffs of players 1 and 2, respectively, in a constant-sum game: $u_1 + u_2 = k$. Now consider a new set of payoffs of the form

$$\begin{aligned} v_1 &= u_1 - u_2 \\ v_2 &= u_2 - u_1 \end{aligned}$$

Clearly, $v_1 + v_2 = 0$, so we now have a zero-sum game. Using the relation $u_1 + u_2 = k$, we can rewrite the payoffs v_1 and v_2 as:

$$\begin{aligned} v_1 &= 2u_1 - k \\ v_2 &= 2u_2 - k \end{aligned}$$

These are positive, linear transformations of utility, and so, according to the expected utility theorem (see 45-749), they have no effect on decisions. For the Battle of the Networks example, the new table becomes

		Network 2	
		Sitcom	Sports
Network 1	Sitcom	(12%, −12%)	(2%, −2%)
	Sports	(0%, 0%)	(−8%, 8%)

It represents each network's advantage in viewer share over the other network. Whether you reason in terms of total viewer shares or in terms of advantage in viewer shares, the solution is the same.

Example 10.1.2 *Competitive Advantage*

In a technologically advanced economy like that of the United states, firms constantly encounter the following situation. A new technological advance becomes available. If one firm adopts the new technology, it gains an advantage over its competitors, a competitive advantage. *If all firms adopt the new technology, then the advantage vanishes. This is represented in the next table, where* a *measures the size of the competitive advantage.*

		Firm 2	
		Adopt	*Stay put*
Firm 1	*Adopt*	$(0, 0)$	$(a, -a)$
	Stay put	$(-a, a)$	$(0, 0)$

Each firm has two strategies, either Stay put, or Adopt the new technology. Firm 1 has an incentive to adopt the new technology: in the event Firm 2 stays put, then Firm 1 gets the competitive advantage a, and in the event Firm 2 adopts the new technology, then Firm 1 erases its competitive disadvantage $-a$. So, whatever Firm 2's decision is, Firm 1 is better off adopting the new technology. This is Firm 1's dominant strategy. Of course, the situation for Firm 2 is identical. So the equilibrium of Competitive Advantage is for both firms to adopt the new technology. As a result, both firms get a payoff of 0. This may seem like a paradox, since the payoffs would have been the same if both firms had stayed put. But, in Competitive Advantage, neither firm can afford to stay put. The firms in this game are driven to adopt any technology that comes along. Take, for example, the hospital industry. Magnetic Resonance Imaging (MRI) is a new technology that enhances conventional X rays. It allows doctors to see body damage in ways that were not previously possible. Once MRI became available, any hospital that installed an MRI unit gained a competitive advantage over other hospitals in the area. From a public policy standpoint, it may not make much sense for every hospital to have its own MRI unit. These units are expensive to buy and to operate. They can eat up millions of dollars. Often, one MRI unit could handle the traffic of several hospitals. But an individual hospital would be at a competitive disadvantage if it did not have its own MRI. As long as hospitals play Competitive Advantage, they are going to adopt every new technology that comes along.

The two-person games we have encountered so far have had unique pure strategy equilibria. However, a two-person zero-sum game may have multiple equilibria. For example, consider the game:

		Player 2		
		A	B	C
Player 1	A	$(0,0)$	$(1,-1)$	$(0,0)$
	B	$(-1,1)$	$(0,0)$	$(-1,1)$
	C	$(0,0)$	$(1,-1)$	$(0,0)$

Each player can play indifferently strategy A or C and so there are four pure strategy equilibria, corresponding to the four corners of the above table. Note that these equilibria all have the same payoff. This is not a coincidence. It can be shown that this is always the case. *Every equilibrium of a 2-person zero-sum game has the same payoff.* For pure strategy equilibria, there is a simple proof: Suppose two equilibria had payoffs $(a,-a)$ and $(b,-b)$ where $a \neq b$. If these two solutions lied on the same row or column, we would get an immediate contradiction to the definition of an equilibrium. Let $(a,-a)$ lie in row i and column k and let $(b,-b)$ lie in row j and column ℓ. The subtable corresponding to these two rows and columns looks as follows

		Player 2	
		k	ℓ
Player 1	i	$(a,-a)$	$(c,-c)$
	j	$(d,-d)$	$(b,-b)$

Since $(a,-a)$ is an equilibrium, we must have $a \geq d$ (from Player 1) and $-a \geq -c$ (from Player 2). Similarly, since $(b,-b)$ is an equilibrium, we must have $b \geq c$ and $-b \geq -d$. Putting these inequalities together, we get

$$a \geq d \geq b \geq c \geq a$$

This implies that $a = b$, which completes the proof.

Example 10.1.3 *Cigarette Advertising on Television*

On January 1, 1971, cigarette advertizing was banned on american television. It came as something of a surprise to the industry, and a welcome one at that, that profits rose by \$91 million in 1971. Why did this happen? Consider the strategic interaction in the industry before and after the agreement went into effect. Although there were four large tobacco companies involved, let us consider the strategic interaction between only two of them, for simplicity. In 1970, the payoff matrix was:

		Company 2	
		No TV ads	*TV Ads*
Company 1	*No TV Ads*	$(50, 50)$	$(20, 60)$
	TV Ads	$(60, 20)$	$(27, 27)$

It is clear from this table that advertising on television is a powerful marketing tool. If Company 1 switches from not advertising to advertising on television, its profits go up 20%, everything else being equal, when Company 2 does not advertise on TV, and they go up 35% when Company 2 advertises on TV; the same is true if the roles are reversed. In other words, advertising on television is a dominant strategy for each of the companies. Hence the equilibrium is when both companies advertise on television, with a payoff of $27 million. We say that an outcome (a_{ij}, b_{ij}) is *efficient* if there is no other outcome (a_{kl}, b_{kl}) that pays both players at least as much, and one or both players strictly more. Namely, the pair i, j is efficient if there is no pair k, l satisfying $a_{kl} \geq a_{ij}$ and $b_{kl} \geq b_{ij}$ and $a_{kl} + b_{kl} > a_{ij} + b_{ij}$. Obviously, the equilibrium in the cigarette advertizing game is not efficient. There is one efficient outcome, however, in this game: that is when neither company advertises on TV. Then both companies enjoy a payoff of $50 million. The ban of cigarette advertizing on television in 1971 only left the strategy "No TV Ads" and forced the industry into the efficient outcome! In large part, this is why profits went up in 1971.

It often happens in variable-sum games that the solution is not efficient. One of the biggest differences between constant-sum and variable-sum games is that solutions to the former are always efficient, whereas solutions to the latter rarely are.

10.1.2 Mixed Strategies

Example 10.1.4 *Market Niche*

Two firms are competing for a single market niche. If one firm occupies the market niche, it gets a return of 100. If both firms occupy the market niche, each loses 50. If a firm stays out of the market, it breaks even. The payoff table is:

		Firm B	
		Enter	*Stay out*
Firm A	*Enter*	$(-50, -50)$	$(100, 0)$
	Stay out	$(0, 100)$	$(0, 0)$

This game has two pure strategy equilibria, namely one of the two firms enters the market niche and the other stays out. But, unlike the games we have encountered thus far, neither player has a dominant strategy. When a player has no dominant strategy, she should consider playing a *mixed strategy.* In a mixed strategy, each of the various pure strategies is played with some probability, say p_1 for Strategy 1, p_2 for Strategy 2, etc with $p_1 + p_2 + \ldots = 1$. What would be the best mixed strategies for Firms A and B? Denote by p_1 the probability that Firm A enters the market niche. Therefore $p_2 = 1 - p_1$ is the probability that Firm A stays out. Similarly, Firm B enters the niche with probability q_1 and stays out with probability $q_2 = 1 - q_1$. The key insight to a mixed strategy equilibrium is the following. *Every pure strategy that is played as part of a mixed strategy equilibrium has the same expected value.* If one pure strategy is expected to pay less than another, then it should not be played at all. The pure strategies that are not excluded should be expected to pay the same. We now apply this principle. The expected value of the "Enter" strategy for Firm A, when Firm B plays its mixed strategy, is

$$EV_{Enter} = -50q_1 + 100q_2.$$

The expected value of the "Stay out" strategy for Firm A is $EV_{Stay\ out} = 0$. Setting $EV_{Enter} = EV_{Stay\ out}$ we get

$$-50q_1 + 100q_2 = 0.$$

Using $q_1 + q_2 = 1$, we obtain

$$q_1 = \frac{2}{3} \qquad q_2 = \frac{1}{3}.$$

Similarly,

$$p_1 = \frac{2}{3} \qquad p_2 = \frac{1}{3}.$$

As you can see, the payoffs of this mixed strategy equilibrium, namely $(0, 0)$, are inefficient. One of these firms could make a lot of money by entering the market niche, if it was sure that the other would not enter the same niche. This assurance is precisely what is missing. Each firm has exactly the same right to enter the market niche. The only way for *both* firms to exercise this right is to play the inefficient, but symmetrical, mixed strategy equilibrium. In many industrial markets, there is only room for a few firms – a situation known as *natural oligopoly.* Chance plays a major role in the identity of the firms that ultimately enter such markets. If too many firms enter, there are losses all around and eventually some firms must exit. From the mixed strategy equilibrium, we can actually predict how often two firms enter a market niche when there is only room for one: with the above data, the probability of entry by either firm is 2/3, so the probability that both firms enter is $(2/3)^2 = 4/9$. That is a little over 44% of the time! This is the source of the inefficiency. The efficient solution has total payoff of 100, but is not symmetrical. The fair solution pays each player the same but is inefficient. These two principles, efficiency and fairness, cannot be reconciled in a game like Market Niche. Once firms recognize this, they can try to find mecanisms to reach the efficient solution. For example, they may consider *side payments.* Or firms might simply attempt to scare off competitors by announcing their intention of moving into the market niche before they actually do so.

An important case where mixed strategies at equilibrium are always efficient is the case of constant-sum games. Interestingly, the optimal mixed strategies can be computed using linear programming, one linear program for each of the two players. We illustrate these ideas on an example.

Example 10.1.5 *Stone, Paper, Scissors*

Each of two players simultaneously utters one of the three words stone, paper, *or* scissors. *If both players utter the same word, the game is a draw. Otherwise, one player pays \$1 to the other according to the following: Scissors defeats (cuts) paper, paper defeats (covers) stone, and stone defeats (breaks) scissors. Find the optimal strategies for this game.*

The payoff matrix is as follows.

		Player 2		
		Stone	Paper	Scissors
Player 1	Stone	$(0,0)$	$(-1,1)$	$(1,-1)$
	Paper	$(1,-1)$	$(0,0)$	$(-1,1)$
	Scissors	$(-1,1)$	$(1,-1)$	$(0,0)$

Suppose Player 1 chooses the mixed strategy (x_1, x_2, x_3), where

$$
\begin{aligned}
x_1 &= \text{Probability that Player 1 chooses Stone} \\
x_2 &= \text{Probability that Player 1 chooses Paper} \\
x_3 &= \text{Probability that Player 1 chooses Scissors}
\end{aligned}
$$

The expected payoff of Player 1 is $x_2 - x_3$ when Player 2 plays Stone, $x_3 - x_1$ when Player 2 plays Paper and $x_1 - x_2$ when Player 2 plays Scissors. The best Player 2 can do is to achieve the minimum of these three values, namely $v = min(x_2 - x_3, x_3 - x_1, x_1 - x_2)$. Player 1, on the other hand, should try to make the quantity v as *large* as possible. So, Player 1 should use probabilities x_1, x_2 and x_3 that maximize $v = min(x_2 - x_3, x_3 - x_1, x_1 - x_2)$. This can be written as the following linear program.

$$
\begin{aligned}
\max\ & v \\
v &\leq x_2 - x_3 \\
v &\leq x_3 - x_1 \\
v &\leq x_1 - x_2 \\
& x_1 + x_2 + x_3 = 1 \\
& x_1, x_2, x_3 \geq 0.
\end{aligned}
$$

Similarly, we obtain Player 2's optimal mixed strategy (y_1, y_2, y_3), where

$$
\begin{aligned}
y_1 &= \text{Probability that Player 2 chooses Stone} \\
y_2 &= \text{Probability that Player 2 chooses Paper} \\
y_3 &= \text{Probability that Player 2 chooses Scissors}
\end{aligned}
$$

by solving the linear program

$$
\begin{aligned}
\max\ & z \\
z &\leq y_2 - y_3 \\
z &\leq y_3 - y_1 \\
z &\leq y_1 - y_2 \\
& y_1 + y_2 + y_3 = 1 \\
& y_1, y_2, y_3 \geq 0.
\end{aligned}
$$

For the Stone, Paper, Scissors game, the optimal solutions are $x_1^* = x_2^* = x_3^* = 1/3$ and $y_1^* = y_2^* = y_3^* = 1/3$ with $v^* = z^* = 0$.

In general, for any 2-person zero-sum game, let v^* and w^* be the optima of the linear programs for Players 1 and 2 respectively. Because of the special form of these linear programs, it can be shown that $z^* = -v^*$. So, the payoff at equilibrium is $(v^*, -v^*)$. The fact that, for each player, the equilibrium solutions are the optimal solutions of a linear has another consequence: *All equilibrium solutions have the same payoff.*

10.2 n-Person Games

When there are more than two players, fundamental differences arise. For example, the theorem about 2-person zero-sum games which says that all equilibrium solutions have the same payoff cannot be generalized to three players. Once there are more than two players, a zero-sum game can have equilibrium solutions that do not have the same payoffs. Here is an example with three players: each player has two strategies, say A and B. Whatever Players 1 and 2 play, the payoff vector is $(1, \frac{-1}{2}, \frac{-1}{2})$ when Player 3 chooses Strategy A and it is $(\frac{-1}{2}, 1, \frac{-1}{2})$ when Player 3 chooses Strategy B. We will call the player that receives the payoff of 1 the *winner* and the players that receive a payoff of $\frac{-1}{2}$, the *loosers*. In this game, although Player 3 can never be the winner, he determines which of Players 1 or 2 will be the winner. When this situation occurs, we say that Player 3 is a *spoiler*. For example, take the 1992 presidential election: 1=Bush, 2=Clinton, 3=Perot. Perot could not win, but he could play the spoiler in this campaign. In a two-way race, noone can play a spoiler; in a three-way race, there is room for a spoiler.

Example 10.2.1 *Competitive Advantage with Three Firms*

As before, each firm has the choice between staying put or adopting the new technology. If no firm adopts the new technology, then there is no competitive advantage and the payoff vector is $(0, 0, 0)$. *If exactly one firm adopts the new technology, then that firm gets the competitive advantage a, while each firm at a competitive disadvantage looses* $\frac{a}{2}$. *Thus if only Firm 1 adopts the new technology, then the payoff vector is* $(a, \frac{-a}{2}, \frac{-a}{2})$. *You can think of this as Firm 1 taking market share from both Firm 2 and Firm 3. If exactly two firms adopt the new technology, then these two firms split the competitive advantage, each gaining* $\frac{a}{2}$, *and the firm at a disadvantage looses a. Finally, if all firms adopt the new technology, there is no competitive advantage and the payoff vector is* $(0, 0, 0)$.

We can represent the payoffs in a table as follows.

Firm 3 Adopt

		Firm 2	
		Adopt	Stay put
Firm 1	Adopt	$(0, 0, 0)$	$(\frac{a}{2}, -a, \frac{a}{2})$
	Stay put	$(-a, \frac{a}{2}, \frac{a}{2})$	$(\frac{-a}{2}, \frac{-a}{2}, a)$

Firm 3 Stay put

		Firm 2	
		Adopt	Stay put
Firm 1	Adopt	$(\frac{a}{2}, \frac{a}{2}, -a)$	$(a, \frac{-a}{2}, \frac{-a}{2})$
	Stay put	$(\frac{-a}{2}, a, \frac{-a}{2})$	$(0, 0, 0)$

Each firm has a dominant strategy, which is to adopt the new technology. So the unique equilibrium occurs when all three firms play the pure strategy: Adopt the new technology. This is presicely what happened with two firms. No firm can be left behind in the race to adopt the new technology. This is just as true for n players as it is for two or three. So, here, the insight from the 2-person game version of Competitive Advantage generalizes nicely to n-person games.

An important aspect of n-person games is that subsets of players can form coalitions. A useful concept in this case is the *core* of the game. Another important aspect of game theory is that of fairness: the Shapley value provides an answer. Game theory also considers games with a sequential structure, games with imperfect information, market games (you learned about Cournot and Bertrand competition in 45-749). To learn more about these and other aspects of game theory, you are referred to the excellent book "Games for Business and Economics" by Roy Gardner (1995), which was used as a basis for this chapter.

Exercise 76 Here is a game that arises in resource economics.

(a) There are two fishing spots, one good and the other excellent. The good spot has 6 fish whereas the excellent spot has 10 fish. Two fishermen know how good each of the spots is. Each fisherman must choose one of the two spots to go fishing. If they end up going to the same spot, they divide the catch equally. If they go to different spots, the fisherman at the good spot gets 6 fish, the one at the excellent spot gets 10. Find three equilibria for this game.

(b) In such fishing games, conflicts often break over the fishing grounds. How might an assymetrical equilibrium, such as the one provided by the 200-mile fishing limit among nations, lead to a more efficient outcome?

Exercise 77 Find each player's optimal strategy and the value of the two-person zero-sum game with the following payoff table.

		Player 2			
		S1	S2	S3	S4
Player 1	T1	$(4,-4)$	$(5,-5)$	$(1,-1)$	$(4,-4)$
	T2	$(2,-2)$	$(1,-1)$	$(6,-6)$	$(3,-3)$
	T3	$(1,-1)$	$(0,0)$	$(0,0)$	$(2,-2)$

Chapter 11

Network Optimization

11.1 Introduction

Network optimization is a special type of linear programming model. Network models have three main advantages over linear programming:

1. They can be solved very quickly. Problems whose linear program would have 1000 rows and 30,000 columns can be solved in a matter of seconds. This allows network models to be used in many applications (such as real-time decision making) for which linear programming would be inappropriate.
2. They have naturally integer solutions. By recognizing that a problem can be formulated as a network program, it is possible to solve special types of integer programs without resorting to the ineffective and time consuming integer programming algorithms.
3. They are intuitive. Network models provide a language for talking about problems that is much more intuitive than the "variables, objective, and constraints" language of linear and integer programming.

Of course these advantages come with a drawback: network models cannot formulate the wide range of models that linear and integer programs can. However, they occur often enough that they form an important tool for real decision making.

11.2 Terminology

A *network* or *graph* consists of points, and lines connecting pairs of points. The points are called *nodes* or vertices. The lines are called *arcs*. The arcs may have a direction on them, in which case they are called *directed arcs*. If an arc has no direction, it is often called an *edge*. If all the arcs in a network are directed, the network is a *directed network*. If all the arcs are undirected, the network is an *undirected network*.

Two nodes may be connected by a series of arcs. A *path* is a sequence of distinct arcs (no nodes repeated) connecting the nodes. A *directed path* from node i to node j is a sequence of arcs, each of whose direction (if any) is towards j. An *undirected path* may have directed arcs pointed in either direction.

A path that begins and ends at the same node is a *cycle* and may be either directed or undirected.

A network is *connected* if there exists an undirected path between any pair of nodes. A connected network without any cycle is called a *tree*, mainly because it looks like one.

11.3 Examples

There are many examples of using network flows in practice. Here are a few:

11.3.1 Shortest Paths

Consider a phone network. At any given time, a message may take a certain amount of time to traverse each line (due to congestion effects, switching delays, and so on). This time can vary greatly minute by minute and telecommunication companies spend a lot of time and money tracking these delays and communicating these delays throughout the system. Assuming a centralized switcher knows these delays, there remains the problem of routing a call so as to minimize the delays. So, in figure 11.1, what is the least delay path from LA to Boston? How can we find that path quickly?

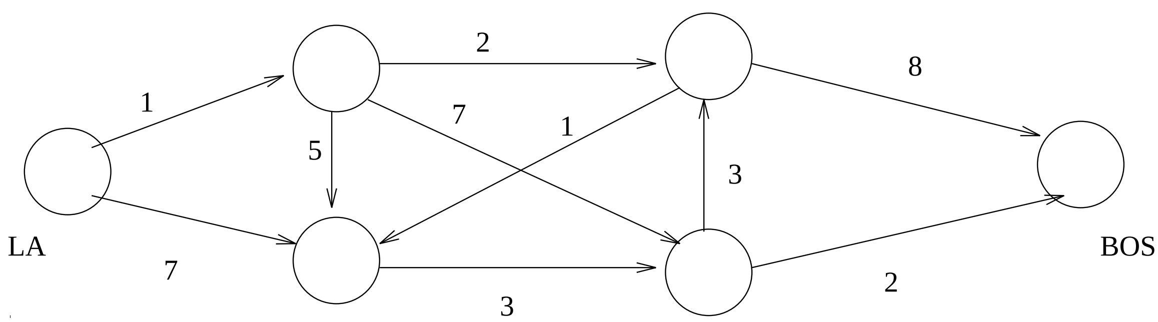


Figure 11.1: Phone Network

This is an example of a particular type of network model, called the *shortest path problem*. In such a problem, you have a network with costs on the edges and two special nodes: a start node and a finish node. The goal is to find a path from the start node to the finish node whose total weight is minimized.

Here is another problem that might not appear to be a shortest path problem, but really is:

At a small but growing airport, the local airline company is purchasing a new tractor for a tractor-trailer train to bring luggage to and from the airplanes. A new mechanized luggage system will be installed in 3 years, so the tractor will not be needed after that. However, because it will receive heavy use, and maintenance costs are high, it may still be more economical to replace the tractor after 1 or 2 years. The following table gives the total net discounted cost associated with purchasing a tractor in year i and trading it in in year j (where year 0 is now):

	j = 1	2	3
i = 0	8	18	31
1	—	10	21
2	—	—	12

The problem is to determine at what times the tractor should be replaced (if ever) to minimize the total costs for tractors. How can this be formulated as a shortest path problem?

11.3.2 Maximum Flow

Another type of model again has a number on each arc, but now the number corresponds to a capacity. This limits the flow on the arc to be no more than that value. For instance, in a distribution system, the capacity might be the limit on the amount of material (in tons, say) that can go over a particular distribution channel. We would then be concerned with the capcity of the network: how much can be sent from a source node to the destination node? Using the same network as above, treating the numbers as capacities, how much can be sent from LA to Boston?

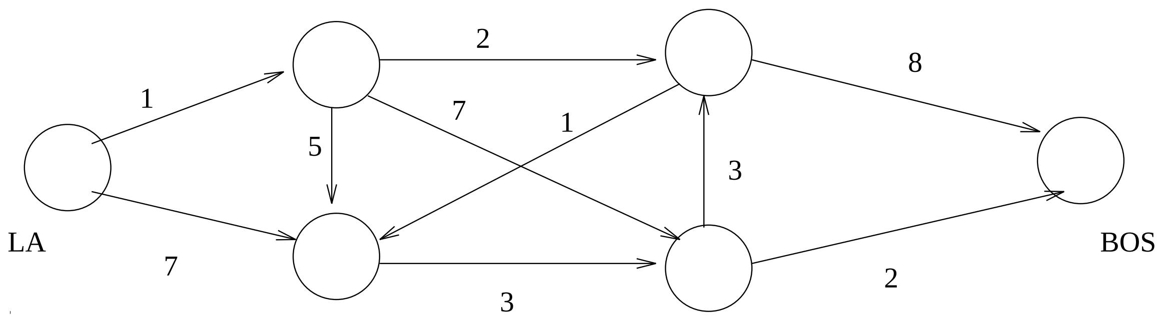


Figure 11.2: Distribution Network

Associated with the maximum flow is a bottleneck: a set of arcs whose total capacity equals the maximum flow, and whose removal leaves no path from source to destination in the network. It is actually a nontrivial result to show that the maximum flow equals the size of the minimum bottleneck. It is also an interesting task to find the bottleneck in the above example.

Maximum flow models occur in applications where cost is not an issue, and the goal is to maximize the number of items handled (in a broad sense). Here is a similar problem that can be formulated as a maximum flow problem (this was part of a conversation with a Chicago consulting firm on how to improve their internal operations):

Over the next three months, there are four projects to be completed. Project A requires 8 person-months of work, and must be done by the end of the second month. Project B requires 3 person-months and must be done by end of the first month. Project C can start at the beginning of the second month, and requires 7 person-months of work, no more than 2 of which are in the second month. Project D requires 6 person-months, no more than 3 of which occur in any month. There are 9 consultants available in the first and second month, and 6 in the third. Is it possible to meet the project deadlines? Surprisingly, this is a maximum flow problem. One nice (and surprising) aspect of the result is that this shows that no consultant need split a month's work between jobs.

Exercise 78 The Smith's, Jones', Johnson's, and Brown's are attending a picnic. There are four members of each family. Four cars are available to transport them: two cars that can hold four people, and two that can hold three. No more than 2 members of any family can be in any car. The goal is to maximize the number of people who can attend the picnic. Formulate this as a maximum flow problem. Give the nodes and arcs of the network and state which is the starting node and which is the destination. For each arc, give its capacity.

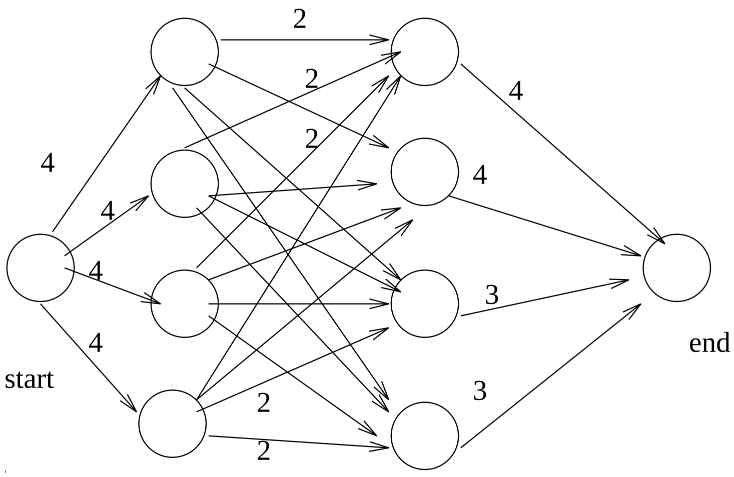


Figure 11.3: Solution of the exercise.

11.3.3 Transportation Problem

Consider the following snow removal problem: there are a number of districts in a city. After a snowfall, the snow in each area must be moved out of the district into a convenient location. In Montreal (from where this example is taken), these locations are large grates (leading to the sewer system), a couple large pits, and a couple entry points to the river. Each of these destinations has a capacity. The goal is to minimize the distance traveled to handle all of the snow.

This problem is an example of a transportation problem. In such a problem, there are a set of nodes called sources, and a set of nodes called destinations. All arcs go from a source to a destination. There is a per–unit cost on each arc. Each source has a supply of material, and each destination has a demand. We assume that the total supply equals the total demand (possibly adding a fake source or destination as needed). For the snow removal problem, the network might look like that in figure 11.4.

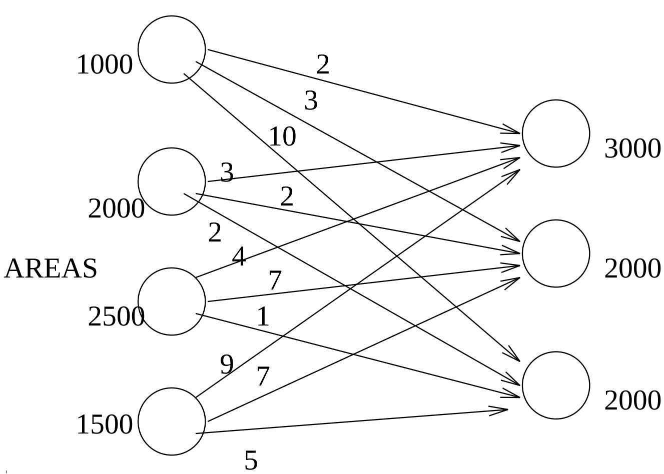


Figure 11.4: Snow Transportation Network

Transportation problems are often used in, surprise, transportation planning. For instance, in an application where goods are at a warehouse, one problem might be to assign customers to a warehouse so as to meet their demands. In such a case, the warehouses are the sources, the customers are the destinations, and the costs represent the per–unit transportation costs.

Here is another example of this:

One of the main products of P&T Company is canned peas. The peas are prepared at three canneries (near Bellingham, Washington; Eugene, Oregon; and Albert Lea, Minnesota) and are then shipped by truck to four distributing warehouses in Sacremento, California; Salt Lake City, Utah; Rapid City, South Dakota; and Albuquerque, New Mexico. Because shipping costs are a major expense, management has begun a study to reduce them. For the upcoming season, an estimate has been made of what the output will be from each cannery, and how much each warehouse will require to satisfy its customers. The shipping costs from each cannery to each warehouse has also been determined. This is summarized in the next table.

Shipping cost	Warehouse				
per truckload	1	2	3	4	Output
1	464	513	654	867	75
Cannery 2	352	416	690	791	125
3	995	682	388	685	100
Requirement	80	65	70	85	

You should find it an easy exercise to model this as a linear program. If we let x_{ij} be the number of truckloads shipped from cannery i to warehouse j, the problem is to

$$\begin{array}{ll} \text{minimize} & 464x_{11} + 513x_{12} + 654x_{13} + 867x_{14} + 352x_{21} + \ldots + 685x_{34} \\ \text{subject to} & \\ & x_{11} + x_{12} + x_{13} + x_{14} = 75 \\ & x_{21} + x_{22} + x_{23} + x_{24} = 125 \\ & x_{31} + x_{32} + x_{33} + x_{34} = 100 \\ & x_{11} + x_{21} + x_{31} = 80 \\ & x_{12} + x_{22} + x_{32} = 65 \\ & x_{13} + x_{23} + x_{33} = 70 \\ & x_{14} + x_{24} + x_{34} = 85 \\ & x_{ij} \geq 0 \text{ for all } i \text{ and } j. \end{array}$$

This is an example of the **transportation model**. As has been pointed out, this problem has a lot of nice structure. All the coefficients are 1 and every variable appears in exactly two constraints. It is this structure that lets the simplex algorithm be specialized into an extremely efficient algorithm.

What defines a transportation model? In general, the transportation model is concerned with distributing (literally or figuratively) a commodity from a group of supply centers, called **sources** to a group of receiving centers, called **destinations** to minimize total cost.

In general, source i has a supply of s_i units, and destination j has a demand for d_j units. The cost of distributing items from a source to a destination is proportional to the number of units. This data can be conveniently represented in a table like that for the sample problem.

We will generally assume that the total supply equals the total demand. If this is not true for a particular problem, *dummy* sources or destinations can be added to make it true. The text refers to such a problem as a *balanced transportation problem.* These dummy centers may have zero distribution costs, or costs may be assigned to represent unmet supply or demand.

For example, suppose that cannery 3 makes only 75 truckloads. The total supply is now 25 units too little. A dummy supply node can be added with supply 25 to balance the problem, and the cost from the dummy to each warehouse can be added to represent the cost of not meeting the warehouse's demand.

The transportation problem has a couple of nice properties:

Feasibility. As long as the supply equals demand, there exists a feasible solution to the problem.

Integrality. If the supplies and demands are integer, every basic solution (including optimal ones) have integer values. Therefore, it is not necessary to resort to integer programming to find integer solutions. Linear programming suffices. Note that this does not mean that each destination will be supplied by exactly one source.

In the pea shipping example, a basic solution might be to ship 20 truckloads from cannery 1 to warehouse 2 and the remaining 55 to warehouse 4, 80 from cannery 2 to warehouse 1 and 45 to warehouse 2 and, finally, 70 truckloads from cannery 3 to warehouse 3 and 30 to warehouse 4. Even though the linear programming formulation of the pea shipping example has seven constraints other than nonnegativity, a basic solution has only six basic variables! This is because the constraints are linearly dependent: the sum of the first three is identical to the sum of the last four. As a consequence, the feasible region defined by the constraints would remain the same if we only kept six of them. In general, a basic solution to the transportation model will have a number of basic variables equal to the number of sources plus the number of destinations minus one.

Exercise 79 Formulate the following production problem as a transportation model. The demands for a given item are 150, 250, 200 units for the next three months. The demand may be satisfied by

- excess production in an earlier month held in stock for later consumption,
- production in the current month,
- excess production in a later month backordered for preceding months.

The variable production cost per unit in any month is \$6.00. A unit produced for later consumption will incur a storage cost at the rate of \$1 per unit per month. On the other hand, backordered items incur a penalty cost of \$3.00 per unit per month. The production capacity in each of the next three months is 200 units. The objective is to devise a minimum cost production plan.

Exercise 80 Avertz RentaCar needs to redeploy its automobiles to correct imbalances in the system. Currently Avertz has too many cars in New York (with 10 cars excess) and Chicago (12 cars excess). Pittsburgh would like up to 6 cars, Los Angeles up to 14 cars, and Miami up to 7 cars (note that more cars are demanded than are available). The cost of transporting a car from one city to another is given by:

	Pittsburgh	Los Angeles	Miami
New York	50	250	100
Chicago	25	200	125

(a) Formulate this problem as a transportation problem. Clearly give the nodes and arcs. For each node, give the supply or demand at the node. For each arc, give the cost. Assume that unmet demand at a city has no cost but that no city can end up with excess cars.

(b) It turns out that unmet demand costs \$50/car in Pittsburgh, \$75/car in LA, and \$100/car in Miami. Update your transportation formulation in (a) to account for this change.

11.3.4 Assignment Problem

A special case of the transportation problem is the *assignment problem* which occurs when each supply is 1 and each demand is 1. In this case, the integrality implies that every supplier will be assigned one destination and every destination will have one supplier. The costs give the charge for assigning a supplier and destination to each other.

Example 11.3.1 *A company has three new machines of different types. There are four different plants that could receive the machines. Each plant can only receive one machine, and each machine can only be given to one plant. The expected profit that can be made by each plant if assigned a machine is as follows:*

	Plant			
	1	*2*	*3*	*4*
1	*13*	*16*	*12*	*11*
Machine 2	*15*	*0*	*13*	*20*
3	*5*	*7*	*10*	*6*

This is a transportation problem with all supplies and demands equal to 1, so it is an assignment problem.

Note that a balanced problem must have the same number of supplies and demands, so we must add a dummy machine (corresponding to receiving no machine) and assign a zero cost for assigning the dummy machine to a plant.

Exercise 81 Albert, Bob, Carl, and David are stranded on a desert island with Elaine, Francine, Gert, and Holly. The "compatibility measures" in the next table indicate the happiness each couple would experience if they spent all their time together. If a couple spends only a partial amount of time together, then the happiness is proportional to the fraction of time spent. So if Albert and Elaine spend half their time together, they earn happiness of 7/2.

	E	F	G	H
A	7	5	8	2
B	7	8	9	4
C	3	5	7	9
D	5	5	6	7

(a) Let x_{ij} be the fraction of time that man i spends with woman j. Formulate a linear program that maximizes the total happiness of the island (assume that no man spends time with any other man, no woman spends time with any other woman, and no one spends time alone).

(b) Explain why exactly four x_{ij} will be 1 and the rest will be 0 in an optimal solution to this linear program. This result is called the *Marriage Theorem* for obvious reasons.

(c) Do you think that this theorem will hold if we allow men to spend time with men and women to spend time with women?

11.4 Unifying Model: Minimum Cost Network Flows

All of the above models are special types of network flow problems: they each have a specialized algorithm that can find solutions hundreds of times faster than plain linear programming.

They can all also be seen as examples of a much broader model, the minimum cost network flow model. This model represents the broadest class of problem that can be solved much faster than linear programming while still retaining such nice properties as integrality of solution and appeal of concept.

Like the maximum flow problem, it considers flows in networks with capacities. Like the shortest path problem, it considers a cost for flow through an arc. Like the transportation problem, it allows multiple sources and destinations. In fact, all of these problems can be seen as *special cases* of the minimum cost flow problem.

Consider a directed network with n nodes. The decision variables are x_{ij}, the flow through arc (i, j). The given information includes:

- c_{ij}: cost per unit of flow from i to j (may be negative),
- u_{ij}: capacity (or upper bound) on flow from i to j,
- b_i: net flow generated at i.

This last value has a sign convention:

- $b_i > 0$ if i is a supply node,
- $b_i < 0$ if i is a demand node,
- $b_i = 0$ if i is a transshipment node.

The objective is to minimize the total cost of sending the supply through the network to satisfy the demand.

Note that for this model, it is not necessary that every arc exists. We will use the convention that summations are only taken over arcs that exist. The linear programming formulation for this problem is:

$$\begin{array}{ll} \text{Minimize} & \sum_i \sum_j c_{ij} x_{ij} \\ \text{Subject to} & \sum_j x_{ij} - \sum_j x_{ji} = b_i \text{ for all nodes } i, \\ & 0 \leq x_{ij} \leq u_{ij} \text{ for all arcs } (i, j). \end{array}$$

Again, we will assume that the network is balanced, so $\sum_i b_i = 0$, since dummies can be added as needed. We also still have a nice integrality property. If all the b_i and u_{ij} are integral, then the resulting solution to the linear program is also integral.

Minimum cost network flows are solved by a variation of the simplex algorithm and can be solved more than 100 times faster than equivalently sized linear programs. From a modeling point of view, it is most important to know the sort of things that can and cannot be modeled in a single network flow problem:

Can do

1. Lower bounds on arcs. If a variable x_{ij} has a lower bound of l_{ij}, upper bound of u_{ij}, and cost of c_{ij}, change the problem as follows:
   - Replace the upper bound with $u_{ij} - l_{ij}$,
   - Replace the supply at i with $b_i - l_{ij}$,
   - Replace the supply at j with $b_i + l_{ij}$,

   Now you have a minimum cost flow problem. Add $c_{ij}l_{ij}$ to the objective after solving and l_{ij} to the flow on arc (i, j) to obtain a solution of the original problem.
2. Upper bounds on flow through a node. Replace the node i with nodes i' and i''. Create an arc from i' to i'' with the appropriate capacity, and cost 0. Replace every arc (j, i) with one from j to i' and every arc (i, j) with one from i'' to j. Lower bounds can also be handled this way.
3. Convex, piecewise linear costs on arc flows (for minimization). This is handled by introducing multiple arcs between the nodes, one for each portion of the piecewise linear function. The convexity will assure that costs are handled correctly in an optimal solution.

Can't do

1. Fixed cost to use a node.
2. Fixed cost to use an arc.
3. "Proportionality of flow." That is, if one unit enters node i, then you insist that .5 units go to node j and .5 to node k.
4. Gains and losses of flow along arcs, as in power distribution.

Note that although these cannot be done in a single network, it may be possible to use the solutions to multiple networks to give you an answer. For instance, if there is only one arc with a fixed cost, you can solve both with and without the arc to see if it is advantageous to pay the fixed cost.

Exercise 82 Here is an example of a problem that doesn't look like a network flow problem, but it really is:

A company must meet the following demands for cash at the beginning of each of the next six months:

Month	1	2	3	4	5	6
Needs	\$200	\$100	\$50	\$80	\$160	\$140

At the beginning of month 1, the company has \$150 in cash and \$200 worth of bond 1, \$100 worth of bond 2 and \$400 worth of bond 3. Of course, the company will have to sell some bonds to meet demands, but a penalty will be charged for any bonds sold before the end of month 6. The penalties for selling \$1 worth of each bond are shown in the table below.

		Month of sale					
		1	2	3	4	5	6
	1	\$0.07	\$0.06	\$0.06	\$0.04	\$0.03	\$0.03
Bond	2	\$0.17	\$0.17	\$0.17	\$0.11	\$0	\$0
	3	\$0.33	\$0.33	\$0.33	\$0.33	\$0.33	\$0

(a) Assuming that all bills must be paid on time, formulate a balanced transportation problem that can be used to minimize the cost of meeting the cash demands for the next six months.

(b) Assume that payment of bills can be made after they are due, but a penalty of \$0.02 per month is assessed for each dollar of cash demands that is postponed for one month. Assuming all bills must be paid by the end of month 6, develop a transshipment model that can be used to minimize the cost of paying the next six months' bills.

[*Hint:* Transshipment points are needed, in the form C_t = cash available at beginning of month t after bonds for month t have been sold, but before month t demand is met. Shipments into C_t occur from bond sales and C_{t-1}. Shipments out of C_t occur to C_{t+1} and demands for months $1, 2, \ldots, t$.]

Exercise 83 Oil R' Us has oil fields in San Diego and Los Angeles. The San Diego field can produce up to 500,000 barrels per day (bpd); Los Angeles can produce up to 400,000 bpd. Oil is sent to a refinery, either in Dallas or Houston. It costs \$700 to refine 100,000 bpd in Dallas and \$900 to refine 100,000 bpd in Houston. The refined oil is then shipped to either Chicago or New York. Chicago requires exactly 400,000 bpd and New York requires exactly 300,000 bpd. The shipping costs (per 100,000 bpd) are given as follows:

	Dallas	Houston	New York	Chicago
LA	\$300	\$110	—	—
San Diego	\$420	\$100	—	—
Dallas	—	—	\$450	\$550
Houston	—	—	\$470	\$530

(a) Formulate this problem as a minimum cost flow problem. Clearly give the network. Give the cost and capacity on each arc and the net supply requirement on each node.

(b) Suppose Dallas could process no more than 300,000 bpd. How would you modify your formulation?

Exercise 84 A company produces a single product at two plants, A and B. A customer demands 12 units of the product this month, and 20 units of the product next month. Plant A can produce 7 items per month. It costs \$12 to produce each item at A this month and \$18 to produce at A next month. Plant B can produce 14 items per month. It costs \$14 to produce each item at B this month and \$20 to produce each item at B next month. Production from this month can be held in inventory to satisfy next month's demand at a cost of \$3/unit.

(a) Formulate the problem of determining the least cost way of meeting demand as a balanced transportation problem. Give the costs and node supplies, and describe the interpretation of each node and arc.

(b) Suppose no more than six items can be held in inventory between this month and next month. Formulate the resulting problem as a minimum cost flow problem. Draw the network, give the interpretation for each node and arc, and give the supply/demand at each node and the cost and capacity for each arc.

11.5 Generalized Networks

The final model I would like to familiarize you with is called the generalized network model. In this model, there may be gains or loses as flow goes along an arc. Each arc has a *multiplier* to represent these gains and loses. For instance, if 3 units of flow enter an arc with a multiplier of .5, then 1.5 unit of flow exit the arc. Such a model can still be represented as a linear program, and there are specialized techniques that can solve such models much faster than linear programming (but a bit slower than regular network flow problems). The optimal solution need not be integer however.

The standard example of a generalized network is in power generation: as electricty moves over wires, there is some unavoidable loss along the way. This loss is represented as a mutliplier. Here is another example on how generalized networks might be used:

Consider the world's currency market. Given two currencies, say the Yen and the USDollar, there is an exchange rate between them (currently about 110 Yen to the Dollar). It is axiomatic of a arbitrage-free market that there is no method of converting, say, a Dollar to Yen then to Deutsch Marks, to Francs, then Pounds, and to Dollars so that you end up with more than a dollar. How would you recognize when there is an arbitrage possibility?

For this model, we use a node to represent each currency. An arc between currency i and j gives the exchange rate from i to j. We begin with a single dollar and wish to send flow through the network and maximize the amount we generate back at the USDollar node. This can be drawn as in figure 11.5.

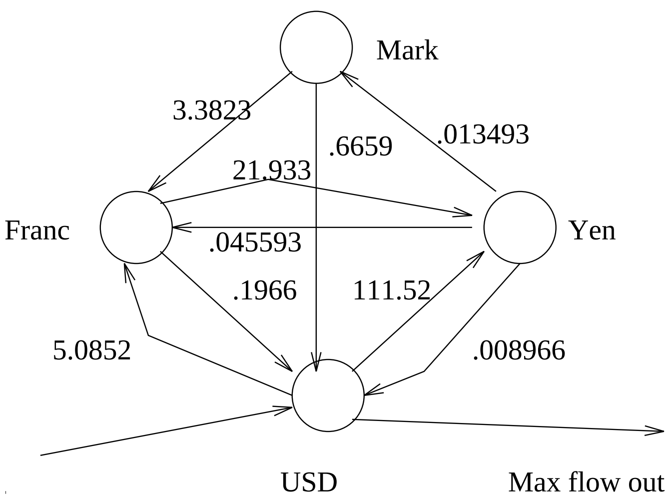


Figure 11.5: Financial Exchange Network

These are actual trades made on November 10, 1996. It is not obvious, but the Dollar-Yen-Mark-Dollar conversion actually makes \$0.002. How would you formulate a linear program to recognize this?

11.6 Conclusions

One class can only provide a bare introduction to this area. The points we would like you to take away are:

1) Networks are an important subclass of linear programs that are intuitive, easy to solve, and have nice integrality properties.

2) Problems that might not look like networks might be networks.

3) Networks provide a useful way to think about problems even if there are additional constraints or variables that preclude use of networks for modeling the whole problem.

In practice, you would generally save using the fastest network codes until the final implementation phase. Until then, linear programming codes will tend to be sufficiently fast to prove the concepts.

Chapter 12

Data Envelopment Analysis

Data Envelopment Analysis (DEA) is an increasingly popular management tool. This write-up is an introduction to Data Envelopment Analysis (DEA) for people unfamiliar with the technique. For a more in-depth discussion of DEA, the interested reader is referred to Seiford and Thrall [1990] or the seminal work by Charnes, Cooper, and Rhodes [1978].

DEA is commonly used to evaluate the efficiency of a number of producers. A typical statistical approach is characterized as a central tendency approach and it evaluates producers relative to an average producer In contrast, DEA compares each producer with only the "best" producers. By the way, in the DEA literature, a producer is usually referred to as a decision making unit or DMU. DEA is not always the right tool for a problem but is appropriate in certain cases. (See Strengths and Limitations of DEA.)

In DEA, there are a number of *producers*. The production process for each producer is to take a set of inputs and produce a set of outputs. Each producer has a varying level of inputs and gives a varying level of outputs. For instance, consider a set of banks. Each bank has a certain number of tellers, a certain square footage of space, and a certain number of managers (the inputs). There are a number of measures of the output of a bank, including number of checks cashed, number of loan applications processed, and so on (the outputs). DEA attempts to determine which of the banks are most efficient, and to point out specific inefficiencies of the other banks.

A fundamental assumption behind this method is that if a given producer, A, is capable of producing Y(A) units of output with X(A) inputs, then other producers should also be able to do the same if they were to operate efficiently. Similarly, if producer B is capable of producing Y(B) units of output with X(B) inputs, then other producers should also be capable of the same production schedule. Producers A, B, and others can then be combined to form a composite producer with composite inputs and composite outputs. Since this composite producer does not necessarily exist, it is typically called a virtual producer.

The heart of the analysis lies in finding the "best" virtual producer for each real producer. If the virtual producer is better than the original producer by either making more output with the same input or making the same output with less input then the original producer is inefficient. The subtleties of DEA are introduced in the various ways that producers A and B can be scaled up or down and combined.

12.1 Numerical Example

To illustrate how DEA works, let's take an example of three banks. Each bank has exactly 10 tellers (the only input), and we measure a bank based on two outputs: Checks cashed and Loan applications. The data for these banks is as follows:

- Bank A: 10 tellers, 1000 checks, 20 loan applications
- Bank B: 10 tellers, 400 checks, 50 loan applications
- Bank C: 10 tellers, 200 checks, 150 loan applications

Now, the key to DEA is to determine whether we can create a virtual bank that is better than one or more of the real banks. Any such dominated bank will be an inefficient bank.

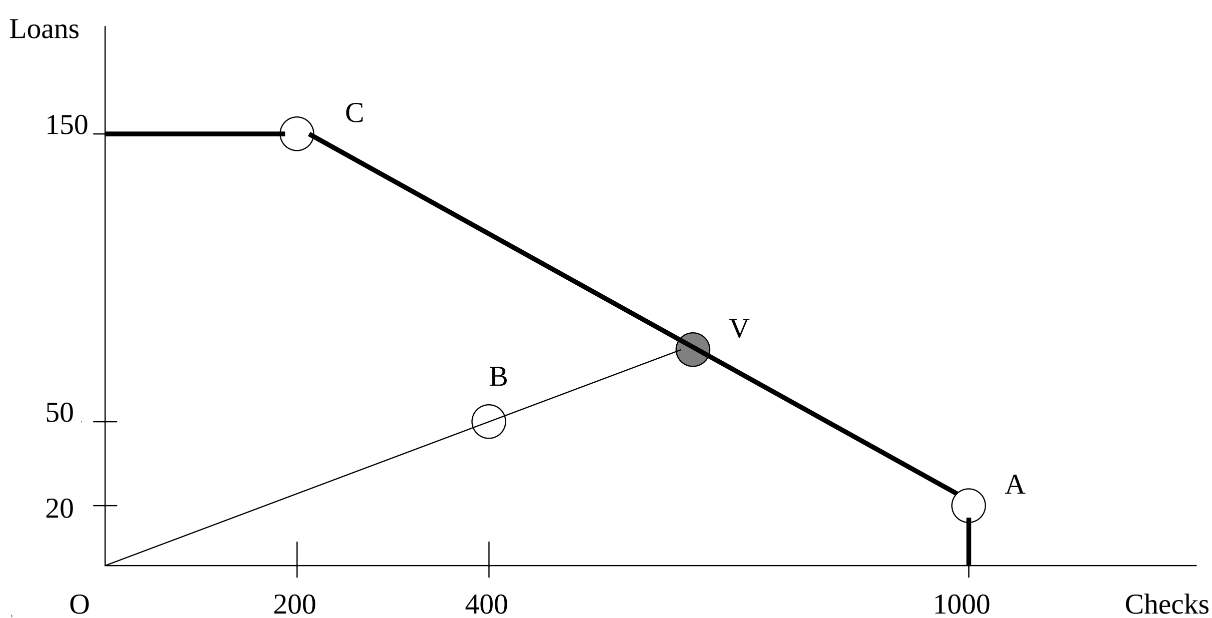
Consider trying to create a virtual bank that is better than Bank A. Such a bank would use no more inputs than A (10 tellers), and produce at least as much output (1000 checks and 20 loans). Clearly, no combination of banks B and C can possibly do that. Bank A is therefore deemed to be efficient. Bank C is in the same situation.

However, consider bank B. If we take half of Bank A and combine it with half of Bank C, then we create a bank that processes 600 checks and 85 loan applications with just 10 tellers. This dominates B (we would much rather have the virtual bank we created than bank B). Bank B is therefore inefficient.

Another way to see this is that we can scale down the inputs to B (the tellers) and still have at least as much output. If we assume (and we do), that inputs are linearly scalable, then we estimate that we can get by with 6.3 tellers. We do that by taking .34 times bank A plus .29 times bank B. The result uses 6.3 tellers and produces at least as much as bank B does. We say that bank B's efficiency rating is .63. Banks A and C have an efficiency rating of 1.

12.2 Graphical Example

The single input two-output or two input-one output problems are easy to analyze graphically. The previous numerical example is now solved graphically. (An assumption of constant returns to scale is made and explained in detail later.) The analysis of the efficiency for bank B looks like the following:



If it is assumed that convex combinations of banks are allowed, then the line segment connecting banks A and C shows the possibilities of virtual outputs that can be formed from these two banks. Similar segments can be drawn between A and B along with B and C. Since the segment AC lies beyond the segments AB and BC, this means that a convex combination of A and C will create the most outputs for a given set of inputs.

This line is called the efficiency frontier. The efficiency frontier defines the maximum combinations of outputs that can be produced for a given set of inputs.

Since bank B lies below the efficiency frontier, it is inefficient. Its efficiency can be determined by comparing it to a virtual bank formed from bank A and bank C. The virtual player, called V, is approximately 54% of bank A and 46% of bank C. (This can be determined by an application of the lever law. Pull out a ruler and measure the lengths of AV, CV, and AC. The percentage of bank C is then AV/AC and the percentage of bank A is CV/AC.)

The efficiency of bank B is then calculated by finding the fraction of inputs that bank V would need to produce as many outputs as bank B. This is easily calculated by looking at the line from the origin, O, to V. The efficiency of player B is OB/OV which is approximately 63%. This figure also shows that banks A and C are efficient since they lie on the efficiency frontier. In other words, any virtual bank formed for analyzing banks A and C will lie on banks A and C respectively. Therefore since the efficiency is calculated as the ratio of OA/OV or OA/OV, banks A and C will have efficiency scores equal to 1.0.

The graphical method is useful in this simple two dimensional example but gets much harder in higher dimensions. The normal method of evaluating the efficiency of bank B is by using an linear programming formulation of DEA.

Since this problem uses a constant input value of 10 for all of the banks, it avoids the complications caused by allowing different returns to scale. Returns to scale refers to increasing or decreasing efficiency based on size. For example, a manufacturer can achieve certain economies of scale by producing a thousand circuit boards at a time rather than one at a time - it might be only 100 times as hard as producing one at a time. This is an example of increasing returns to scale (IRS.)

On the other hand, the manufacturer might find it more than a trillion times as difficult to produce a trillion circuit boards at a time though because of storage problems and limits on the worldwide copper supply. This range of production illustrates decreasing returns to scale (DRS.) Combining the two extreme ranges would necessitate variable returns to scale (VRS.)

Constant Returns to Scale (CRS) means that the producers are able to linearly scale the inputs and outputs without increasing or decreasing efficiency. This is a significant assumption. The assumption of CRS may be valid over limited ranges but its use must be justified. As an aside, CRS tends to lower the efficiency scores while VRS tends to raise efficiency scores.

12.3 Using Linear Programming

Data Envelopment Analysis, is a linear programming procedure for a frontier analysis of inputs and outputs. DEA assigns a score of 1 to a unit only when comparisons with other relevant units do not provide evidence of inefficiency in the use of any input or output. DEA assigns an efficiency score less than one to (relatively) inefficient units. A score less than one means that a linear combination of other units from the sample could produce the same vector of outputs using a smaller vector of inputs. The score reflects the radial distance from the estimated production frontier to the DMU under consideration.

There are a number of equivalent formulations for DEA. The most direct formulation of the exposition I gave above is as follows:

Let X_i be the vector of inputs into DMU i. Let Y_i be the corresponding vector of outputs. Let X_0 be the inputs into a DMU for which we want to determine its efficiency and Y_0 be the outputs. So the X's and the Y's are the data. The measure of efficiency for DMU_0 is given by the following linear program:

$$\begin{array}{lrcl} Min & \theta & & \\ & & & \\ s.t. & \sum \lambda_i X_i & \leq & \theta X_0 \\ & \sum \lambda_i Y_i & \geq & Y_0 \\ & \lambda \geq 0 & & \end{array}$$

where λ_i is the weight given to DMU i in its efforts to dominate DMU 0 and θ is the efficiency of DMU 0. So the λ's and θ are the variables. Since DMU 0 appears on the left hand side of the equations as well, the optimal θ cannot possibly be more than 1. When we solve this linear program, we get a number of things:

1. The efficiency of DMU 0 (θ), with $\theta = 1$ meaning that the unit is efficient.
2. The unit's "comparables" (those DMU with nonzero λ).
3. The "goal" inputs (the difference between X_0 and $\sum \lambda_i X_i$)
4. Alternatively, we can keep inputs fixed and get goal outputs ($\frac{1}{\theta} \sum_i Y_i$)

DEA assumes that the inputs and outputs have been correctly identified. Usually, as the number of inputs and outputs increase, more DMUs tend to get an efficiency rating of 1 as they become too specialized to be evaluated with respect to other units. On the other hand, if there are too few inputs and outputs, more DMUs tend to be comparable. In any study, it is important to focus on correctly specifying inputs and outputs.

Example 12.3.1 *Consider analyzing the efficiencies of 3 DMUs where 2 inputs and 3 outputs are used. The data is as follows:*

DMU	*Inputs*		*Outputs*		
1	*5*	*14*	*9*	*4*	*16*
2	*8*	*15*	*5*	*7*	*10*
3	*7*	*12*	*4*	*9*	*13*

The linear programs for evaluating the 3 DMUs are given by:

- **LP for evaluating DMU 1:**

```
min THETA
st
5L1+8L2+7L3 - 5THETA <= 0
14L1+15L2+12L3 - 14THETA <= 0
9L1+5L2+4L3 >= 9
4L1+7L2+9L3 >= 4
16L1+10L2+13L3 >= 16
L1, L2, L3 >= 0
```

- **LP for evaluating DMU 2:**

```
min THETA
st
5L1+8L2+7L3 - 8THETA <= 0
14L1+15L2+12L3 - 15THETA <= 0
9L1+5L2+4L3 >= 5
4L1+7L2+9L3 >= 7
16L1+10L2+13L3 >= 10
L1, L2, L3 >= 0
```

- **LP for evaluating DMU 3:**

```
min THETA
st
5L1+8L2+7L3 - 7THETA <= 0
14L1+15L2+12L3 - 12THETA <= 0
9L1+5L2+4L3 >= 4
4L1+7L2+9L3 >= 9
16L1+10L2+13L3 >= 13
L1, L2, L3 >= 0
```

The solution to each of these is as follows:

- **DMU 1.**

Adjustable Cells

Cell	Name	Final Value	Reduced Cost	Objective Coefficien	Allowable Increase	Allowable Decrease
$B$10	theta	1	0	1	1E+30	1
$B$11	L1	1	0	0	0.92857142	0.619047619
$B$12	L2	0	0.24285714	0	1E+30	0.242857143
$B$13	L3	0	0	0	0.36710963	0.412698413

Constraints

Cell	Name	Final Value	Shadow Price	Constraint R.H. Side	Allowable Increase	Allowable Decrease
$B$16	IN1	-0.103473	0	0	1E+30	0
$B$17	IN2	-0.289724	-0.07142857	0	0	1E+30
$B$18	OUT1	9	0.085714286	9	0	0
$B$19	OUT2	4	0.057142857	4	0	0
$B$20	OUT3	16	0	16	0	1E+30

- **DMU 2.**

Adjustable Cells

Cell	Name	Final Value	Reduced Cost	Objective Coefficien	Allowable Increase	Allowable Decrease
$B$10	theta	0.773333333	0	1	1E+30	1
$B$11	L1	0.261538462	0	0	0.866666667	0.577777778
$B$12	L2	0	0.22666667	0	1E+30	0.226666667
$B$13	L3	0.661538462	0	0	0.342635659	0.385185185

Constraints

Cell	Name	Final Value	Shadow Price	Constraint R.H. Side	Allowable Increase	Allowable Decrease
$B$16	IN1	-0.24820512	0	0	1E+30	0.248205128
$B$17	IN2	-0.452651	-0.0666667	0	0.46538461	1E+30
$B$18	OUT1	5	0.08	5	10.75	0.655826558
$B$19	OUT2	7	0.0533333	7	1.05676855	3.41509434
$B$20	OUT3	12.78461538	0	10	2.78461538	1E+30

- **DMU 3.**

Adjustable Cells

Cell	Name	Final Value	Reduced Cost	Objective Coefficien	Allowable Increase	Allowable Decrease
$B$10	theta	1	0	1	1E+30	1
$B$11	L1	0	0	0	1.08333333	0.722222222
$B$12	L2	0	0.283333333	0	1E+30	0.283333333
$B$13	L3	1	0	0	0.42829457	0.481481481

Constraints

Cell	Name	Final Value	Shadow Price	Constraint R.H. Side	Allowable Increase	Allowable Decrease
$B$16	IN1	-0.559375	0	0	1E+30	0
$B$17	IN2	-0.741096	-0.08333333	0	0	1E+30
$B$18	OUT1	4	0.1	4	16.25	0
$B$19	OUT2	9	0.066666667	9	0	0
$B$20	OUT3	13	0	13	0	1E+30

Note that DMUs 1 and 3 are overall efficient and DMU 2 is inefficient with an efficiency rating of 0.773333.

Hence the efficient levels of inputs and outputs for DMU 2 are given by:

- Efficient levels of Inputs:

$$0.261538 \begin{bmatrix} 5 \\ 14 \end{bmatrix} + 0.661538 \begin{bmatrix} 7 \\ 12 \end{bmatrix} = \begin{bmatrix} 5.938 \\ 11.6 \end{bmatrix}$$

- Efficient levels of Outputs:

$$0.261538 \begin{bmatrix} 9 \\ 4 \\ 16 \end{bmatrix} + 0.661538 \begin{bmatrix} 4 \\ 9 \\ 13 \end{bmatrix} = \begin{bmatrix} 5 \\ 7 \\ 12.785 \end{bmatrix}$$

Note that the outputs are at least as much as the outputs currently produced by DMU 2 and inputs are at most as big as the 0.773333 times the inputs of DMU 2. This can be used in two different ways: The inefficient DMU should target to cut down inputs to equal at most the efficient levels. Alternatively, an equivalent statement can be made by finding a set of efficient levels of inputs and outputs by dividing the levels obtained by the efficiency of DMU 2. This focus can then be used to set targets primarily for outputs rather than reduction of inputs.

Alternate Formulation

There is another, probably more common formulation, that provides the same information. We can think of DEA as providing a price on each of the inputs and a value for each of the outputs. The efficiency of a DMU is simply the ratio of the inputs to the outputs, and is constrained to be no more than 1. The prices and values have nothing to do with real prices and values: they are an artificial construct. The goal is to find a set of prices and values that puts the target DMU in the best possible light. The goal, then is to

$$\begin{array}{lll} Max & \frac{u^T Y_0}{v^T X_0} & \\ s.t. & \frac{u^T Y_j}{v^T X_j} \leq & 1, \ \ j = 0, ..., n, \\ & u^T \geq & 0, \\ & v^T \geq & 0. \end{array}$$

Here, the variables are the u's and the v's. They are vectors of prices and values respectively.

This fractional program can be equivalently stated as the following linear programming problem (where Y and X are matrices with columns Y_j and X_j respectively).

$$\begin{array}{llll} Max & u^T Y_0 & & \\ s.t. & v^T X_0 & = & 1, \\ & u^T Y - v^T X & \leq & 0, \\ & u^T & \geq & 0, \\ & v^T & \geq & 0. \end{array}$$

We denote this linear program by (D). Let us compare it with the one introduced earlier, which we denote by (P):

$$\begin{array}{llll} Min & \theta & & \\ s.t. & \sum \lambda_i X_i & \leq & \theta X_0 \\ & \sum \lambda_i Y_i & \geq & Y_0 \\ & \lambda \geq 0. & & \end{array}$$

To fix ideas, let us write out explicitely these two formulations for DMU 2, say, in our example.

Formulation (P) for DMU 2:

```
min                      THETA
st
-5 L1 - 8 L2 - 7 L3 + 8 THETA >= 0
-14L1 -15 L2 -12 L3 + 15THETA >= 0
 9 L1 + 5 L2 + 4 L3           >= 5
 4 L1 + 7 L2 + 9 L3           >= 7
 16L1 +10 L2 +13 L3           >= 10
L1>=0, L2>=0, L3>=0
```

Formulation (D) for DMU 2:

```
max             5 U1 + 7 U2 + 10 U3
st
- 5 V1 - 14V2 + 9 U1 + 4 U2 + 16 U3 <= 0
- 8 V1 - 15V2 + 5 U1 + 7 U2 + 12 U3 <= 0
- 7 V1 - 12V2 + 4 U1 + 9 U2 + 13 U3 <= 0
  8 V1 + 15V2                        = 1
    V1>=0, V2>=0, U1>=0, U2>=0,  U3>=0
```

Formulations (P) and (D) are dual linear programs! These two formulations actually give the same information. You can read the solution to one from the shadow prices of the other. We will not discuss linear programming duality in this course. You can learn about it in some of the OR electives.

Exercise 85 Consider the following baseball players:

Name	At Bat	Hits	HomeRuns
Martin	135	41	6
Polcovich	82	25	1
Johnson	187	40	4

(You need know nothing about baseball for this question). In order to determine the efficiency of each of the players, At Bats is defined as an input while Hits and Home Runs are outputs. Consider the following linear program and its solution:

```
 MIN     THETA
 SUBJECT TO
          - 187 THETA + 135 L1 + 82 L2 + 187 L3 <=   0
           41 L1 + 25 L2 + 40 L3 >=   40
            6 L1 + L2 + 4 L3 >=   4
              L1, L2, L3 >= 0
```

Adjustable Cells

Cell	Name	Final Value	Reduced Cost	Objective Coefficient
$B$10	THETA	0.703135	0	1
$B$11	L1	0.550459	0	0
$B$12	L2	0.697248	0	0
$B$13	L3	0	0.296865	0

Constraints

Cell	Name	Final Value	Shadow Price	Constraint R.H. Side
$B$16	AT BATS	0	-0.005348	0
$B$17	HITS	0	0.017515	40
$B$18	HOME RUNS	0	0.000638	4

(a) For which player is this a DEA analysis? Is this player efficient? What is the efficiency rating of this player? Give the "virtual producer" that proves that efficiency rating (you should give the At bats, Hits, and Home Runs for this virtual producer).

(b) Formulate the linear program for Jonhson using the alternate formulation and solve using Solver. Compare the "Final Value" and "Shadow Price" columns from your Solver output with the solution given above.

12.4 Applications

The simple bank example described earlier may not convey the full view on the usefulness of DEA. It is most useful when a comparison is sought against "best practices" where the analyst doesn't want the frequency of poorly run operations to affect the analysis. DEA has been applied in many situations such as: health care (hospitals, doctors), education (schools, universities), banks, manufacturing, benchmarking, management evaluation, fast food restaurants, and retail stores.

The analyzed data sets vary in size. Some analysts work on problems with as few as 15 or 20 DMUs while others are tackling problems with over 10,000 DMUs.

12.5 Strengths and Limitations of DEA

As the earlier list of applications suggests, DEA can be a powerful tool when used wisely. A few of the characteristics that make it powerful are:

- DEA can handle multiple input and multiple output models.
- It doesn't require an assumption of a functional form relating inputs to outputs.
- DMUs are directly compared against a peer or combination of peers.
- Inputs and outputs can have very different units. For example, X1 could be in units of lives saved and X2 could be in units of dollars without requiring an a priori tradeoff between the two.

The same characteristics that make DEA a powerful tool can also create problems. An analyst should keep these limitations in mind when choosing whether or not to use DEA.

- Since DEA is an extreme point technique, noise (even symmetrical noise with zero mean) such as measurement error can cause significant problems.
- DEA is good at estimating "relative" efficiency of a DMU but it converges very slowly to "absolute" efficiency. In other words, it can tell you how well you are doing compared to your peers but not compared to a "theoretical maximum."
- Since DEA is a nonparametric technique, statistical hypothesis tests are difficult and are the focus of ongoing research.
- Since a standard formulation of DEA creates a separate linear program for each DMU, large problems can be computationally intensive.

12.6 References

DEA has become a popular subject since it was first described in 1978. There have been hundreds of papers and technical reports published along with a few books. Technical articles about DEA have been published in a wide variety of places making it hard to find a good starting point. Here are a few suggestions as to starting points in the literature.

1. Charnes, A., W.W. Cooper, and E. Rhodes. "Measuring the efficiency of decision making units." European Journal of Operations Research (1978): 429-44.
2. Banker, R.D., A. Charnes, and W.W. Cooper. "Some models for estimating technical and scale inefficiencies in data envelopment analysis." Management Science 30 (1984): 1078-92.
3. Dyson, R.G. and E. Thanassoulis. "Reducing weight flexibility in data envelopment analysis." Journal of the Operational Research Society 39 (1988): 563-76.
4. Seiford, L.M. and R.M. Thrall. "Recent developments in DEA: the mathematical programming approach to frontier analysis." Journal of Econometrics 46 (1990): 7-38.
5. Ali, A.I., W.D. Cook, and L.M. Seiford. "Strict vs. weak ordinal relations for multipliers in data envelopment analysis." Management Science 37 (1991): 733-8.
6. Andersen, P. and N.C. Petersen. "A procedure for ranking efficient units in data envelopment analysis." Management Science 39 (1993): 1261-4.
7. Banker, R.D. "Maximum likelihood, consistency and data envelopment analysis: a statistical foundation." Management Science 39 (1993): 1265-73.

The first paper was the original paper describing DEA and results in the abbreviation CCR for the basic constant returns-to-scale model. The Seiford and Thrall paper is a good overview of the literature. The other papers all introduce important new concepts. This list of references is certainly incomplete.

A good source covering the field of productivity analysis is The Measurement of Productive Efficiency edited by Fried, Lovell, and Schmidt, 1993, from Oxford University Press. There is also a recent book from Kluwer Publishers, Data Envelopment Analysis: Theory, Methodology, and Applications by Charnes, Cooper, Lewin, and Seiford.

To stay more current on the topics, some of the most important DEA articles appear in Management Science, The Journal of Productivity Analysis, The Journal of the Operational Research Society, and The European Journal of Operational Research. The latter just published a special issue, "Productivity Analysis: Parametric and Non-Parametric Approaches" edited by Lewin and Lovell which has several important papers.