Elements of Mathematics *for* Economics and Finance

VASSILIS C. MAVRON • TIMOTHY N. PHILLIPS

Springer

Elements of Mathematics for Economics and Finance

Vassilis C. Mavron and Timothy N. Phillips

Elements of Mathematics for Economics and Finance

With 77 Figures

Springer

Vassilis C. Mavron, MA, MSc, PhD
Institute of Mathematical and Physical Sciences
University of Wales Aberystwyth
Aberystwyth SY23 3BZ
Wales, UK

Timothy N. Phillips, MA, MSc, DPhil, DSc
Cardiff School of Mathematics
Cardiff University
Senghennydd Road
Cardiff CF24 4AG
Wales, UK

Mathematics Subject Classification (2000): 91-01; 91B02

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Control Number: 2006928729

ISBN-10: 1-84628-560-7 e-ISBN 1-84628-561-5 Printed on acid-free paper
ISBN-13: 978-1-84628-560-8

Printed in the United States of America (HAM)

9 8 7 6 5 4 3 2 1

Springer Science + Business Media, LLC
springer.com

Preface

The mathematics contained in this book for students of economics and finance has, for many years, been given by the authors in two single-semester courses at the University of Wales Aberystwyth. These were mathematics courses in an economics setting, given by mathematicians based in the Department of Mathematics for students in the Faculty of Social Sciences or School of Management. The choice of subject matter and arrangement of material reflect this collaboration and are a result of the experience thus obtained.

The majority of students to whom these courses were given were studying for degrees in economics or business administration and had not acquired any mathematical knowledge beyond pre-calculus mathematics, i.e., elementary algebra. Therefore, the first-semester course assumed little more than basic pre-calculus mathematics and was based on Chapters 1–7. This course led on to the more advanced second-semester course, which was also suitable for students who had already covered basic calculus. The second course contained at most one of the three Chapters 10, 12, and 13. In any particular year, their inclusion or exclusion would depend on the requirements of the economics or business studies degree syllabuses. An appendix on differentials has been included as an optional addition to an advanced course.

The students taking these courses were chiefly interested in learning the mathematics that had applications to economics and were not primarily interested in theoretical aspects of the subject per se. The authors have not attempted to write an undergraduate text in economics but instead have written a text in mathematics to complement those in economics.

The simplicity of a mathematical theory is sometimes lost or obfuscated by a dense covering of applications at too early a stage. For this reason, the aim of the authors has been to present the mathematics in its simplest form, highlighting threads of common mathematical theory in the various topics of economics.

Some knowledge of theory is necessary if correct use is to be made of the techniques; therefore, the authors have endeavoured to introduce some basic theory in the expectation and hope that this will improve understanding and incite a desire for a more thorough knowledge.

Students who master the simpler cases of a theory will find it easier to go on to the more difficult cases when required. They will also be in a better position to understand and be in control of calculations done by hand or calculator and also to be able to visualise problems graphically or geometrically. It is still true that the best way to understand a technique thoroughly is through practice. Mathematical techniques are no exception, and for this reason the book illustrates theory through many examples and exercises.

We are grateful to Noreen Davies and Joe Hill for invaluable help in preparing the manuscript of this book for publication.

Above all, we are grateful to our wives, Nesta and Gill, and to our children, Nicholas and Christiana, and Rebecca, Christopher, and Emily, for their patience, support, and understanding: this book is dedicated to them.

Vassilis C. Mavron
Aberystwyth
United Kingdom

Timothy N. Phillips
Cardiff
United Kingdom

March 2006

Contents

1 Essential Skills

1.1 Introduction

Many models and problems in modern economics and finance can be expressed using the language of mathematics and analysed using mathematical techniques. This book introduces, explains, and applies the basic quantitative methods that form an essential foundation for many undergraduate courses in economics and finance. The aim throughout this book is to show how a range of important mathematical techniques work and how they can be used to explore and understand the structure of economic models.

In this introductory chapter, the reader is reacquainted with some of the basic principles of arithmetic and algebra that formed part of their previous mathematical education. Since economics and finance are quantitative subjects it is vitally important that students gain a familiarity with these principles and are confident in applying them. Mathematics is a subject that can only be learnt by doing examples, and therefore students are urged to work through the examples in this chapter to ensure that these key skills are understood and mastered.

1.2 Numbers

For most, if not all, of us, our earliest encounter with numbers was when we were taught to count as children using the so-called counting numbers $1, 2, 3, 4, \ldots$. The counting numbers are collectively known as the **natural numbers**. The natural numbers can be represented by equally spaced points on a line as shown in Fig. 1.1. The direction in which the arrow is pointing in Fig. 1.1 indicates the direction in which the numbers are getting larger, i.e., the natural numbers are ordered in the sense that if you move along the line to the right, the numbers progressively increase in magnitude.

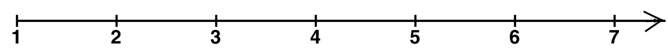

Figure 1.1 The natural numbers.

It is sometimes useful and necessary to talk in terms of numbers less than zero. For example, a person with an overdraft on their bank account essentially has a negative balance or debt, which needs to be cancelled before the account is in credit again. In the physical world, negative numbers are used to report temperatures below 0^0 Centigrade, which is the temperature at which water freezes. So, for example, -5^0C is 5^0 C below freezing.

If the line in Fig. 1.1 is extended to the left, we can mark equally spaced points that represent zero and the negatives of the natural numbers. The natural numbers, their negatives, and the number zero are collectively known as the **integers**. All these numbers can be represented by equally spaced points on a number line as shown in Fig. 1.2. If we move along the line to the right, the numbers become progressively larger, while if we move along the line to the left, the numbers become smaller. So, for example, -4 is smaller than -1 and we write $-4 < -1$ where the symbol '$<$' means 'is less than' or, equivalently, -1 is greater than -4 and we write $-1 > -4$ where the symbol '$>$' means 'is greater than'. Note that these symbols should not be confused with the symbols '$\leq$' and '$\geq$', which mean 'less than or equal to' and 'greater than or equal to', respectively.

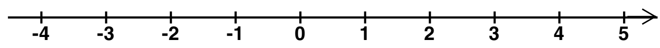

Figure 1.2 Integers on the number line.

1.2.1 Addition and Subtraction

Initially, numerical operations involving negative numbers may seem rather confusing. We give the rules for adding and subtracting numbers and then appeal to the number line for some justification. If a and b are any two numbers, then we have the following rules

$$a + (-b) = a - b, \tag{1.1}$$

$$a - (+b) = a - b, \tag{1.2}$$

$$a - (-b) = a + b. \tag{1.3}$$

Thus we can regard $-(-b)$ as equal to $+b$.

We consider a few examples:

$$4 + (-1) = 4 - 1 = 3,$$

and

$$3 - (-2) = 3 + 2 = 5.$$

The last example makes sense if we regard $3 - (-2)$ as the difference between 3 and -2 on the number line. Note that $a - b$ will be negative if and only if $a < b$. For example,

$$-2 - (-1) = -2 + 1 = -1 < 0.$$

1.2.2 Multiplication and Division

If a and b are any two positive numbers, then we have the following rules for multiplying positive and negative numbers:

$$a \times (-b) = -(a \times b), \tag{1.4}$$

$$(-a) \times b = -(a \times b), \tag{1.5}$$

$$(-a) \times (-b) = a \times b. \tag{1.6}$$

So multiplication of two numbers of the same sign gives a positive number, while multiplication of two numbers of different signs gives a negative number.

For example, to calculate $2 \times (-5)$, we multiply 2 by 5 and then place a minus sign before the answer. Thus,

$$2 \times (-5) = -10.$$

It is usual in mathematics to write ab rather than $a \times b$ to express the multiplication of two numbers a and b. We say that ab is the product of a and b. Thus, we can write (1.6) in the form

$$(-a)(-b) = ab.$$

These multiplication rules give, for example,

$$(-2) \times (-3) = 6,\ (-4) \times 5 = -20,\ 7 \times (-5) = -35.$$

The same rules hold for division because it is the same sort of operation as multiplication, since

$$\frac{a}{b} = a \times \frac{1}{b}.$$

So the division of a number by another of the same sign gives a positive number, while division of a number by another of the opposite sign gives a negative number. For example, we have

$$(-15) \div (-3) = 5,\ (-16) \div 2 = -8,\ 2 \div (-4) = -1/2.$$

1.2.3 Evaluation of Arithmetical Expressions

The order in which operations in an arithmetical expression are performed is important. Consider the calculation

$$12 + 8 \div 4.$$

Different answers are obtained depending on the order in which the operations are executed. If we first add together 12 and 8 and then divide by 4, the result is 5. However, if we first divide 8 by 4 to give 2 and then add this to 12, the result is 14. Therefore, the order in which the mathematical operations are performed is important and the convention is as follows: brackets, exponents, division, multiplication, addition, and subtraction. So that the evaluation of expressions within brackets takes precedence over addition and the evaluation of any number or expressions raised to a power (an exponential) takes precedence over division, for example. This convention has the acronym BEDMAS. However, the main point to remember is that if you want a calculation to be done in a particular order, you should use brackets to avoid any ambiguity.

Example 1.1

Evaluate the expression $2^3 \times 3 + (5 - 1)$.

Solution. Following the BEDMAS convention, we evaluate the contents of the bracket first and then evaluate the exponential. Therefore,

$$\begin{aligned} 2^3 \times 3 + (5 - 1) &= 2^3 \times 3 + 4 \\ &= 8 \times 3 + 4. \end{aligned}$$

Finally, since multiplication takes precedence over addition, we have

$$2^3 \times 3 + (5 - 1) = 24 + 4 = 28.$$

1.3 Fractions

A **fraction** is a number that expresses part of a whole. It takes the form a/b where a and b are any integers except that $b \neq 0$. The integers a and b are known as the **numerator** and **denominator** of the fraction, respectively. Note that a can be greater than b. The formal name for a fraction is a **rational number** since they are formed from the ratio of two numbers. Examples of statements that use fractions are 3/5 of students in a lecture may be female or 1/3 of a person's income may be taxed by the government.

Fractions may be simplified to obtain what is known as a **reduced fraction** or a **fraction in its lowest terms**. This is achieved by identifying any common factors in the numerator and denominator and then cancelling those factors by dividing both numerator and denominator by them. For example, consider the simplification of the fraction 27/45. Both the numerator and denominator have 9 as a common factor since $27 = 9 \times 3$ and $45 = 9 \times 5$ and therefore it can be cancelled:

$$\frac{27}{45} = \frac{3 \times 9}{5 \times 9} = \frac{3}{5}.$$

We say that 27/45 and 3/5 are **equivalent fractions** and that 3/5 is a reduced fraction.

To compare the relative sizes of two fractions and also to add or subtract two fractions, we express them in terms of a common denominator. The common denominator is a number that each of the denominators of the respective fractions divides, i.e., each is a **factor** of the common denominator. Suppose we wish to determine which is the greater of the two fractions 4/9 and 5/11. The common denominator is $9 \times 11 = 99$. Each of the denominators (9 and

11) of the two fractions divides 99. The simplest way to compare the relative sizes is to multiply the numerator and denominator of each fraction by the denominator of the other, i.e.,

$$\frac{4}{9} = \frac{4 \times 11}{9 \times 11} = \frac{44}{99}, \text{ and } \frac{5}{11} = \frac{5 \times 9}{11 \times 9} = \frac{45}{99}.$$

So $5/11 > 4/9$ since $45/99 > 44/99$.

We follow a similar procedure when we want to add two fractions. Consider the general case first of all in which we add the fractions a/b and c/d with $b \neq 0$ and $d \neq 0$:

$$\begin{aligned} \frac{a}{b} + \frac{c}{d} &= \frac{a \times d}{b \times d} + \frac{c \times b}{d \times b} \\ &= \frac{a \times d + b \times c}{b \times d}. \end{aligned}$$

Therefore, we have

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}. \tag{1.7}$$

For example,

$$\frac{2}{7} + \frac{3}{5} = \frac{2 \times 5 + 3 \times 7}{7 \times 5} = \frac{10 + 21}{35} = \frac{31}{35}.$$

The result for the subtraction of two fractions is similar, i.e.,

$$\frac{a}{b} - \frac{c}{d} = \frac{ad - bc}{bd}. \tag{1.8}$$

Example 1.2

Simplify

$$\frac{13}{24} - \frac{5}{16}.$$

Solution. The idea is to express each of these fractions as equivalent fractions having a common denominator. Therefore, we have

$$\begin{aligned} \frac{13}{24} - \frac{5}{16} &= \frac{13 \times 16}{24 \times 16} - \frac{5 \times 24}{16 \times 24} \\ &= \frac{208 - 120}{384} \\ &= \frac{88}{384} \\ &= \frac{11 \times 8}{48 \times 8} \\ &= \frac{11}{48}. \end{aligned}$$

Note that a smaller common denominator, namely 48, could have been used in this example since the two denominators, viz. 16 and 24, are both factors of 48. Thus

$$\frac{13}{24} = \frac{2 \times 13}{2 \times 24} = \frac{26}{48}$$

and

$$\frac{5}{16} = \frac{3 \times 5}{3 \times 16} = \frac{13}{48}.$$

Therefore,

$$\frac{13}{24} - \frac{5}{16} = \frac{26 - 15}{48} = \frac{11}{48}.$$

1.3.1 Multiplication and Division

To multiply together two fractions, we simply multiply the numerators together and multiply the denominators together:

$$\frac{a}{b} \times \frac{c}{d} = \frac{a \times c}{b \times d} = \frac{ac}{bd}. \tag{1.9}$$

To divide one fraction by another, we multiply by the **reciprocal** of the divisor where the reciprocal of the fraction a/b is defined to be b/a provided $a, b \neq 0$. That is

$$\frac{a}{b} \div \frac{c}{d} = \frac{a}{b} \times \frac{d}{c} = \frac{a \times d}{b \times c} = \frac{ad}{bc}. \tag{1.10}$$

Example 1.3

Simplify the following fractions

1. $\frac{5}{8} \times \frac{16}{27}$,
2. $\frac{9}{13} \div \frac{12}{25}$.

Solution.

1. The product is the fraction

$$\frac{5 \times 16}{8 \times 27}.$$

To simplify this fraction, we note that 8 is a factor of the numerator and denominator (since $16 = 8 \times 2$) and can be cancelled. Therefore, we have

$$\frac{5}{8} \times \frac{16}{27} = \frac{5 \times 16}{8 \times 27} = \frac{5 \times 8 \times 2}{8 \times 27} = \frac{10}{27}.$$

2. Using the rule (1.10) for the division of two fractions, we have

$$\frac{9}{13} \div \frac{12}{25} = \frac{9}{13} \times \frac{25}{12} = \frac{9 \times 25}{13 \times 12}.$$

Then noting that 3 is a common factor of the numerator and denominator, we have

$$\frac{5}{8} \times \frac{16}{27} = \frac{3 \times 3 \times 25}{13 \times 4 \times 3} = \frac{3 \times 25}{13 \times 4} = \frac{75}{52}.$$

1.4 Decimal Representation of Numbers

A fraction or rational number may be converted to its equivalent decimal representation by dividing the numerator by the denominator. For example, the decimal representation of 3/4 is found by dividing 3 by 4 to give 0.75. This is an example of a **terminating decimal** since it ends after a finite number of digits. The following are examples of rational numbers that have a terminating decimal representation:

$$\frac{1}{8} = 0.125,$$

and

$$\frac{3}{25} = 0.12.$$

Some fractions do not possess a finite decimal representation – they go on forever. The fraction 1/3 is one such example. Its decimal representation is 0.3333... where the dots denote that the 3s are repeated and we write

$$\frac{1}{3} = 0.\dot{3},$$

where the dot over the number indicates that it is repeated indefinitely. This is an example of a **recurring decimal**. All rational numbers have a decimal representation that either terminates or contains an infinitely repeated finite sequence of numbers. Another example of a recurring decimal is the decimal representation of 1/13:

$$\frac{1}{13} = 0.0769230769230\ldots = 0.0\dot{7}6923\dot{0},$$

where the dots indicate the first and last digits in the repeated sequence.

All numbers that do not have a terminating or recurring decimal representation are known as **irrational numbers**. Examples of irrational numbers are $\sqrt{2}$ and π. All the irrational numbers together with all the rational numbers form the **real numbers**. Every point on the number line in Fig. 1.2 corresponds to a real number, and the line is known as the **real line**.

To convert a decimal to a fraction, you simply have to remember that the first digit after the decimal point is a tenth, the second a hundredth, and so on. For example,

$$0.2 = \frac{2}{10} = \frac{1}{5},$$

and

$$0.375 = \frac{375}{1{,}000} = \frac{3}{8}.$$

Sometimes we are asked to express a number correct to a certain number of **decimal places** or a certain number of **significant figures**. Suppose that we wish to write the number 23.541638 correct to two decimal places. To do this, we truncate the part of the number following the second digit after the decimal point:

$$23.54 \mid 1638.$$

Then, since the first neglected digit, 1 in this case, lies between 0 and 4, then the truncated number, 23.54, is the required answer. If we wish to write the same number correct to three decimal places, the truncated number is

$$23.541 \mid 638,$$

and since the first neglected digit, 6 in this case, lies between 5 and 9, then the last digit in the truncated number is rounded up by 1. Therefore, the number 23.541638 is 23.542 correct to three decimal places or, for short, 'to three decimal places'.

To express a number to a certain number of significant figures, we employ the same rounding strategy used to express numbers to a certain number of decimal places but we start counting from the first non-zero digit rather than from the first digit after the decimal point. For example,

$$\begin{aligned} 72{,}648 &= 70{,}000 \text{ (correct to 1 significant figure)} \\ &= 73{,}000 \text{ (correct to 2 significant figures)} \\ &= 72{,}600 \text{ (correct to 3 significant figures)} \\ &= 72{,}650 \text{ (correct to 4 significant figures)}, \end{aligned}$$

and

$$\begin{aligned} 0.004286 &= 0.004 \text{ (correct to 1 significant figure)} \\ &= 0.0043 \text{ (correct to 2 significant figures)} \\ &= 0.00429 \text{ (correct to 3 significant figures)}. \end{aligned}$$

Note that $497 = 500$ correct to 1 significant figure and also correct to 2 significant figures.

1.4.1 Standard Form

The distance of the Earth from the Sun is approximately 149,500,000 km. The mass of an electron is 0.000000000000000000000000000911 g. Numbers such as these are displayed on a calculator in **standard** or **scientific form**. This is a shorthand means of expressing very large or very small numbers. The standard form of a number expresses it in terms of a number lying between 1 and 10 multiplied by 10 raised to some power or exponent. More precisely, the standard form of a number is

$$a \times 10^b,$$

where $1 \leq a < 10$, and b is an integer. A practical reason for the use of the standard form is that it allows calculators and computers to display more significant figures than would otherwise be possible.

For example, the standard form of 0.000713 is 7.13×10^{-4} and the standard form of 459.32 is 4.5932×10^2. The power gives the number of decimal places the decimal point needs to be moved to the right in the case of a positive power or the number of decimal places the decimal point needs to be moved to the left in the case of a negative power. For example, $5.914 \times 10^3 = 5914$ and $6.23 \times 10^{-4} = 0.000623$. Returning to the above examples, the Earth is about 1.495×10^8 km from the Sun and the mass of an electron is 9.11×10^{-28} g. Similarly, a budget deficit of 257,000,000,000 is 2.57×10^{11} in standard form.

1.5 Percentages

To convert a fraction to a percentage, we multiply the fraction by 100%. For example,

$$\frac{3}{4} = \frac{3}{4} \times 100\% = 75\%,$$

and

$$\frac{3}{13} = \frac{3}{13} \times 100\% = 23.077\% \text{ (to three decimal places).}$$

To perform the reverse operation and convert a percentage to a fraction, we divide the number by 100. The resulting fraction may then be simplified to obtain a reduced fraction. For example,

$$45\% = \frac{45}{100} = \frac{9}{20},$$

where the fraction has been simplified by dividing the numerator and denominator by 5 since this is a common factor of 45 and 100.

To find the percentage of a quantity, we multiply the quantity by the number and divide by 100. For example,

$$25\% \text{ of } 140 \text{ is } \frac{25}{100} \times 140 = 35,$$

and

$$4\% \text{ of } 5,200 \text{ is } \frac{4}{100} \times 5,200 = 208.$$

If a quantity is increased by a percentage, then that percentage of the quantity is added to the original. Suppose that an investment of £300 increases in value by 20%. In monetary terms, the investment increases by

$$\frac{20}{100} \times 300 = £60,$$

and the new value of the investment is

$$£300 + £60 = £360.$$

In general, if the percentage increase is $r\%$, then the new value of the investment comprises the original and the increase. The new value can be found by multiplying the original value by the factor

$$1 + \frac{r}{100}.$$

It is easy to work in the reverse direction and determine the original value if the new value and percentage increase is known. In this case, one simply divides by the factor

$$1 + \frac{r}{100}.$$

Example 1.4

The cost of a refrigerator is £350.15 including sales tax at 17.5%. What is the price of the refrigerator without sales tax?

Solution. To determine the price of the refrigerator without sales tax, we divide £350.15 by the factor

$$1 + \frac{17.5}{100} = 1.175.$$

So the price of the refrigerator without VAT is

$$\frac{350.15}{1.175} = £298.$$

Similarly, if a quantity decreases by a certain percentage, then that percentage of the original quantity is subtracted from the original to obtain its new value. The new value may be determined by multiplying the original value by the quantity

$$1 - \frac{r}{100}.$$

Example 1.5

A person's income is €25,000 of which €20,000 is taxable. If the rate of income tax is 22%, calculate the person's net income.

Solution. The person's net income comprises the part of his salary that is not taxable (€5,000) together with the portion of his taxable income that remains after the tax has been taken. The person's net income is therefore

$$\begin{aligned} 5{,}000 + \left(1 - \frac{22}{100}\right) \times 20{,}000 &= 5{,}000 + \frac{78}{100} \times 20{,}000 \\ &= 5{,}000 + 78 \times 200 \\ &= 5{,}000 + 15{,}600 \\ &= €20{,}600. \end{aligned}$$

1.6 Powers and Indices

Let x be a number and n be a positive integer, then x^n denotes x multiplied by itself n times. Here x is known as the **base** and n is the **power** or **index** or **exponent**. For example,

$$x^5 = x \times x \times x \times x \times x.$$

There are rules for multiplying and dividing two algebraic expressions or numerical values involving the same base raised to a power. In the case of multiplication, we add the indices and raise the expression or value to that new power to obtain the product rule

$$x^a \times x^b = x^a x^b = x^{a+b}.$$

For example,

$$x^2 \times x^3 = (x \times x) \times (x \times x \times x) = x^5.$$

In the case of division, we subtract the indices and raise the expression or value to that new power to obtain the quotient rule

$$x^a \div x^b = \frac{x^a}{x^b} = x^{a-b}.$$

For example,

$$x^2 \div x^4 = \frac{x \times x}{x \times x \times x \times x} = \frac{1}{x^2},$$

and using the quotient rule we have

$$\frac{x^2}{x^4} = x^{2-4} = x^{-2}.$$

More generally, we have

$$\frac{1}{x^n} = x^{-n}.$$

Suppose now that we multiply an expression with a fractional power as many times as the denominator of the fraction. For example, multiply $x^{1/3}$ by itself three times. We have

$$x^{1/3} \times x^{1/3} \times x^{1/3} = x^{1/3+1/3+1/3} = x^1 = x.$$

However, the number that when multiplied by itself three times gives x is known as the cube root of x, and an alternative notation for $x^{1/3}$ is $\sqrt[3]{x}$. The symbol $\sqrt[n]{x}$, which sometimes appears on a calculator as $x^{1/n}$, is known as the nth root of x. In the case $n = 2$, the n is omitted in the former symbol. So we write $\sqrt{x}$ rather than $\sqrt[2]{x}$ for the square root $x^{1/2}$ of x.

Suppose we wish to raise an expression with a power to a power, for example $(x^2)^4$. We may rewrite this as

$$(x^2)(x^2)(x^2)(x^2) = x^{2+2+2+2} = x^8,$$

using the product rule. More generally, we have

$$(x^m)^n = x^{mn}.$$

These rules for simplifying expressions involving powers may be used to evaluate arithmetic expressions without using a calculator. For example,

$$\begin{aligned} 2^3 &= 2 \times 2 \times 2 = 8, \\ 3^4 &= 3 \times 3 \times 3 \times 3, \\ \sqrt{81} &= 9, \\ \sqrt[3]{27} &= 3, \\ 2^{-3} &= \frac{1}{2^3} = \frac{1}{8}. \end{aligned}$$

Note the following two conventions related to the use of powers:

1. $x^1 = x$ (An exponent of 1 is not expressed.)
2. $x^0 = 1$ for $x \neq 0$ (Any nonzero number raised to the zero power is equal to 1.)

To summarise, we have the following rules governing indices or powers:

Rules of Indices

$$x^a x^b = x^{a+b} \tag{1.11}$$
$$\frac{x^a}{x^b} = x^{a-b} \tag{1.12}$$
$$(x^a)^b = x^{ab} \tag{1.13}$$
$$\frac{1}{x^a} = x^{-a} \tag{1.14}$$
$$\sqrt[a]{x} = x^{\frac{1}{a}} \tag{1.15}$$
$$\sqrt[a]{x^b} = x^{\frac{b}{a}} \tag{1.16}$$

Finally, consider the product of two numbers raised to some power. For example, consider $(xy)^3$. Now

$$(xy)^3 = (x \times y) \times (x \times y) \times (x \times y) = (x \times x \times x) \times (y \times y \times y) = x^3 y^3,$$

since it does not matter in which order numbers are multiplied. More generally, we have

$$(xy)^a = x^a y^a.$$

Similarly, we have

$$\left(\frac{x}{y}\right)^a = \frac{x^a}{y^a}.$$

Example 1.6

Simplify the following using the rules of indices:

1. $\dfrac{x^2}{x^{3/2}}$,
2. $\dfrac{x^2 y^3}{x^4 y}$.

Solution.

1. Using the quotient rule (1.12), we have

$$\frac{x^2}{x^{3/2}} = x^{2-3/2} = x^{1/2} = \sqrt{x}$$

2. Using the quotient and reciprocal rules, we have

$$\begin{aligned}\frac{x^2y^3}{x^4y} &= \left(\frac{x^2}{x^4}\right)\left(\frac{y^3}{y}\right)\\ &= (x^{2-4})(y^{3-1}) \quad \text{(using the quotient rule (1.12))}\\ &= x^{-2}y^2\\ &= \frac{y^2}{x^2} \qquad \text{(using the reciprocal rule (1.14))}\\ &= \left(\frac{y}{x}\right)^2.\end{aligned}$$

Example 1.7

Write down the values of the following without using a calculator:

1.	3^{-3}	2.	$16^{3/4}$	3.	$16^{-3/4}$
4.	$27^{-1/3}$	5.	$4^{3/2}$	6.	19^0.

Solution.

1. $3^{-3} = \frac{1}{3^3} = \frac{1}{27}$.
2. $16^{3/4} = (16^{1/4})^3 = (\sqrt[4]{16})^3 = 2^3 = 8$.
3. $16^{-3/4} = \frac{1}{16^{3/4}} = \frac{1}{8}$.
4. $27^{-1/3} = \frac{1}{27^{1/3}} = \frac{1}{\sqrt[3]{27}} = \frac{1}{3}$.
5. $4^{3/2} = (4^{1/2})^3 = (\sqrt{4})^3 = 2^3 = 8$.
6. $19^0 = 1$.

Note that we could also evaluate $4^{3/2}$ as follows:

$$4^{3/2} = (4^3)^{1/2} = 64^{1/2} = \sqrt{64} = 8.$$

1.7 Simplifying Algebraic Expressions

In the algebraic expression

$$7x^3,$$

x is called the **variable**, and 7 is known as the **coefficient** of x^3. Expressions consisting simply of a coefficient multiplying one or more variables raised to the power of a positive integer are called **monomials**. Monomials can be added or subtracted to form **polynomials**. Each of the monomials comprising a polynomial is called a **term**. For example, the terms in the polynomial $3x^2+2x+1$ are $3x^2$, $2x$, and 1. The coefficient of x^2 is 3, the coefficient of x is 2, and the constant term is 1.

To add or subtract two polynomials, we collect like terms and add or subtract their coefficients. For example, if we wish to add $7x+2$ and $5-2x$, then we collect the terms in x and the constant terms:

$$(7x+2)+(5-2x)=(7+(-2))x+(2+5)=5x+7.$$

Example 1.8

Simplify the following:

1. $(3x^2+2x+1)+(5x^2-x-7)$,
2. $(9x^4+12x^3+6x+1)-(x^4+2x^2-4)$,
3. $(x^3+4x-5)+(2x^2-x+8)$.

Solution.

1. $(3+5)x^2+(2-1)x+(1-7)=8x^2+x-6$.
2. $(9-1)x^4+12x^3-2x^2+6x+(1+4)=8x^4+12x^3-2x^2+6x+5$.
3. $x^3+2x^2+(4-1)x+(-5+8)=x^3+2x^2+3x+3$.

1.7.1 Multiplying Brackets

There are occasions when mathematical expressions may be simplified by removing any brackets present. This process, which is also known as expanding the brackets or multiplying out the brackets, culminates in an equivalent expression without brackets. The removal of brackets is based on the following basic rule:

$$a(b+c)=ab+ac, \tag{1.17}$$

where a, b, and c are any three numbers. Since the order in which multiplication is performed is not important, we also have

$$(b+c)a = ba + ca, \tag{1.18}$$

The rules (1.17) and (1.18), which are examples of what is known as the distributive law, may be generalized to include expressions involving polynomials. For example,

$$3(x+2y) = 3x + 6y,$$

and

$$-2(3x^2 - 5y) = -6x^2 + 10y.$$

It is important to take care multiplying out brackets when there is a negative sign outside the brackets. In this case, the sign of each term inside the brackets is changed when the brackets are removed. For example,

$$-(2x^2 - 3x - 2y + 5) = -2x^2 + 3x + 2y - 5.$$

We also have the following rule for multiplying two brackets:

$$(a+b)(c+d) = ac + bc + ad + bd, \tag{1.19}$$

where a, b, c, and d are any three numbers. So to multiply out two brackets we simply multiply each term in the second bracket by each term in the first bracket and add together all contributions. For example,

$$\begin{aligned} (x+2)(2x-3) &= (x)(2x) + (2)(2x) + (x)(-3) + (2)(-3) \\ &= 2x^2 + 4x - 3x - 6 \\ &= 2x^2 + x - 6. \end{aligned}$$

The rule (1.19) extends to brackets containing more than two terms. The important thing to remember is that each term in the second bracket is multiplied by each term in the first before all contributions are added together. For example,

$$\begin{aligned} (2x - y + 5)(x-3) &= (2x)(x) + (-y)(x) + (5)(x) \\ &\quad +(2x)(-3) + (-y)(-3) + (5)(-3) \\ &= 2x^2 - xy + 5x - 6x + 3y - 15 \\ &= 2x^2 - xy - x + 3y - 15. \end{aligned}$$

Example 1.9

Multiply out the brackets and simplify the following:

1. $(2x+3)(7-5x)$,
2. $\dfrac{(120 - 24x)}{4.8}$,
3. $(x+3y)(2x-5y-1)$.

Solution.

1. Using the rule (1.19), we have

$$\begin{aligned}(2x+3)(7-5x) &= (2x)(7)+(3)(7)+(2x)(-5x)+(3)(-5x)\\ &= 14x+21-10x^2-15x\\ &= 21-x-10x^2.\end{aligned}$$

2. In this example, we just note that division of $120-24x$ by 0.48 is the same as multiplication of $120-24x$ by $1/(4.8)$, and therefore we can use the rule (1.17):

$$\begin{aligned}\frac{(120-24x)}{4.8} &= \frac{1}{4.8}(120-24x)\\ &\quad \frac{120}{4.8}+\frac{-24x}{4.8}\\ &= 25-5x.\end{aligned}$$

3. Using the generalization of rule (1.19), we have

$$\begin{aligned}(x+3y)(2x-5y-1) &= (x)(2x)+(3y)(2x)+(x)(-5y)\\ &\quad +(3y)(-5y)+(x)(-1)+(3y)(-1)\\ &= 2x^2+6xy-5xy-15y^2-x-3y\\ &= 2x^2+xy-15y^2-x-3y.\end{aligned}$$

1.7.2 Factorization

Factorization is the reverse process to multiplying out the brackets. It involves taking a mathematical expression and rewriting it by expressing it in terms of a product of factors. There are a number of techniques that can be used to factorize an expression:

1. The simplest technique is to identify a **common factor** in two or more terms. The equivalent factorized expression can then be written in terms of the common factor multiplying a bracketed expression. For example,

   a) $ab-ac=a(b-c)$,

   b) $4x^2+6x=2x(2x+3)$,

   c) $ax^2-a^2x=ax(x-a)$,

   d) $-36x^2-9x=-9x(4x+1)$,

e) $\dfrac{5x+10y}{10x-5y} = \dfrac{5(x+2y)}{5(2x-y)} = \dfrac{x+2y}{2x-y}$.

2. The second technique is based on the following **identity** involving the **difference of two squares**:

$$a^2 - b^2 = (a-b)(a+b).$$

An identity is a formula valid for all values of the variables; in this case, a and b. The following are examples of the application of this identity:

a) $x^2 - 36 = (x-6)(x+6)$;

b) $9a^2 - 16x^2 = (3a)^2 - (4x)^2 = (3a-4x)(3a+4x)$;

c) $9 - 36x^2 = 9(1-4x^2) = 9(1^2 - (2x)^2) = 9(1-2x)(1+2x)$.

An additional technique that can be used for factorizing quadratic expressions of the form $ax^2 + bx + c$ or $ax^2 + bxy + cy^2$ will be discussed in Chapter 3.

EXERCISES

1.1. Evaluate

$$3^5 - 8 \div 2^2 + 5 + 2^3 \times 4.$$

1.2. Express the following fractions using decimal notation:

a) $\dfrac{3}{10}$,

b) $\dfrac{5}{16}$,

c) $\dfrac{3}{4}$,

d) $\dfrac{3}{13}$,

e) $\dfrac{2}{7}$,

f) $\dfrac{1}{19}$.

1.3. Simplify the following fractions:

a) $\dfrac{2}{5} + \dfrac{3}{8}$,

b) $\dfrac{5}{16} - \dfrac{3}{32}$,

c) $\frac{15}{54} \times \frac{18}{35}$,

d) $\frac{32}{49} \div \frac{56}{21}$.

1.4. Find which is the larger of the two fractions: 11/32, 7/24 by expressing the numbers as:

a) fractions with the same denominator;

b) decimals.

1.5. Write each of the following numbers correct to two decimal places:

a) 51.2361

b) 7.896

c) 362.275

1.6. Write each of the following numbers correct to three significant figures:

a) 5,889

b) 0.0002817

c) 72,961

d) 0.09274

1.7. Write each of the following numbers in standard form:

a) 495,200

b) 0.000000837

1.8. The computing equipment belonging to a company is valued at \$45,000. Each year, 12% of the value is written off for depreciation. Find the value of the equipment at the end of two years.

1.9. Death duties of 20% are paid on a legacy to three children of £180,000. The eldest child is bequeathed 50%, the middle child 30%, and the youngest child the remainder. How much does each child receive? What percentage of the original legacy does the youngest child receive?

1.10. Simplify the following:

a) $x^{2/3}x^{7/3}$,

b) $\frac{x^5}{x^2}$,

c) $(x^{2/3})^6$,

d) $\dfrac{x^3y^2}{x^2y^5}$.

1.11. Write down the values of the following without using a calculator:

a) $16^{5/4}$,

b) $81^{1/4}$,

c) $\left(\dfrac{27}{125}\right)^{2/3}$,

d) $81^{-3/4}$.

1.12. Multiply out the brackets and simplify the following:

a) $(2x+9)(3x-8)$,

b) $(x+4)(6x+3)$,

c) $(3x-2)(11-4x)$,

d) $\dfrac{(15-24x+18y)}{0.75}$,

e) $(x-4y+7)(5x-2y-3)$.

1.13. Factorize the following expressions:

a) $96x-32$,

b) $-21x+49x^2$,

c) $4x^2-49$.

2
Linear Equations

2.1 Introduction

In this book, we will be concerned primarily with the analysis of the relationship between two or more variables. For example, we will be interested in the relationship between economic entities or variables such as

– *total cost* and *output*,

– *price* and *quantity* in an analysis of demand and supply,

– *production* and *factors of production* such as *labour* and *capital*.

If one variable, say y, changes in an entirely predictable way in terms of another variable, say x, then, under certain conditions (to be defined precisely in Chapter 4), we say that y is a **function** of x. A function provides a rule for providing values of y given values of x. The simplest function that relates two or more variables is a linear function. In the case of two variables, the linear function takes the form of the linear equation $y = ax + b$ for $a \neq 0$. For example, $y = 3x + 5$ is an example of a linear function. Given a value of x, one can determine the corresponding value of y using this functional relationship. For instance, when $x = 2$, $y = 3 \times 2 + 5 = 11$ and when $x = -3$, $y = 3 \times (-3) + 5 = -4$. We will say more about functions in Chapter 4. Linear equations or functions may be portrayed by a straight line on a graph. In this chapter, we introduce graphs and give a number of examples showing how linear equations can be used to model situations in economics and how to interpret properties of their graphs.

2.2 Solution of Linear Equations

A mathematical statement setting two algebraic expressions equal to each other is called an **equation**. The ability to solve equations is one of the most important algebraic techniques to master. Equipped with this skill, you will be able to solve a range of economic problems. The simplest type of equation is the **linear equation** in a single variable or unknown, which we will denote by x for the moment. In a linear equation, the unknown x only occurs raised to the **power** 1. The following are examples of linear equations:

1. $5x + 3 = 11$,
2. $1 - 4x = 3x + 7$,
3. $\dfrac{2 + 3x}{5} = \dfrac{2x - 1}{6}$.

A linear equation may be solved by rearranging it so that all terms involving x appear on one side of the equation and all the constant terms appear on the other side. This is achieved by performing a series of algebraic operations. The key is to remember that you must perform the same operations to both sides of the equation. You must be completely impartial so that each stage of the rearrangement process yields an equivalent equation. Two equations are said to be **equivalent** if and only if when one holds then so does the other. Equivalent equations, therefore, have precisely the same solutions if they have any at all. However, it is important that you never multiply or divide through an equation by 0. For example, take the equation $1 = 2$, which is not valid, and multiply both sides by 0. Then we obtain the equation $0 = 0$, which is true. So the two equations are not equivalent. If an equation contains a fraction, then the equation may be simplified by multiplying through by the denominator. Remember that the value of a fraction a/b is the same if the numerator and denominator are multiplied (or divided) by the same nonzero number. That is,

$$\frac{a}{b} = \frac{ta}{tb},$$

for any number $t \neq 0$. It is instructive to look at an example.

Example 2.1

Solve the equation

$$\frac{7x - 4}{2} = 2x + 4.$$

Solution. To determine the value of x that satisfies this equation, we rearrange the equation so that all terms involving the unknown x appear on the one side of the equation and all the constant terms appear on the other.

1. Multiply both sides by 2, which is the denominator of the fraction on the left-hand side of this equation:

$$\begin{aligned} 7x - 4 &= 2 \times (2x + 4) \\ &= (2 \times 2)x + (2 \times 4) \\ &= 4x + 8. \end{aligned}$$

2. Subtract $4x$ from both sides so that all terms involving x are on the left-hand side:

$$7x - 4 - 4x = 4x + 8 - 4x,$$

$$3x - 4 = 8.$$

3. Add 4 to both sides so that all the constant terms are on the right-hand side:

$$3x - 4 + 4 = 8 + 4,$$

$$3x = 12.$$

4. Finally divide both sides by 3:

$$\frac{3x}{3} = \frac{12}{3},$$

$$x = 4.$$

So the solution to this equation is $x = 4$.

We can check to see if this answer is correct by replacing x by 4 in the original equation. If $x = 4$ is the correct solution, then the left- and right-hand sides of the equation should give the same numerical value.

$$\begin{aligned} \text{LHS} &= \frac{(7 \times 4) - 4}{2} \\ &= \frac{28 - 4}{2} \\ &= \frac{24}{2} \\ &= 12 \end{aligned}$$

$$\begin{aligned} \text{RHS} &= 2 \times 4 + 4 \\ &= 12. \end{aligned}$$

Example 2.2

Solve the equation

$$\frac{x}{4} - 3 = \frac{x}{5} + 1. \tag{2.1}$$

Solution. Again, we go through the solution step-by-step. The idea is to rearrange the equation so that all terms involving x appear on the left-hand side and all the constant terms appear on the right-hand side. Once this is done, the terms involving fractions are simplified.

1. Subtract $x/5$ from both sides:

$$\frac{x}{4} - \frac{x}{5} - 3 = 1$$

2. Add 3 to both sides

$$\frac{x}{4} - \frac{x}{5} = 1 + 3 = 4$$

3. Simplify the left-hand side by expressing it as a single fraction. This is achieved by expressing each of the fractions in terms of their lowest common denominator, 20. In the case of the first fraction, both the numerator and denominator are multiplied by 5, and in the case of the second fraction they are both multiplied by 4, i.e.,

$$\frac{x}{4} = \frac{5x}{5 \times 4} = \frac{5x}{20} \text{ and } \frac{x}{5} = \frac{4x}{4 \times 5} = \frac{4x}{20}.$$

Therefore

$$\begin{aligned} \frac{5x}{20} - \frac{4x}{20} &= 4 \\ \frac{5x - 4x}{20} &= 4 \\ \frac{x}{20} &= 4. \end{aligned}$$

4. Finally multiply both sides by 20:

$$x = 80.$$

The solution to this equation is $x = 80$. Again we can check that this is the correct solution by substituting $x = 80$ into the left- and right-hand sides of (2.1).

2.3 Solution of Simultaneous Linear Equations

A number of economic models are built on linear relationships between variables. For example, the economic concept of equilibrium requires the solution of a system of equations.

The next degree of difficulty is to solve two linear equations in two unknowns. Suppose the two unknowns are denoted by x and y. The most general form of system of simultaneous linear equations in the unknowns x and y is

$$a_1x + b_1y = c_1, \tag{2.2}$$
$$a_2x + b_2y = c_2. \tag{2.3}$$

where a_1, b_1, c_1, a_2, b_2, and c_2 are constants. In the first equation (2.2), the coefficient of x is a_1 and that of y is b_1. We are going to describe the **elimination method** for solving this system of equations. As its name suggests, the method involves eliminating one of the variables from the system. This allows us to determine the value of the unknown that remains by solving a single linear equation in one unknown. The value of the eliminated unknown is then determined by substituting the known value into either of the original equations and solving another linear equation.

Suppose we wish to eliminate the variable y from (2.2)–(2.3). To do this, we multiply (2.2) by b_2 and (2.3) by b_1 so that the coefficients of y in the equivalent equations are the same:

$$b_2a_1x + b_2b_1y = b_2c_1, \tag{2.4}$$
$$b_1a_2x + b_1b_2y = b_1c_2. \tag{2.5}$$

Next we eliminate the variable y by subtracting (2.5) from (2.4):

$$(b_2a_1 - b_1a_2)x = b_2c_1 - b_1c_2, \tag{2.6}$$

from which we deduce

$$x = \frac{b_2c_1 - b_1c_2}{b_2a_1 - b_1a_2}. \tag{2.7}$$

Note that we can only perform this last step provided that $(b_2a_1 - b_1a_2) \neq 0$. The quantity $(b_2a_1 - b_1a_2)$ is known as the **determinant** (see Chapter 10) of the system of equations (2.2)–(2.3). The condition for this system to possess a unique solution is that the determinant is nonzero.

Similarly, we can eliminate x from equations (2.2)–(2.3) to obtain

$$y = \frac{c_2a_1 - c_1a_2}{b_2a_1 - b_1a_2}; \tag{2.8}$$

or we can obtain y by substituting the value of x we have obtained (2.7) in either (2.2) or (2.3) and solving the resulting linear equation.

There is no guarantee that a system of two or more simultaneous equations will possess a unique solution. Consider the system of equations

$$\begin{aligned} 2x + y &= 10, \\ 2x + y &= 5. \end{aligned}$$

This system of equations does not have a solution. In fact, the equations are **inconsistent**. They cannot hold simultaneously since $10 \neq 5$! We shall see later in this chapter that the solution of a system of simultaneous linear equations may be interpreted as the point of intersection of two straight lines. For the example under consideration, the two lines are parallel and therefore never intersect.

Next consider the system of equations

$$\begin{aligned} 2x + y &= 10, \\ -6x - 3y &= -30. \end{aligned}$$

At first sight this might seem to be an innocuous system of equations. However, the second equation is just a multiple of the first; obtained by multiplying the first equation by -3. In this case, the equations are not **independent**. The second equation does not provide any additional information over the first equation. Since there are two unknowns to be determined, there is no unique solution – in fact there are infinitely many solutions. For the above system one can verify that $x = s$ and $y = 10 - 2s$ is a solution for any number s.

To obtain a unique solution to a system of simultaneous linear equations, the equations must be consistent and independent and there must be as many equations as unknowns (variables).

Example 2.3

Solve the system of equations

$$\begin{aligned} 3x + 2y &= 1 \\ -2x + y &= 2. \end{aligned}$$

Solution. We solve this system of equations using the elimination method in which we eliminate the variable x. To do this, we arrange for the coefficients of x in both equations to differ only in sign by multiplying the two equations by appropriate factors. The variable can then be eliminated by adding or subtracting the two equations. For example, suppose we multiply the first equation by 2 and the second by 3:

$$\begin{aligned} 6x + 4y &= 2 \\ -6x + 3y &= 6. \end{aligned}$$

The variable x is eliminated by adding the two equations:

$$7y = 8,$$

which, after division by 7, gives

$$y = \frac{8}{7}.$$

This value can now be substituted in either of the original two equations to obtain the corresponding value of x. Let us use the first equation, then

$$\begin{aligned} 3x + 2\left(\frac{8}{7}\right) &= 1 \\ 3x + \frac{16}{7} &= 1 \\ 3x &= 1 - \frac{16}{7} \\ 3x &= \frac{7-16}{7} \text{ (since } 1 = 7/7) \\ 3x &= -\frac{9}{7} \\ x &= \frac{1}{3} \times \left(-\frac{9}{7}\right) \\ x &= -\frac{3}{7} \end{aligned}$$

Therefore, the solution is $x = -3/7$, $y = 8/7$. Of course, we can check that we have the correct solution by substituting it back into the original set of equations and checking that the equations are satisfied.

An alternative but equivalent method for solving simultaneous linear equations is known as the **substitution method**. The idea is to rearrange one of the equations in order to isolate one of the variables on the left-hand side. The expression for this variable is then substituted into the second equation to yield a linear equation for the other variable. We demonstrate this by means of an example.

Example 2.4

At the beginning of the year, an investor had £50,000 in two bank accounts, each of which paid interest annually. The interest rates were 4% and 6% per annum, respectively. If the investor has made no withdrawals during the year and has earned a total of £2,750 interest, what was the initial balance in each of the two accounts?

Solution. Let x and y denote the initial balances in the accounts with interest rates 4% and 6%, respectively. Since the total amount invested at the start of the year was £50,000, we have

$$x + y = 50{,}000.$$

The amount of interest earned on the two bank accounts during the year is given by

$$0.04x \text{ and } 0.06y,$$

respectively. Since the total amount of interest earned during the year is £2,750,

$$0.04x + 0.06y = 2{,}750,$$

or, after multiplying through by 100

$$4x + 6y = 275{,}000.$$

Therefore, we have two equations with which to determine initial balances in the two bank accounts:

$$x + y = 50{,}000 \tag{2.9}$$
$$4x + 6y = 275{,}000. \tag{2.10}$$

Multiplying (2.9) by 4, we obtain

$$4x + 4y = 200{,}000. \tag{2.11}$$

Then subtracting (2.11) from (2.10) yields

$$2y = 75{,}000,$$

so that $y = 37{,}500$. Finally, it follows from (2.9) that $x = 12{,}500$. Therefore, the initial balance in each of the two accounts was £12,500 and £37,500, respectively.

2.4 Graphs of Linear Equations

Consider the linear equation

$$y = 3x - 2.$$

Given a value of x, one can use this equation to determine the corresponding value of y. For example, when $x = 0$, $y = 3 \times 0 - 2 = -2$, and when $x = 2$, $y = 3 \times 2 - 2 = 6 - 2 = 4$. The collection of all such pairs of values of x and y that satisfy this linear equation can be represented on a **graph**.

Consider the two perpendicular lines shown in Fig. 2.1. The horizontal line is referred to as the x-axis and the vertical line as the y-axis. The point where these lines intersect is known as the origin and is denoted by the letter O. At this point, both variables take the value zero. Each axis is assigned a numerical scale that is chosen appropriately for the situation being considered. On the x-axis, the scale takes positive values to the right of the origin and negative values to the left. Moreover, the further we move away from the origin, the larger these values become. On the y-axis, the scale takes positive values above the origin and negative values below. Again, the further we move away from the origin in the vertical direction, the larger these values become. These axes enable us to define uniquely any point, P, in terms of its coordinates, (x, y). We write the coordinates (x, y) alongside the point P as in Fig. 2.1. The first number, x, denotes the horizontal distance along the x-axis and the second number y denotes the vertical distance along the y-axis. The arrows on the axis denote the positive direction. The collection of all points (x, y) satisfying a linear equation lie on a **straight line**. That is, any equation of the form

$$y = ax + b, \tag{2.12}$$

where a and b are constants is a linear equation and can be represented by a straight line graph. We sometimes say that y is a linear function of x since in the equation defining y, the variable x only occurs linearly.

Note also that the equation $x = k$, where k is any constant, is also represented by a straight line graph: the ‘vertical’ line, parallel to the y-axis, through the point $(k, 0)$.

Example 2.5

Plot the following points $A : (-2, 3)$, $B : (-3, -4)$, $C : (3, 5)$, $D : (1, -4)$.

Solution. The position of A is determined by the pair of values $x = -2$ and $y = 3$, and therefore it is located 2 units in the negative x-direction and 3 units in the positive y-direction as shown in Fig. 2.2. The other points are plotted in a similar way.

The general form of a linear equation is

$$cx + dy = e, \tag{2.13}$$

where c, d, and e are constants. We assume that c and d are not both zero. This equation contains multiples of x and y and a constant. These are the only terms involving x that are present in a linear equation; otherwise the equation

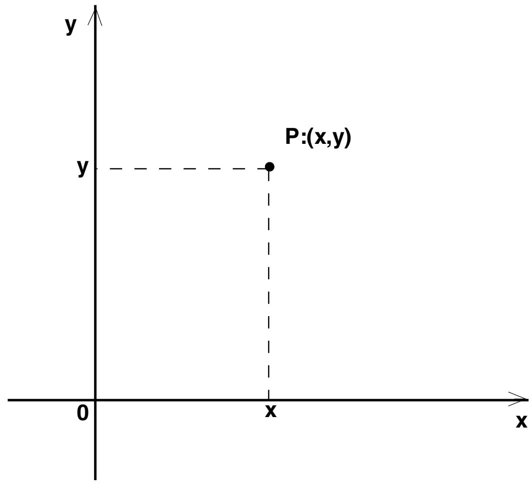

Figure 2.1 The coordinate axes and the position of a general point P.

is said to be **nonlinear**. The values c and d are referred to as the coefficients of x and y, respectively,. For example, the coefficients of the linear equation

$$2x - y = -3$$

are 2 and -1. More specifically, the coefficient of x is 2 and the coefficient of y is -1.

Any equation of the form (2.13) can be rearranged into the form (2.12) provided $d \neq 0$. First subtract cx from both sides of (2.13):

$$dy = -cx + e.$$

Then divide both sides by d provided $d \neq 0$:

$$y = -\frac{c}{d}x + \frac{e}{d}. \qquad (2.14)$$

If we now compare this equation with (2.12) by comparing the coefficients of x and the constant terms in both equations, we see that (2.14) is just (2.12) with

$$a = -\frac{c}{d}, \quad b = \frac{e}{d}.$$

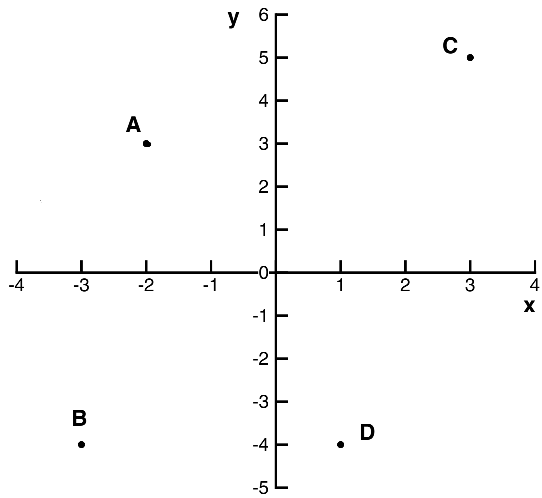

Figure 2.2 The location of the points specified in Example 2.5.

Note that when $d = 0$, the linear equation (2.13) reduces to

$$cx = e \text{ or } x = \frac{e}{c}.$$

This is represented by a straight line parallel to the y-axis passing through the point $(e/c, 0)$ on the x-axis.

To sketch the graph of a straight line, it is sufficient to draw a line through any two points lying on it.

Example 2.6

Sketch the graph of the straight line

$$y = 2x + 3,$$

for values of x lying between 0 and 4.

Solution. We determine the coordinates of two points on the line. When $x = 0$, we have that $y = 3$ and when $x = 4$, we have $y = 11$. Therefore, the points $(0, 3)$ and $(4, 11)$ lie on the line. The graph is formed by drawing a straight line through these points as shown in Fig. 2.3.

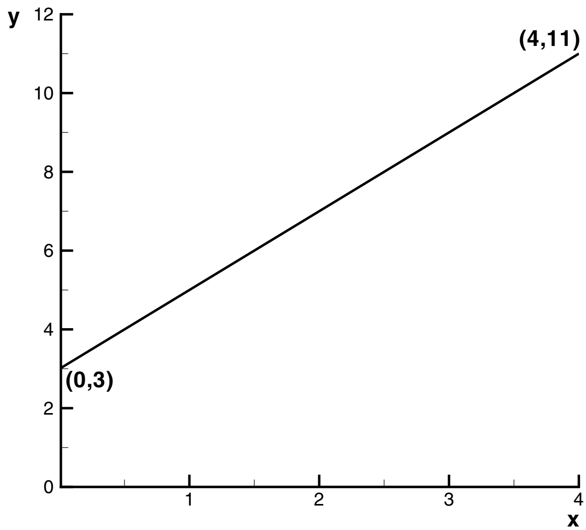

Figure 2.3 The graph of the equation $y = 2x + 3$.

Example 2.7

Sketch the straight line

$$2x + y = 5.$$

Solution. Setting $x = 0$ gives $y = 5$. Hence $(0, 5)$ lies on the line. Setting $y = 0$ gives $2x = 5$ or $x = 5/2$. Hence $(5/2, 0)$ lies on the line.

2.4.1 Slope of a Straight Line

The coefficients a and b in the linear equation $y = ax + b$ of (2.12) have special significance and can be related to features of its graph. When $x = 0$, $y = b$ and therefore the constant b represents the **intercept** on the y-axis, i.e., it is the value of y corresponding to the point of intersection of the straight line with the y-axis. The value of x for which $y = 0$ is the solution of the linear equation

$$ax + b = 0.$$

This equation has solution $x = -b/a$, provided $a \neq 0$.

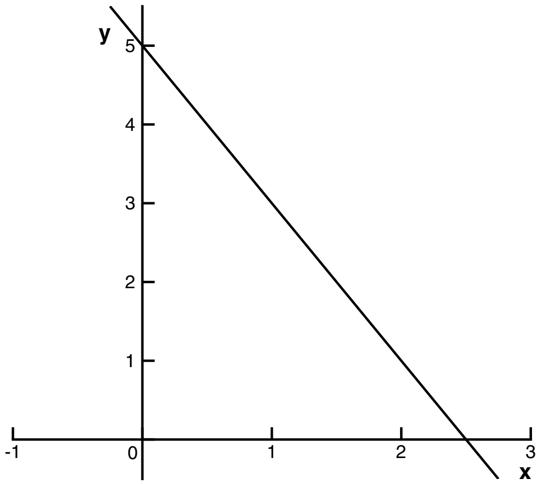

Figure 2.4 The graph of the equation $2x + y = 5$.

The coefficient a in the equation $y = ax + b$ defines the **slope** or **gradient** of the straight line with that equation. The slope of a straight line provides important information about the behaviour of the relationship between the variables x and y. Let $A : (x_1, y_1)$ and $B : (x_2, y_2)$ be any two distinct points lying on a straight line as shown in Fig. 2.5. The slope or gradient of the line measures the ratio of the change in the vertical direction with respect to the change in the horizontal direction as one moves from A to B. We illustrate this with reference to Fig. 2.5. Since $y_1 = ax_1 + b$ and $y_2 = ax_2 + b$, then

$$y_2 - y_1 = ax_2 - ax_1 = a(x_2 - x_1).$$

Therefore,

$$\frac{BC}{AC} = \frac{y_2 - y_1}{x_2 - x_1} = \frac{a(x_2 - x_1)}{x_2 - x_1} = a,$$

i.e.

$$a = \frac{y_2 - y_1}{x_2 - x_1} = \frac{BC}{AC}. \tag{2.15}$$

The value of a is independent of the choice of points A, B on the line. Positive values of a correspond to straight lines where y increases as x increases, while negative values of a correspond to straight lines where y decreases as x increases. Larger values of a correspond to straight lines with steeper slopes. For example,

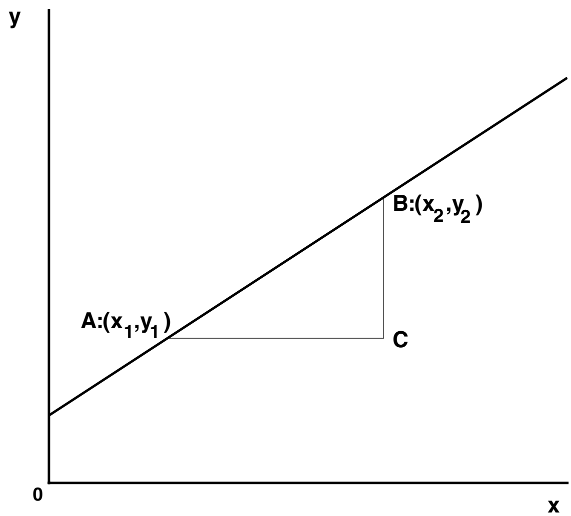

Figure 2.5 The graph of a linear equation and its slope.

the slope of the straight line $y = 6x - 3$ is steeper than that of $y = x + 3$. Another way of viewing the slope a is that it is the change in y when x increases by one unit, as then $x_2 - x_1 = 1$ and therefore $a = y_2 - y_1$.

Example 2.8

Determine the slope and intercept of the straight line $9x + 3y = 4$.

Solution. We need to write this equation in the form $y = ax + b$.

$$\begin{aligned} 9x + 3y &= 4 \\ 3y &= -9x + 4 \\ y &= -3x + \frac{4}{3} \end{aligned}$$

One can say immediately that the slope of this straight line is -3 and the intercept is $4/3$.

Example 2.9

Find the slope of the straight line that passes through the points $(2, -1)$ and $(-2, -11)$.

Solution. The slope of a straight line passing through the points (x_1, y_1), (x_2, y_2) is

$$a = \frac{y_2 - y_1}{x_2 - x_1}.$$

Therefore the required slope is

$$a = \frac{-11 - (-1)}{-2 - 2} = \frac{-10}{-4} = \frac{5}{2}.$$

2.5 Budget Lines

Suppose that a company or an individual has a given budget, B, that can be used to purchase two goods. If the cost or price of each of these goods is known, then it is possible to determine the different combinations of the two goods that can be bought with the given budget. Suppose that the two goods are denoted by X and Y, and their respective prices are P_X and P_Y. The quantities purchased of these goods is also denoted by X and Y. Then the equation of the budget line is

$$P_X X + P_Y Y = B. \tag{2.16}$$

Example 2.10

An electrical company has a budget of £6,000 a week to spend on the manufacture of toasters and kettles. It costs £5 to manufacture a toaster and £12 to manufacture a kettle. Write down the equation of the budget line and sketch its graph.

Solution. Let T and K denote the number of toasters and kettles that are manufactured each week. Then the cost of manufacture and the available budget means that the budget line has the equation

$$5T + 12K = 6{,}000.$$

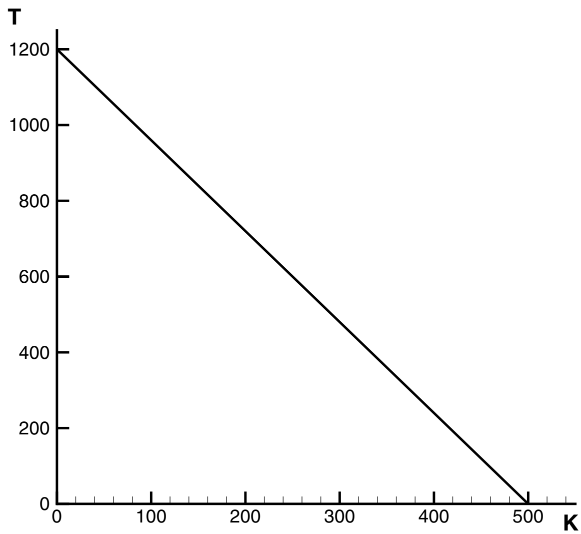

Figure 2.6 The graph of the budget line $5T + 12K = 6{,}000$.

To sketch the graph of this budget line, it is sufficient to determine the coordinates of two points on the line. When $T = 0$, $12K = 6{,}000$ and therefore $K = 500$. Similarly, when $K = 0$, $5T = 6{,}000$ and therefore $T = 1{,}200$. The graph of the budget line is given by the straight line joining the points $T = 0$, $K = 500$ and $T = 1{,}200$, $K = 0$. The graph of the budget line is sketched in Fig. 2.6.

Example 2.11

A person has £120 to spend on two goods (X, Y) whose respective prices are £3 and £5.

1. Draw a budget line showing all the different combinations of the two goods that can be bought with the given budget (B).
2. What happens to the original budget line if the budget falls by 25%?
3. What happens to the original budget line if the price of X doubles?
4. What happens to the original budget line if the price of Y falls to £4?

Draw the new budget lines in each case.

Solution.

1. The general equation of a budget line is

$$P_X X + P_Y Y = B$$

where P_X is the price of X and P_Y is the price of Y. Now if $P_X = 3,\ P_Y = 5,\ B = 120$, then the equation of the budget line is

$$3X + 5Y = 120.$$

We can rearrange this equation to give

$$Y = -\frac{3}{5}X + 24.$$

The graph of this budget line is represented by the solid line in Fig. 2.7.

2. If the budget falls by 25% it is reduced by 25% of £120, i.e., £30. The new budget $B = £120 - £30 = £90$. The equation for the new budget line is

$$3X + 5Y = 90,$$

which, after rearrangement, can be written in the form

$$Y = -\frac{3}{5}X + 18.$$

This line has the same slope as the original budget line but lies to the left of it. This is the dashed line in Fig. 2.7.

3. If $P_X = 6$ the budget equation becomes

$$6X + 5Y = 120$$

or

$$Y = -\frac{6}{5}X + 24.$$

This time the intercept remains the same as the original budget line but the slope is steeper – the slope is $-6/5$ compared with the slope of $-3/5$ of the original budget line. The graph of this budget line is represented by the long dashed line in Fig. 2.7.

4. If $P_Y = 4$, then the budget equation is

$$3X + 4Y = 120,$$

or

$$Y = -\frac{3}{4}X + 30.$$

This time both the slope and the intercept change. See the dash-dot line in Fig. 2.7.

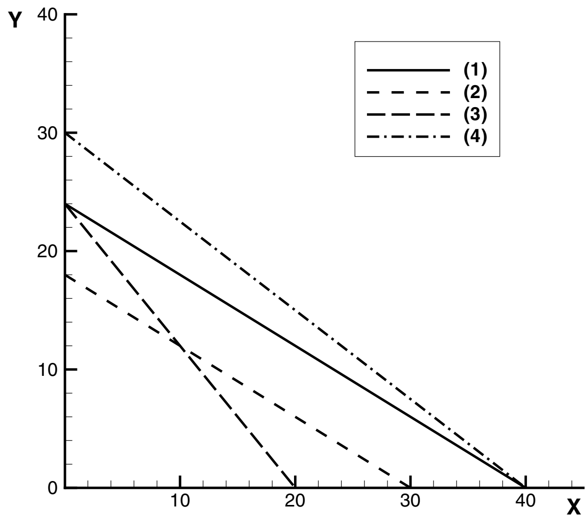

Figure 2.7 The graph of the budget lines in Example 2.11.

2.6 Supply and Demand Analysis

Microeconomics is concerned with the analysis of the economic theory and policy of individual firms and markets. The mathematics we have introduced so far can be used to calculate the market equilibrium in which the demand and supply of a particular good balance.

The quantity demanded, Q, of a particular good depends on the market price, P. We shall refer to the way Q depends on P as the **demand equation** or **demand function**. Functions will be defined in more detail later in the book (Chapter 4). Economists normally plot the relationship between price and quantity with Q on the horizontal axis and P on the vertical axis. We assume that this relationship is linear, i.e.,

$$P = aQ + b,$$

for some appropriate constants (parameters) a and b. A graph of a typical linear demand function is the dashed line in Fig. 2.8. Elementary theory shows that demand usually falls as the price of the good rises so the slope of the line is negative, i.e., $a < 0$. We say that P is a decreasing function of Q.

Similarly, the **supply equation** or **supply function** is the relation between the quantity, Q, of a good that producers plan to bring to the market and

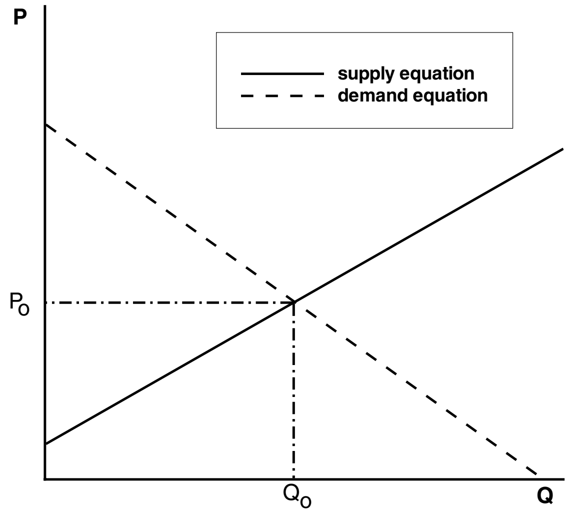

Figure 2.8 The graph of typical linear demand and supply equations. The point of intersection provides the point of equilibrium for the model.

the price, P, of the good. A typical linear supply curve is the solid line in Fig. 2.8. Economic theory indicates that as the price rises, so does the supply. Mathematically, P is then said to be an increasing function of Q. Note that the supply Q is zero when $P = b$. It is only when the price exceeds this threshold level that the producers decide that it is worth supplying any good whatsoever.

We are interested in the interplay between supply and demand. Of particular significance is the point of intersection of the demand and supply curves (see Fig. 2.8). At this point, the market is said to be in **equilibrium** because the quantity demanded is equal to the quantity supplied. The corresponding price, P_0, and quantity, Q_0, are called the **equilibrium price and quantity**. It is also of interest to observe the effect of a shift of the market price away from its equilibrium price.

Example 2.12

The demand and supply equations of a good are given by

$$\begin{aligned} 4P &= -Q_d + 240, \\ 5P &= Q_s + 30. \end{aligned}$$

Determine the equilibrium price and quantity.

Solution. At market equilibrium, we have

$$Q_d = Q_s = Q, \text{ say },$$

where Q is the equilibrium quantity. In this case, the demand and supply equations become

$$\begin{aligned} 4P &= -Q + 240, \\ 5P &= Q + 30. \end{aligned}$$

This is a system of two simultaneous equations in the unknowns P and Q. We can eliminate Q from the system by adding the two equations. This gives

$$9P = 270.$$

Then, dividing both sides by 9 gives the equilibrium price

$$P = 30.$$

Finally, the equilibrium quantity Q is determined by substituting this value into either of the demand or supply equations. The supply equation gives

$$5 \times 30 = Q + 30,$$

which, after rearrangement yields the equilibrium quantity

$$Q = 120.$$

Example 2.13

The demand and supply functions of a good are given by

$$\begin{aligned} P &= -Q_d + 125, \\ 2P &= 3Q_s + 30. \end{aligned}$$

Determine the equilibrium price and quantity. Determine also the effect on the market equilibrium if the government decides to impose a fixed tax of £5 on each good. Who pays the tax?

Solution. At market equilibrium, we have

$$Q_d = Q_s = Q, \text{ say },$$

where Q is the equilibrium quantity. In this case, the demand and supply equations become

$$\begin{aligned} P &= -Q + 125, && (2.17) \\ 2P &= 3Q + 30. && (2.18) \end{aligned}$$

This is a system of two simultaneous equations in the unknowns P and Q. We can eliminate Q from the system by multiplying the demand equation (Eq. (2.17)) by 3:

$$3P = -3Q + 375, \tag{2.19}$$

and adding the resulting equation (2.17) to the supply equation (2.18). This gives

$$5P = 405,$$

which, after dividing both sides by 5 gives the equilibrium price

$$P = 81.$$

Finally, the equilibrium quantity Q is determined by substituting this value into either of the demand or supply equations. The demand equation gives

$$81 = -Q + 125,$$

which, after rearrangement yields the equilibrium quantity

$$Q = 125 - 81 = 44.$$

If the government imposes a fixed tax of £5 on each good, then the original supply equation needs to be modified. This is because the amount the supplier receives as a result of each sale is the amount that the consumer pays (P) less the tax (£5), i.e., $P-5$. Thus, the new supply equation is obtained by replacing P by $P - 5$ in the original supply equation:

$$2(P - 5) = 3Q_s + 30. \tag{2.20}$$

This equation can be simplified by multiplying out the bracket on the left-hand side and taking the constant term to the right-hand side. The new supply equation becomes

$$2P - 10 = 3Q_s + 30,$$

or

$$2P = 3Q_s + 40. \tag{2.21}$$

We then proceed as before to determine the equilibrium price and quantity for the new situation. At market equilibrium, we have

$$Q_d = Q_s = Q, \text{ say },$$

where Q is the equilibrium quantity. In this case, the demand and supply equations become

$$P = -Q + 125, \tag{2.22}$$

$$2P = 3Q + 40. \tag{2.23}$$

We can eliminate Q from the system by multiplying the demand equation (Eq. (2.22)) by 3:

$$3P = -3Q + 375, \tag{2.24}$$

and adding the resulting equation (2.24) to the supply equation (2.23). This gives

$$5P = 415,$$

which, after dividing both sides by 5 gives the equilibrium price

$$P = 83.$$

Finally, the equilibrium quantity Q is determined by substituting this value into either of the demand or supply equations. The demand equation gives

$$83 = -Q + 125,$$

which, after rearrangement yields the equilibrium quantity

$$Q = 125 - 83 = 42.$$

The influence of government taxation on the equilibrium price is to increase it from £81 to £83. Therefore, not of all of the tax is passed on to the consumer. The consumer pays an extra £2 per good after tax has been imposed. The remaining part of the tax is borne by the supplier.

2.6.1 Multicommodity Markets

At the beginning of this section, we looked at supply and demand analysis for a single good. We extend these ideas now to a multicommodity market. Suppose that there are two goods in related markets, which we call good 1 and good 2.

The demand for either good depends on the prices of both good 1 and good 2. If the corresponding demand functions are linear, then

$$\begin{aligned} Q_{d_1} &= a_1 + b_1 P_1 + c_1 P_2 \\ Q_{d_2} &= a_2 + b_2 P_1 + c_2 P_2 \end{aligned}$$

where P_i and Q_{d_i} denote the price and demand for the ith good, and a_i, b_i, and c_i are constants depending on the model. For the first equation $a_1 > 0$, because there is a positive demand when the prices of both goods are zero. Also $b_1 < 0$, because the demand of a good falls as its price rises. The sign of c_1 depends on the nature of the two goods. If the goods are **substitutable**, then an increase in the price of good 2 would mean that consumers would switch from good 2 to good 1, causing Q_{d_1} to increase. Substitutable goods are therefore characterized by a positive value of c_1. On the other hand, if the goods are **complementary**, then a rise in the price of either good would see the demand fall so c_1 is negative. Similar results apply to the signs of a_2, b_2 and c_2.

Example 2.14

The demand and supply functions for two interdependent commodities are given by

$$\begin{aligned} Q_{d_1} &= 145 - 2P_1 + P_2 \\ Q_{s_1} &= -45 + P_1 \\ Q_{d_2} &= 30 + P_1 - 2P_2 \\ Q_{s_2} &= -40 + 5P_2 \end{aligned}$$

where Q_{d_i}, Q_{s_i}, and P_i denote the quantity demanded, quantity supplied, and price of good i, respectively. Determine the equilibrium price and quantity for this two-commodity model. Are these goods substitutable or complementary? Give reasons for your answer.

Solution. At equilibrium, the quantity supplied is equal to the quantity demanded for each good, so that

$$Q_{d_1} = Q_{s_1} \text{ and } Q_{d_2} = Q_{s_2}.$$

Let us write these respective common values as Q_1 and Q_2. Then for good 1 we have

$$\begin{aligned} Q_1 &= 145 - 2P_1 + P_2 \\ Q_1 &= -45 + P_1 \end{aligned}$$

Therefore

$$145 - 2P_1 + P_2 = -45 + P_1$$

which simplifies to give

$$3P_1 - P_2 = 190.$$

Similarly for good 2 we have

$$\begin{aligned} Q_2 &= 30 + P_1 - 2P_2 \\ Q_2 &= -40 + 5P_2 \end{aligned}$$

Therefore

$$30 + P_1 - 2P_2 = -40 + 5P_2$$

which simplifies to give

$$-P_1 + 7P_2 = 70.$$

We have therefore shown that the equilibrium prices satisfy the simultaneous equations

$$3P_1 - P_2 = 190 \qquad (2.25)$$

$$-P_1 + 7P_2 = 70 \qquad (2.26)$$

These equations can be solved by elimination. Multiply (2.26) by 3. This gives

$$\begin{aligned} 3P_1 - P_2 &= 190 \\ -3P_1 + 21P_2 &= 210 \end{aligned}$$

Adding these two equations yields

$$20P_2 = 400,$$

and so $P_2 = 20$. Substituting this value of P_2 back into (2.25):

$$3P_1 = 190 + 20 = 210,$$

which gives $P_1 = 70$. Finally, substituting these values of P_1 and P_2 back into the original supply equations, we obtain

$$Q_1 = 145 - 140 + 20 = 25$$

and

$$Q_2 = -40 + 100 = 60.$$

On inspection of the demand equation for good 1, we see that the demand for this good increases when the price of good 2 increases. This is characterized by a positive coefficient of P_2 in this equation. Therefore, the two goods are substitutable.

EXERCISES

2.1. Solve the following linear equations:

a) $3x - 4 = 2$,

b) $\left(\frac{2x-1}{3}\right) = \left(\frac{3x-1}{4}\right) + 1$.

2.2. Solve the system of equations

$$\begin{aligned} 3x - 2y &= 4 \\ x - 2y &= 2. \end{aligned}$$

2.3. Solve the system of equations

$$\begin{aligned} 3x + 5y &= 19 \\ -5x + 2y &= -11. \end{aligned}$$

2.4. Sketch the graph of the straight line $y = -x + 2$ for $-1 \leq x \leq 5$.

2.5. Sketch the graph of the straight line $y = 2x - 3$ for $0 \leq x \leq 4$.

2.6. Find the slope of the straight line passing through the points $(-1, -3)$ and $(4, 2)$.

2.7. Find the slope of the straight line passing through the points $(0, 0)$ and $(2, 1)$.

2.8. A person has €60 to spend on two goods, X and Y, whose respective prices are €6 and €4.

a) Draw a budget line showing all the different combinations of the two goods that can be bought within the given budget.

b) What happens to the original budget line if the budget is increased by 20%?

c) What happens to the original budget line if the price of X is halved?

2.9. The demand and supply equations for a good are given by

$$\begin{aligned} 2P &= -Q_d + 125, \\ 8P &= Q_s + 45, \end{aligned}$$

where P, Q_d, and Q_s denote the price, quantity demanded, and quantity supplied, respectively.

a) Determine the equilibrium price and quantity.

b) Determine the effect on the market equilibrium if the government decides to impose a fixed tax of £2.50 on each good. Who pays the tax?

2.10. The demand and supply functions of a good are given by

$$P + 2Q_d = 144$$
$$4P - 3Q_s = 136$$

where P, Q_d, and Q_s, denote the price, quantity demanded, and quantity supplied, respectively.

a) Determine the equilibrium price and quantity.

b) Determine the effect on the market equilibrium if the government decides to impose a fixed tax of \$11 on each good. Who pays the tax?

2.11. The demand and supply functions of a good are given by

$$\begin{array}{rcrcr} 4P & = & -Q_d & + & 102 \\ 5P & = & Q_s & + & 6 \end{array}$$

where P, Q_d, and Q_s denote the price, quantity demanded, and quantity supplied, respectively.

a) Determine the equilibrium price and quantity.

b) Determine the effect on the market equilibrium if the government decides to impose a fixed tax of £9 on each good. Who pays the tax?

2.12. The demand and supply equations for two complementary goods, trousers (T) and jackets (J), are given by

$$\begin{array}{rcl} Q_{d_T} & = & 410 - 5P_T - 2P_J \\ Q_{s_T} & = & -60 + 3P_T \end{array}$$

and

$$\begin{array}{rcl} Q_{d_J} & = & 295 - P_T - 3P_J \\ Q_{s_J} & = & -120 + 2P_J \end{array}$$

respectively, where Q_{d_T}, Q_{s_T}, and P_T denote the quantity demanded, quantity supplied, and price of trousers, and Q_{d_J}, Q_{s_J}, and P_J denote the quantity demanded, quantity supplied, and price of jackets. Determine the equilibrium price and quantity for this two-market model.

3
Quadratic Equations

3.1 Introduction

Linear equations and methods for their solution were introduced in the previous chapter. As we have seen, the graphs of linear functions are straight lines and therefore their slopes are constant. This means that the function changes by a constant amount whenever the dependent variable changes by the same fixed value. This type of behaviour is not always observed in real-life applications in economics. It is, therefore, necessary to introduce an added level of sophistication to the mathematical modelling. This is achieved through the introduction of nonlinear functions. The simplest nonlinear function is the quadratic function. This function takes the general form

$$f(x) = ax^2 + bx + c, \tag{3.1}$$

where $a \neq 0$, b and c are constants. The condition $a \neq 0$ is to prevent the occurrence of the degenerate case in which (3.1) reduces to a linear function.

If the profit function for a firm is given by a quadratic expression, then one can determine the level of output for which the firm breaks even by solving a quadratic equation. Additionally, one can determine the maximum profit and the level of output for which it is attained by algebraically manipulating the expression for the function. For more general nonlinear functions, the maximum and/or minimum values of a function can be determined using the techniques of calculus (see Chapter 7), but for a quadratic function this can be achieved using algebra.

Certain total cost and total revenue functions are examples of quadratic functions and are defined in terms of a quadratic expression involving the demand.

3.2 Graphs of Quadratic Functions

In the case of a linear function of the form $f(x) = dx + e$, the parameters d and e can be interpreted in terms of properties of the graph of the function. The value of d, the coefficient of x, gives the slope or gradient of the function, and the value of e, the constant term, tells us where the straight line intercepts the y-axis. A natural question to ask is whether the parameters in the expression defining the general quadratic function $f(x) = ax^2 + bx + c$ can be interpreted in a similar way in order to help us sketch its graph.

If we evaluate the function $f(x) = ax^2 + bx + c$ when $x = 0$ we obtain $f(0) = c$. Therefore, the quadratic function intercepts the y-axis at the location $y = c$. The values of the other parameters cannot be interpreted in such a simple manner. However, the sign of the parameter a tells us something about the shape of the graph. If $a > 0$, then the graph of $f(x)$ has a $\bigcup$ shape, whereas if $a < 0$ the graph of $f(x)$ has a $\bigcap$ shape. This information gives us a rough idea of what the graph of a quadratic function looks like. An additional aid is to tabulate the function at a sequence of integer values of x and to draw a smooth curve through the set of points. For example, let us sketch the graph of the quadratic function $f(x) = x^2$ for $-3 \le x \le 3$. If we compare the coefficients of this function with those of the general quadratic function, we find that $a = 1$, and $b = c = 0$. Therefore, the graph of this function intercepts the y-axis at the origin as $c = 0$ and has a $\bigcup$ shape as $a > 0$. The values of this function are tabulated in Table 3.1 for integer values of x for which $-3 \le x \le 3$, and the graph of the function is shown in Fig. 3.1.

Now consider the function $f(x) = 2x^2 + 3x - 2$. Again comparison with the general quadratic function (3.1) shows that $a = 2$, $b = 3$, and $c = -2$. The graph is again of a $\bigcup$ shape since $a > 0$ and it intercepts the y-axis at $y = -2$. The values of this function for integer values of x between -4 and 2 are shown

Table 3.1 Table of values of the function $f(x) = x^2$ for integer values of x for which $-3 \le x \le 3$. The graph of this function is shown in Fig. 3.1.

x	-3	-2	-1	0	1	2	3
$f(x)$	9	4	1	0	1	4	9

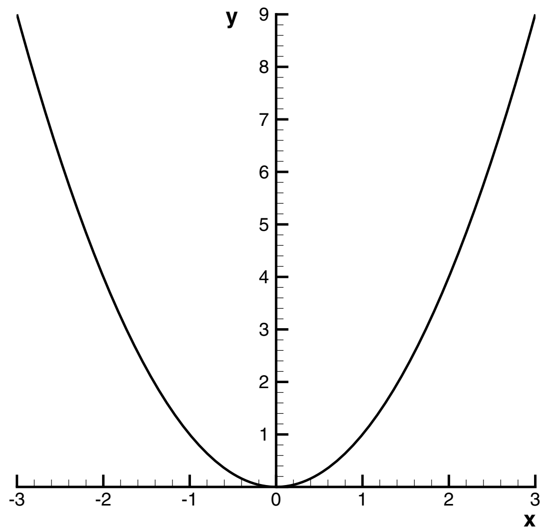

Figure 3.1 The graph of the function $f(x) = x^2$ for $-3 \leq x \leq 3$.

in Table 3.2, and the graph of the function is shown in Fig. 3.2. The graph of this function crosses the x-axis in two places, at $x = -2$ and $x = 1/2$. These values of x satisfy the quadratic equation $2x^2 + 3x - 2 = 0$ since $y = f(x) = 0$ at these two points. The values of x that satisfy the equation $f(x) = 0$ are known as the **roots** or **solutions** of the equation. These two terms are used interchangeably. Therefore, we say that $x = -2$ and $x = 1/2$ are the roots or solutions of the quadratic equation $2x^2 + 3x - 2 = 0$.

The next function we consider is $f(x) = 2x - x^2$. This function has a

Table 3.2 Table of values of the function $f(x) = 2x^2 + 3x - 2$ for integer values of x for which $-4 \leq x \leq 2$. The graph of this function is shown in Fig. 3.2.

x	-4	-3	-2	-1	0	1	2
$2x^2$	32	18	8	2	0	2	8
$3x$	-12	-9	-6	-3	0	3	6
-2	-2	-2	-2	-2	-2	-2	-2
$f(x)$	18	7	0	-3	-2	3	12

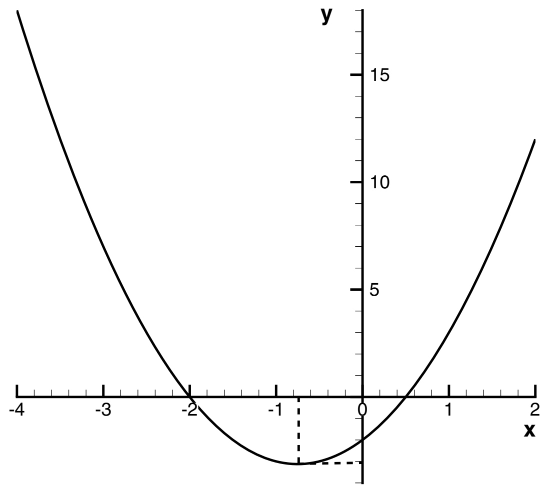

Figure 3.2 The graph of the function $f(x) = 2x^2 + 3x - 2$ for $-4 \leq x \leq 2$.

negative coefficient of x^2. In terms of the general quadratic function (3.1), we have $a = -1$, $b = 2$, and $c = 0$. Since $a < 0$, the graph of the function has a $\bigcap$ shape. The graph intersects the x-axis at the origin since $y = 2x - x^2 =$ when $x = 0$. The other intercept (intersection) with the x-axis is the other root of the equation $2x - x^2 = 0$, namely $x = 2$. This is evident since $2x - x^2 = (2 - x)x$. Therefore, one of the roots of the equation $2x - x^2 = 0$ is $x = 0$. The other root is $x = 2$. The values of this function for integer values of x between -2 and 4 are shown in Table 3.3, and the graph of the function is shown in Fig. 3.3.

Finally, we consider the function $f(x) = x^2 - 2x + 2$. Comparison with the general quadratic function (3.1) gives $a = 1$, $b = -2$, and $c = 2$. The

Table 3.3 Table of values of the function $f(x) = 2x - x^2$ for integer values of x for which $-2 \leq x \leq 4$. The graph of this function is shown in Fig. 3.3.

x	-2	-1	0	1	2	3	4
$2x$	-4	-2	0	2	4	6	8
$-x^2$	-4	-1	0	-1	-4	-9	-16
$f(x)$	-8	-3	0	1	0	-3	-8

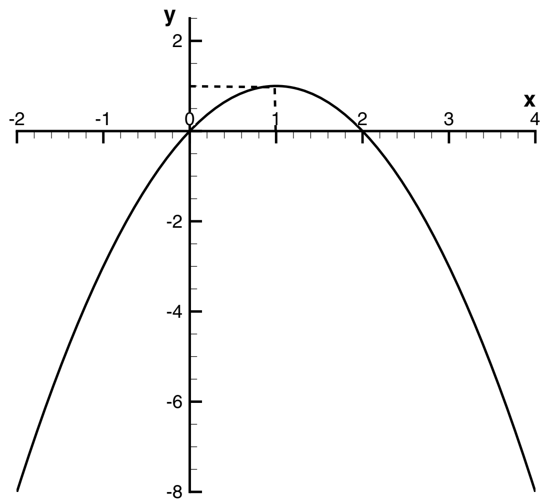

Figure 3.3 The graph of the function $f(x) = 2x - x^2$ for $-2 \leq x \leq 4$.

values of this function at integer values of x between -2 and 4 are shown in Table 3.4, and the graph of the function is shown in Fig. 3.4. Note that the graph of this function does not cross the x-axis. It lies entirely above the x-axis, i.e., $f(x) > 0$ for all values of x. Therefore, there are no real roots of the corresponding equation $x^2 - 2x + 2 = 0$.

The graph of a quadratic function is known as a **parabola**. On inspection of Figs. 3.1–3.4, we observe that a parabola is symmetric about a vertical line $x = h$, where h is some constant. This line is known as the **axis of symmetry** of

Table 3.4 Table of values of the function $f(x) = x^2 - 2x + 2$ for integer values of x for which $-2 \leq x \leq 4$. The graph of this function is shown in Fig. 3.4.

x	-2	-1	0	1	2	3	4
x^2	4	1	0	1	4	9	16
$-2x$	4	2	0	-2	-4	-6	-8
2	2	2	2	2	2	2	2
$f(x)$	10	5	2	1	2	5	10

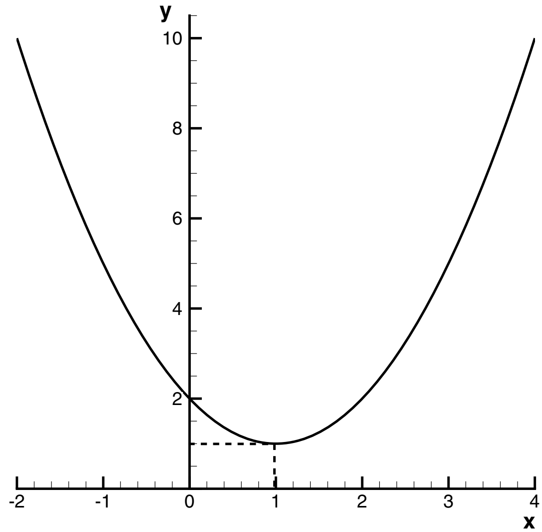

Figure 3.4 The graph of the function $f(x) = x^2 - 2x + 2$ for $-3 \leq x \leq 3$.

the parabola. The point of intersection of a parabola with its axis of symmetry is called the **vertex**. For example, the quadratic function $f(x) = x^2 - 3x + 2$ has $x = 3/2$ as its axis of symmetry and $(3/2, -1/4)$ as its vertex. If $a > 0$, then the y component of the vertex provides the minimum value of the quadratic function. Similarly, if $a < 0$, then the y component of the vertex provides the maximum value of the quadratic function.

If a quadratic function can be expressed in the form

$$f(x) = a(x - h)^2 + k, \tag{3.2}$$

then the axis of symmetry is $x - h = 0$ and the vertex is the point with coordinates (h, k). Let us rearrange the expression defining the general quadratic expression so that it is in this form. To do this, we use a process known as completing the square. First, we extract a factor a from the quadratic expression $ax^2 + bx + c$, i.e.,

$$ax^2 + bx + c = a\left(x^2 + \frac{b}{a}x + \frac{c}{a}\right) \tag{3.3}$$

Then, we express the first two terms inside the bracket on the right-hand side of (3.3), viz. $x^2 + (b/a)x$ as the difference between two squares:

$$x^2 + \frac{b}{a}x = \left(x + \frac{b}{2a}\right)^2 - \left(\frac{b}{2a}\right)^2.$$

Therefore,

$$ax^2 + bx + c = a\left[\left(x + \frac{b}{2a}\right)^2 - \frac{b^2}{4a^2} + \frac{c}{a}\right] \tag{3.4}$$

$$= a\left[\left(x + \frac{b}{2a}\right)^2 + \frac{4ac - b^2}{4a^2}\right], \tag{3.5}$$

in which the last two terms in (3.4) have been combined to form a single fraction. Finally, we arrive at

$$ax^2 + bx + c = a\left(x + \frac{b}{2a}\right)^2 + \frac{4ac - b^2}{4a}. \tag{3.6}$$

So comparing (3.6) with (3.2), we have

$$h = -\frac{b}{2a}, \quad k = \frac{4ac - b^2}{4a}.$$

In the above example, rearrangement gives

$$f(x) = \left(x - \frac{3}{2}\right)^2 - \frac{9}{4} + 2 = \left(x - \frac{3}{2}\right)^2 - \frac{1}{4},$$

from which we deduce that the axis of symmetry is $x - 3/2 = 0$ and the vertex is $(3/2, -1/4)$. Next consider the function $f(x) = 2x - x^2$. This expression can be rearranged as follows to determine the axis of symmetry and vertex:

$$\begin{aligned} f(x) &= 2x - x^2 \\ &= -(x-1)^2 + 1 \end{aligned}$$

Therefore, for this function we have $a = -1$, $h = 1$, and $k = 1$, so the axis of symmetry is the line $x = 1$ and the vertex is located at the point with coordinates $(1, 1)$.

Finally, we consider the function $f(x) = 2x^2 + 3x - 2$. As before we write

$$\begin{aligned} f(x) &= 2x^2 + 3x - 2 \\ &= 2\left[x^2 + \frac{3}{2}x - 1\right] \\ &= 2\left[\left(x + \frac{3}{4}\right)^2 - \frac{9}{16} - 1\right] \\ &= 2\left(x + \frac{3}{4}\right)^2 - \frac{25}{8} \end{aligned}$$

Therefore, for this function we have $a = 2$, $h = -3/4$, and $k = -25/8$, so the axis of symmetry is the line $x = -3/4$ and the vertex is located at the point with coordinates $(-3/4, -25/8)$. In Table 3.5, we provide the axes and vertices of the four quadratic functions we have investigated in this chapter.

Table 3.5 Axes and vertices of some quadratic functions.

$f(x)$	Axis	Vertex
x^2	$x = 0$	$(0, 0)$
$2x - x^2$	$x = 1$	$(1, 1)$
$2x^2 + 3x - 2$	$x = -3/4$	$(-3/4, -25/8)$
$x^2 - 2x + 2$	$x = 1$	$(1, 1)$

3.3 Quadratic Equations

There are a number of techniques for determining the roots of a quadratic equation. Knowledge of the roots of a quadratic equation can be an additional aid to sketching the graph of a quadratic function. If the expression defining a quadratic function can be factorised as a product of linear factors, then equating each of the factors to zero and solving the resulting linear equations will provide the roots.

Example 3.1

Solve $x^2 + 13x + 30 = 0$ using factorization.

Solution. First, we factorize the quadratic expression $x^2 + 13x + 30$ as a product of two linear factors $(x + A)$ and $(x + B)$, where A and B are two constants that need to be determined. Since

$$(x + A)(x + B) = x^2 + (A + B)x + AB,$$

then the constants A and B need to be chosen so that

$$A + B = 13, \quad AB = 30.$$

The possible combinations of integers whose product is 30 are 30×1, 15×2, 10×3, and 6×5. Of course, one also has the combinations in which the integers have been negated such as $(-30) \times (-1)$, but out of these combinations the only one for which the pair of integers sums to 13 is 10×3. Therefore, we choose $A = 10$ and $B = 3$, i.e.,

$$x^2 + 13x + 10 = (x + 10)(x + 3).$$

We now solve the equation $(x + 3)(x + 10) = 0$. For the product of the two linear terms $x + 3$ and $x + 10$ to be zero, at least one of them must be zero. So either $x + 3 = 0$ or $x + 10 = 0$.

If $x + 3 = 0$ then $x = -3$, and if $x + 10 = 0$ then $x = -10$. Therefore, the roots of the equation $x^2 + 13x + 30 = 0$ are $x = -3$ and $x = -10$.

Example 3.2

Solve the quadratic equation $2x^2 - 11x + 12 = 0$ using factorization.

Solution. As in the previous example, the first step is to factorize the quadratic expression $2x^2 - 11x + 12$ as a product of linear factors. These linear factors must be of the form $(2x + A)$ and $(x + B)$ in order to retrieve the quadratic factor $2x^2$, where A and B are two positive constants. Since

$$(2x + A)(x + B) = 2x^2 + (A + 2B)x + AB,$$

then the constants A and B need to be chosen so that

$$A + 2B = -11, \quad AB = 12.$$

The possible combinations of integers whose product is 12 are 12×1, 6×2, 4×3, -4×-3, -6×-2, and -12×-1. The only pair of integers amongst these for which $A + 2B = -11$ is $A = -3$ and $B = -4$. Therefore, we have

$$2x^2 - 11x + 12 = (2x - 3)(x - 4).$$

The problem now is to solve the equation

$$(2x - 3)(x - 4) = 0.$$

Either $2x - 3 = 0$ or $x - 4 = 0$. If $2x - 3 = 0$ then $2x = 3$ and $x = 3/2$. If $x - 4 = 0$, then $x = 4$. Therefore, the two roots of the equation $2x^2 - 11x + 12 = 0$ are $x = 3/2$ and $x = 4$.

Most quadratic expressions, however, do not factorise easily in the sense that they cannot be expressed as a product of linear factors with integer coefficients, even if the coefficients of the quadratic equation are integers. For example, the quadratic equation $3x^2 - 9x + 5 = 0$ cannot be factored into a product of linear factors with integer coefficients. Clearly, a more systematic approach is required.

There is a formula for finding the solution to a quadratic equation

$$ax^2 + bx + c = 0. \tag{3.7}$$

The formula may be derived by the process known as completing the square that was introduced in Section 3.2. We assume that $a \neq 0$. Using (3.6) we see that (3.7) is equivalent to

$$a\left(x + \frac{b}{2a}\right)^2 + \frac{4ac - b^2}{4a} = 0.$$

Dividing both sides by a and taking the last term to the right-hand side yields

$$\left(x + \frac{b}{2a}\right)^2 = \frac{b^2 - 4ac}{4a^2}.$$

Now taking the square root of both sides gives

$$x + \frac{b}{2a} = \pm\sqrt{\frac{b^2 - 4ac}{4a^2}} = \pm\frac{\sqrt{b^2 - 4ac}}{2a}.$$

Finally, subtracting $b/(2a)$ from both sides we arrive at the formula for the roots of a quadratic equation:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \tag{3.8}$$

This is an important formula for the roots (that is, solutions) of a quadratic equation, which we highlight:

> The solutions of the quadratic equation $ax^2 + bx + c = 0$ are
>
> $$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

The number of solutions of a quadratic equation depends on the sign of the expression under the square root sign in this formula. A quadratic equation has two, one or no solutions depending on whether the expression $b^2 - 4ac$ is positive, zero, or negative:

- If $b^2 - 4ac > 0$, there are two solutions

$$x = \frac{-b + \sqrt{b^2 - 4ac}}{2a} \text{ and } x = \frac{-b - \sqrt{b^2 - 4ac}}{2a}.$$

- If $b^2 - 4ac = 0$, then there is one solution

$$x = -\frac{b}{2a}.$$

- If $b^2 - 4ac < 0$, then there are no solutions since the square root of $b^2 - 4ac$ does not exist in this case.

Example 3.3

Solve the quadratic equation

$$4x^2 - 11x + 6 = 0$$

using the formula.

Solution. Compare the coefficients of this equation with those of the general quadratic equation. If we do this, we notice that $a = 4$, $b = -11$, and $c = 6$. Inserting these values into the formula (3.8) gives

$$\begin{aligned} x &= \frac{-(-11) \pm \sqrt{(-11)^2 - 4 \times 4 \times 6}}{2 \times 4} \\ &= \frac{11 \pm \sqrt{121 - 96}}{8} \\ &= \frac{11 \pm \sqrt{25}}{8} \\ &= \frac{11 \pm 5}{8} \end{aligned}$$

Therefore, the two solutions are

$$x = \frac{11+5}{8} = \frac{16}{8} = 2, \quad \text{and } x = \frac{11-5}{8} = \frac{6}{8} = \frac{3}{4}.$$

Example 3.4

Solve the quadratic equation

$$x^2 - 2x - 15 = 0$$

using the formula.

Solution. Compare the coefficients of this equation with those of the general quadratic equation. If we do this, we notice that $a = 1$, $b = -2$, and $c = -15$. Inserting these values into the formula (3.8) gives

$$\begin{aligned} x &= \frac{-(-2) \pm \sqrt{(-2)^2 - 4 \times 1 \times (-15)}}{2 \times 1} \\ &= \frac{2 \pm \sqrt{4 - (-60)}}{2} \\ &= \frac{2 \pm \sqrt{64}}{2} \\ &= \frac{2 \pm 8}{2} \end{aligned}$$

Therefore, the two solutions are

$$x = \frac{2+8}{2} = \frac{10}{2} = 5, \quad \text{and } x = \frac{2-8}{2} = \frac{-6}{2} = -3.$$

Example 3.5

Solve the quadratic equation

$$3x^2 - 9x + 5 = 0$$

using the formula.

Solution. Compare the coefficients of this equation with those of the general quadratic equation. If we do this, we notice that $a = 3$, $b = -9$, and $c = 5$. Inserting these values into the formula (3.8) gives

$$\begin{aligned} x &= \frac{-(-9) \pm \sqrt{(-9)^2 - 4 \times 3 \times 5}}{2 \times 3} \\ &= \frac{9 \pm \sqrt{81 - 60}}{6} \\ &= \frac{9 \pm \sqrt{21}}{6} \end{aligned}$$

Note that 21 is not a perfect square, and therefore the roots of this equation can only be expressed in decimal representation to a specified number of decimal places. Therefore, to four decimal places the two solutions of this equation are

$$x = \frac{9+\sqrt{21}}{6} = 2.2638, \quad \text{and } x = \frac{9-\sqrt{21}}{6} = 0.7362.$$

Example 3.6

Solve the quadratic equation $x^2 - 18x + 45 = 0$ by completing the square.

Solution. In this example $a = 1$, $b = -18$, and $c = 45$. Therefore, using (3.6) we may write the equation in the form

$$(x-9)^2 - 81 + 45 = 0,$$

or

$$(x-9)^2 = 36.$$

Then taking the square root of both sides gives

$$x - 9 = \pm 6.$$

Either $x - 9 = 6$, which means that $x = 15$. Or $x - 9 = -6$, which means that $x = 3$.

3.4 Applications to Economics

One function of particular interest in economics is the **profit function**. We denote this function by the Greek symbol π. The profit function is defined to be the difference between total revenue, TR, and the total cost, TC, i.e.,

$$\pi = TR - TC.$$

The total revenue received from the sale of Q goods at price P is given by the product of P and Q, i.e.,

$$TR = P \times Q.$$

The total cost function relates the cost of production to the level of output, Q, and is the sum of the fixed costs, FC, and variable costs, $VC \times Q$, where VC denotes the variable cost per unit of output. Fixed costs include, for example, the cost of land, rental, equipment, and skilled labour. Variable costs include, for example, the cost of raw materials, energy, and unskilled labour. The total cost in producing Q goods is given by

$$TC = FC + (VC) \times Q.$$

Thus the profit function is

$$\pi = P \times Q - [FC + (VC) \times Q] = PQ - FC - (VC) \times Q.$$

Note that care needs to be exercised in removing the brackets. It is important to remember that the negative sign outside the square brackets negates **all** terms inside the brackets when the brackets are moved.

Example 3.7

If fixed costs are 18, variable costs per unit are 4, and the demand function is

$$P = 24 - 2Q$$

obtain an expression for π in terms of Q and hence sketch a graph of π against Q.

1. For what values of Q does the firm break even?
2. What is the maximum profit?

Table 3.6 Table of values of the profit function $\pi = -2Q^2 + 20Q - 18$ for even integer values of Q for which $0 \leq Q \leq 10$.

Q	0	2	4	6	8	10
$-2Q^2$	0	−8	−32	−72	−128	−200
$20Q$	0	40	80	120	160	200
−18	−18	−18	−18	−18	−18	−18
π	−18	14	30	30	14	−18

Solution. The total revenue function is given by

$$TR = P \times Q = (24 - 2Q)Q = 24Q - 2Q^2,$$

where we have used the demand function $P = 24 - 2Q$ to eliminate P in the expression defining TR. We have expressed TR solely in terms of the level of output, Q. The total cost function is given by

$$TC = FC + (VC) \times Q = 18 + 4Q,$$

since $FC = 18$ and $VC = 4$. We can now obtain an expression for the profit function by subtracting the expression for TC from the expression for TR, i.e.,

$$\begin{aligned} \pi &= TR - TC \\ &= 24Q - 2Q^2 - (18 + 4Q) \\ &= 24Q - 2Q^2 - 18 - 4Q \\ &= -2Q^2 + 20Q - 18, \end{aligned}$$

where we have taken care to change the sign of all terms inside the brackets on their removal.

Since the coefficient of Q^2 in the quadratic expression defining π is negative, the graph of the profit function has a $\bigcap$ shape. When $Q = 0$, $\pi = -18$. The profit function is tabulated in Table 3.6 for $0 \leq Q \leq 10$. From this information, we are able to sketch the graph of the function. This is shown in Fig. 3.5.

1. The value of the profit function will be zero (i.e., $\pi = 0$) for values of Q that satisfy the quadratic equation

$$-2Q^2 + 20Q - 18 = 0.$$

Solving this equation using the formula with $a = -2$, $b = 20$, and $c = -18$ yields

$$\begin{aligned} Q &= \frac{-20 \pm \sqrt{400 - 144}}{-4} \\ &= \frac{-20 \pm \sqrt{256}}{-4} \\ &= \frac{-20 \pm 16}{-4} \end{aligned}$$

Therefore, either

$$Q = \frac{-20 + 16}{-4} = \frac{-4}{-4} = 1,$$

or

$$Q = \frac{-20 - 16}{-4} = \frac{-36}{-4} = 9.$$

The profit is zero when $Q = 1$ and $Q = 9$. Therefore, the firm breaks even when $Q = 1$ and $Q = 9$. For $1 < Q < 9$, the profit function is positive (see Fig. 3.5) and the firm is in profit. For values of Q outside this range, i.e., $Q < 1$ and $Q > 9$, the profit function is negative and therefore the firm makes a loss at these levels of output.

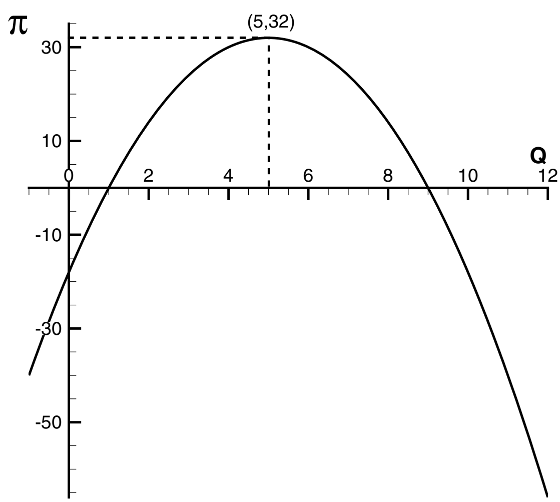

Figure 3.5 The graph of the profit function $\pi = -2Q^2 + 20Q - 18$.

2. To determine the maximum value of the profit function, we complete the square.

$$\begin{aligned}\pi &= -2\left[Q^2 - 10Q + 9\right] \\ &= -2\left[(Q-5)^2 - 25 + 9\right] \\ &= -2\left[(Q-5)^2 - 16\right] \\ &= (-2) \times (Q-5)^2 + (-2) \times (-16) \\ &= -2(Q-5)^2 + 32\end{aligned}$$

Therefore, the maximum profit is $\pi = 32$ since the term $-2(Q-5)^2$ is always negative except when $Q = 5$ when it is zero.

Finally, we return to supply and demand analysis. In Chapter 2, we considered examples in which both the supply and demand functions were linear and determined the equilibrium price and quantity. Although linear models are frequently used in economics because of the simplicity of their mathematical structure, they can also be limiting in the sort of economic behaviour they describe. As we shall see in the next example, it is not necessary for the supply and demand functions to be linear, and, in the case when they are defined by quadratic expressions, the market equilibrium can be determined by solving a quadratic expression.

Example 3.8

Given the supply and demand functions

$$\begin{aligned}P &= Q_s^2 + 12Q_s + 32, \\ P &= -Q_d^2 - 4Q_d + 200,\end{aligned}$$

calculate the equilibrium price and quantity.

Solution. At equilibrium, the quantity supplied is equal to the quantity demanded, so that

$$Q_d = Q_s = Q, \text{ say.}$$

Then the supply and demand equations become

$$\begin{aligned}P &= Q^2 + 12Q + 32, \\ P &= -Q^2 - 4Q + 200.\end{aligned}$$

Equating the expressions on the right-hand sides of these equations, we have

$$Q^2 + 12Q + 32 = -Q^2 - 4Q + 200.$$

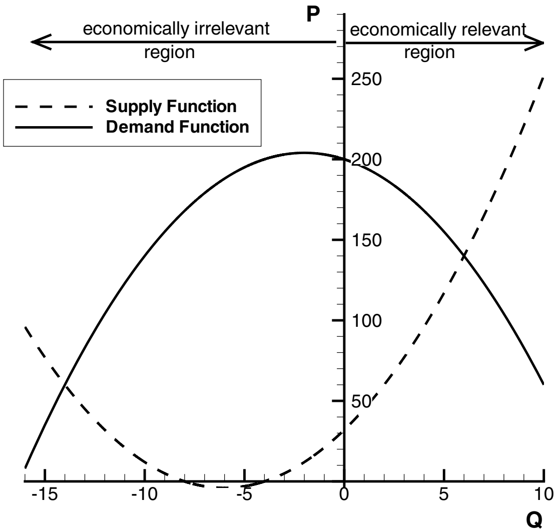

Figure 3.6 The graph of the supply and demand functions in Example 3.8.

We can do this since both expressions are equal to P. Rearranging this equation and collecting like terms yields the quadratic equation

$$2Q^2 + 16Q - 168 = 0.$$

This equation can be simplified by dividing throughout by 2. We then have the quadratic equation

$$Q^2 + 8Q - 84 = 0.$$

Solving this equation using the formula with $a = 1$, $b = 8$, and $c = -84$ yields

$$\begin{aligned} Q &= \frac{-8 \pm \sqrt{8^2 - 4 \times 1 \times (-84)}}{2 \times 1} \\ &= \frac{-8 \pm \sqrt{64 + 336}}{2} \\ &= \frac{-8 \pm \sqrt{400}}{2} \\ &= \frac{-8 \pm 20}{2} \end{aligned}$$

Therefore, either

$$Q = \frac{-8 + 20}{2} = \frac{12}{2} = 6,$$

or

$$Q = \frac{-8 - 20}{2} = \frac{-28}{2} = -14.$$

So the quadratic equation has solutions $Q = 6$ and $Q = -14$. The solution $Q = -14$ can be discarded because a negative quantity does not make sense. Therefore, the equilibrium quantity is 6. The corresponding equilibrium price can be determined by substituting $Q = 6$ into either the supply or demand equation. If we substitute this value into the supply equation, we have

$$P = 6^2 + 12 \times 6 + 32 = 36 + 72 + 32 = 140.$$

Therefore, the equilibrium price is 140.

The graphs of the supply and demand functions are shown in Fig. 3.6. There are two points of intersection. The one for positive Q provides the market equilibrium.

EXERCISES

3.1. Evaluate the function $f(x) = 2x^2 - 9x + 4$ when $x = 0,\ 1,\ 2,\ 3,\ 4,\ 5$. Hence, sketch the graph of this function for $0 \le x \le 5$.

3.2. Evaluate the function $f(x) = -2x^2 - 3x + 3$ when $x = -3, -2, -1, 0, 1, 2$. Hence, sketch the graph of this function for $-3 \le x \le 2$.

3.3. Sketch the graphs of the following functions:

a) $f(x) = 4x^2 - 7x - 2$, for $-2 \le x \le 4;$,

b) $f(x) = 9 - 6x - 8x^2$, for $-3 \le x \le 3$.

3.4. Solve the following quadratic equations using the formula:

a) $x^2 - 4x + 3 = 0$,

b) $3x^2 + 5x - 8 = 0$,

c) $2x^2 - 19x - 10 = 0$.

3.5. Solve the following quadratic equations using factorization:

a) $x^2 + 7x + 10 = 0$,

b) $x^2 - 4x - 5 = 0$,

c) $6x^2 + 19x + 10 = 0$.

3.6. Write the quadratic function $f(x) = x^2 - 8x + 12$ in the form

$$f(x) = a(x - h)^2 + k.$$

What is the equation for the axis of symmetry of this parabola, and what is its vertex? Use this information to sketch the graph of this function.

3.7. If fixed costs are 6, variable costs per unit are 2, and the demand function is

$$P = 15 - 3Q$$

obtain an expression for the profit function π in terms of Q. Hence, sketch a graph of π against Q.

3.8. If fixed costs are 4, variable costs per unit are 3, and the demand function is

$$P = 45 - 4Q$$

obtain an expression for the profit function π in terms of Q. Hence, sketch a graph of π against Q.

4
Functions of a Single Variable

4.1 Introduction

The concept of a function is fundamental to many of the applications that we will encounter in economics. As we have already seen in Chapters 2 and 3, it is a convenient way of expressing a relationship between two variables in terms of a prescribed mathematical rule. More formally, we have the following definition:

Definition 4.1

A **function** f is a rule that assigns to each value of a variable x, called the **independent variable** of the function, one and only one value $f(x)$, referred to as **the value of the function at** x. The variable $y = f(x)$ varies with x and is known as the **dependent variable**.

We sometimes write $f(x)$ to denote the function f if we wish to indicate that the variable is x. The function rule defines the dependent variable in terms of the independent variable. A function of a single variable enables the value of the dependent variable to be determined when the independent variable is specified. A function may therefore be interpreted as a process f that takes an input number x and converts it into only one output number $f(x)$. For example, the function defined by the rule $f(x) = 6x + 2$ is the rule that takes an input number x, multiplies it by 6, and then adds 2 to the product to obtain the output number. Given a value of x, the corresponding value of $f(x)$ can be determined using this rule. For example, if $x = 3$

$$f(x) \ = \ 6 \times 3 + 2 \ = \ 18 + 2 \ = \ 20.$$

We write $f(3) = 20$ and say 'the value of f at $x = 3$ is 20' or 'f of 3 equals 20'.

Example 4.2

Evaluate $f(x) = 2x - 5$ when $x = -1$, $x = 2$ and $x = 4$.

Solution. When $x = -1$,

$$f(x) = 2 \times (-1) - 5 = -2 - 5 = -7,$$

so that $f(-1) = -7$. When $x = 2$,

$$f(x) = 2 \times 2 - 5 = 4 - 5 = -1,$$

so that $f(2) = -1$. When $x = 4$,

$$f(x) = 2 \times 4 - 5 = 8 - 5 = 3,$$

so that $f(4) = -3$.

Functions are generally represented by algebraic formulae that are usually expressed in the form

$$y = f(x),$$

where f defines the precise nature of the functional relationship. We say 'y equals f of x' or 'y is a function of x'. In mathematics, we usually denote functions by letters such as f, g, and h. Examples of functions are :

1. the linear function $y = f(x) = ax + b$;
2. the quadratic function $y = f(x) = ax^2 + bx + c$;
3. the power function $y = f(x) = ax^n$;

where a, b, c and n are constants.

Example 4.3

Given $f(x) = x^2 + 4x - 5$, find $f(2)$ and $f(-3)$.

Solution.

$$f(2) = 2^2 + 4(2) - 5 = 4 + 8 - 5 = 7$$
$$f(-3) = (-3)^2 + 4(-3) - 5 = 9 - 12 - 5 = -8$$

There are occasions when the input number or value of the dependent variable is not admissible in the sense that the function fails to process it. For example, take the reciprocal function $f(x) = 1/x$ and consider the input value 0. If we try to evaluate $f(0)$ on a calculator, an error message will be given because we cannot divide by zero. Some calculators will even deliver a reprimand and inform you that you cannot divide by zero! All the numbers that a function can process are known collectively as the **domain** of the function.

Sometimes we may wish to restrict the domain to a smaller set of numbers than are admissible. In many applications in economics, we are only interested in domains that contain nonzero numbers. For example, the profit function is only of interest for non-negative values of output even though it may well be defined for negative values as well. The smaller set of numbers is called a **restricted domain**. For example, the function defined by

$$f(x) = 2x + 1, \quad -2 \leq x \leq 4, \tag{4.1}$$

has a domain restricted to all the real numbers lying between -2 and 4 even though this function is defined over all the real numbers. The **range** of a function is the collection of all those values of $f(x)$ that correspond to each and every number in the domain of the function. For example, the function $f(x) = x^2$ has a domain that consists of all the real numbers and a range that contains all the non-negative real numbers. The function $f(x)$ defined by (4.1) has domain $-2 \leq x \leq 4$ and range $-3 \leq f(x) \leq 9$.

Note that a function can take the same value for two different values of its argument. For example, the function $f(x) = x^2$ takes the value 4 when $x = -2$ and $x = 2$. Such functions are said to be **many-to-one**. Functions that are such that each element x of the domain is assigned to a different value $f(x)$ are said to be **one-to-one**, i.e., the function f is one-to-one if

$$f(x_1) = f(x_2) \text{ implies } x_1 = x_2.$$

Every linear function

$$f(x) = ax + b, \ a \neq 0,$$

is one-to-one. Relationships which are **one-to-many** can occur, but from our definition they are not functions. For example, $y^2 = 1 - x^2$ is an example of a one-to-many relationship. When $x = 0$, $y^2 = 1$, and so $y = -1$ and $y = 1$. Therefore, there are two values of y that correspond to $x = 0$.

4.2 Limits

Sometimes it is of interest to know how a function behaves as the value of its argument tends to a fixed value. For example, in economics one may wish to know how the average cost of producing a certain good decreases as the number of goods produced increases. For example, suppose that the total cost to an electronics company of producing Q flat screen televisions is

$$TC = 800Q + 1{,}000{,}000.$$

What is the average cost AC of producing Q televisions when Q is very large? We can answer questions such as this using the concept of a **limit**.

The limiting behaviour of a function when the values in its domain are larger than any finite number may be formalised by expressing the limit of a function $f(x)$ as x moves increasingly far to the right on the real line as

$$\lim_{x \to \infty} f(x).$$

So $x \to \infty$ means x increases without bound, and we say x tends to ∞. Similarly, the limit of $f(x)$ as x moves increasingly far to the left on the real line is expressed as

$$\lim_{x \to -\infty} f(x).$$

So $x \to -\infty$ means x decreases without bound, and we say x tends to $-\infty$. In the next section, the concept of the limit of a function will be explored and explained for the reciprocal function $f(x) = 1/x$.

4.3 Polynomial Functions

The properties of linear and quadratic functions were described in Chapters 2 and 3, respectively. In this section, we look at other polynomial functions. First of all, consider the power functions defined by

$$f(x) = x^n,$$

where n is a positive integer. These are sometimes known as **monomials** since they comprise only one term.

If n is even, the graph of $f(x) = x^n$ is similar to that of $f(x) = x^2$ in terms of its shape and its symmetry about the y-axis (see Fig. 4.1). The important difference is that, for $n > 2$, $f(x)$ increases more rapidly as x increases away from $x = \pm 1$ in the positive and negative x-directions. Note that all the graphs pass through the three points $(0, 0)$ (where they attain their minimum values), $(1, 1)$ and $(-1, 1)$.

If n is odd, the graphs of $f(x) = x^n$ are similar, for positive values of x, to those for which n is even. However, for negative values of x they are quite different (see Fig. 4.2). The portion of the graph for negative values of x may be formed as the result of two reflections of the positive portion of the graph, first with respect to the y-axis and then with respect to the x-axis, i.e., if the point (x, y) lies on the graph then so also does the point $(-x, -y)$. All the graphs pass through the points $(0, 0)$, $(1, 1)$, and $(-1, -1)$.

If n is odd, the function $f(x) = x^n$ is an **increasing** function of x since

$$f(x_1) \leq f(x_2) \text{ for } x_1 < x_2.$$

If n is even, the function $f(x) = x^n$ is an increasing function of x for $x \geq 0$. However, for $x \leq 0$, the function is **decreasing** since

$$f(x_1) \geq f(x_2) \text{ for } x_1 < x_2 \leq 0.$$

The general **cubic** function has the form

$$f(x) = ax^3 + bx^2 + cx + d,$$

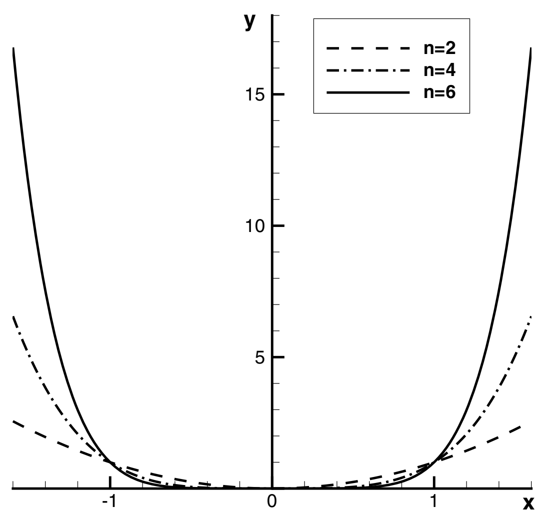

Figure 4.1 The graphs of the even monomials $f(x) = x^n$ for $n = 2, 4, 6$.

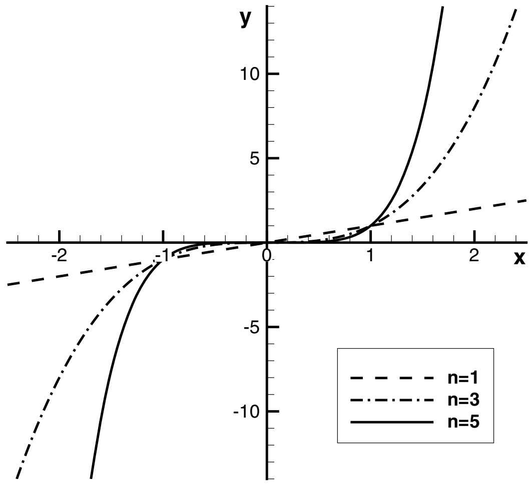

Figure 4.2 The graphs of the odd monomials $f(x) = x^n$ for $n = 1, 3, 5$.

with $a \neq 0$. The simplest cubic function is $f(x) = x^3$. Its graph is the green curve in Fig. 4.2. More generally, the graph of a cubic function has one of the two forms shown in Figs. 4.3 and 4.4 depending on the sign of a. If $a > 0$, $f(x)$ tends to ∞ as x tends to ∞ and tends to $-\infty$ as x tends to $-\infty$. The cubic function $f(x) = x^3 + x^2 - 2x$ has $a = 1 > 0$, and its graph is shown in Fig. 4.3. If $a < 0$, $f(x)$ tends to $-\infty$ as x tends to ∞ and tends to ∞ as x tends to $-\infty$. The cubic function $f(x) = -x^3 + 5x^2 - 2x - 15$ has $a = -1 < 0$, and its graph is shown in Fig. 4.4.

The graph of a cubic function crosses the x-axis at one, two, or three points. Therefore, the equation $f(x) = 0$ has one, two, or three real roots. For example, the graph of the cubic function $f(x) = x^3 + x^2 - 2x$ (see Fig. 4.3) crosses the x-axis when $x = -2$, $x = 0$ and $x = 1$, and the graph of the cubic function $f(x) = -x^3 + 5x^2 - 2x - 15$ (see Fig. 4.4) crosses the x-axis at the single point $x = -3/2$. The graph of the function $f(x) = x^3 - x^2$ crosses the x-axis at $x = 1$ and $x = 0$. At $x = 0$ the function $f(x) = x^3 - x^2$ has two coincident roots.

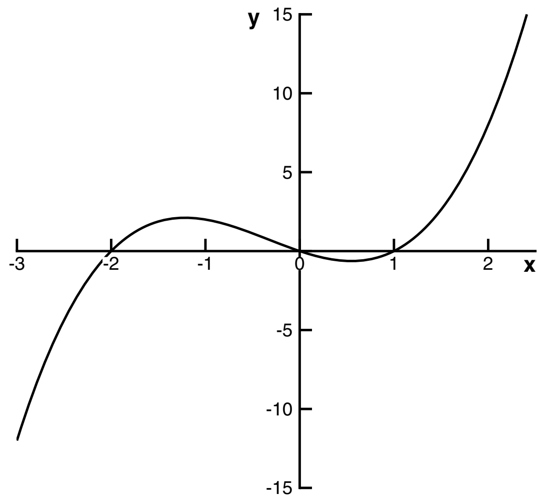

Figure 4.3 The graph of the function $f(x) = x^3 + x^2 - 2x$.

4.4 Reciprocal Functions

Consider the reciprocal function defined by

$$f(x) = \frac{1}{x},$$

for $x > 0$. All the applications considered in this book are for $x > 0$. However, for completeness we also sketch the function for $x < 0$ in Fig. 4.5. Here we see that the part of the graph for $x < 0$ is obtained by reflecting the graph for $x > 0$ in the line $y = -x$. As we have already noted, this function is not defined for $x = 0$.

When x is large and positive, $f(x)$ is small and positive, and as x takes increasingly larger values, $f(x)$ takes values that approach but never reach 0. For example, $f(10) = 0.1$, $f(100) = 0.01$, $f(1{,}000) = 0.001$, etc. As $x \to \infty$, the graph of $f(x)$ gets arbitrarily close to the x-axis and therefore

$$\lim_{x \to \infty} \frac{1}{x} = 0.$$

When x is large and negative, $f(x)$ is small and negative. As $x \to -\infty$, the graph of $f(x)$ gets arbitrarily close to the x-axis approaching it from below and

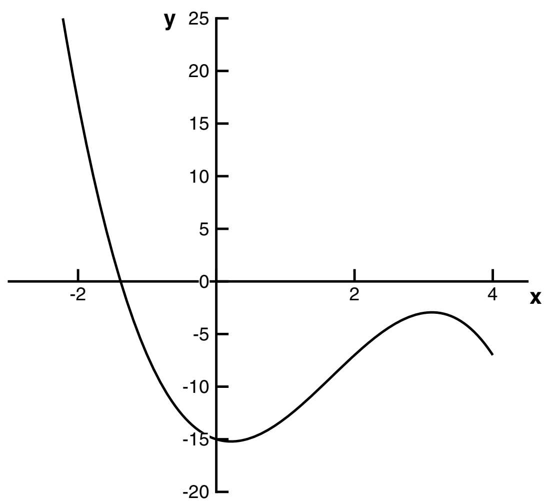

Figure 4.4 The graph of the function $f(x) = -x^3 + 5x^2 - 2x - 15$.

therefore

$$\lim_{x \to -\infty} \frac{1}{x} = 0.$$

The idea of a limit can also be used to describe the unbounded behaviour of functions. For example, consider the limit of the reciprocal function $f(x) = 1/x$ for $x \neq 0$ as x tends to 0 from the right (see Fig. 4.5), i.e., x takes only positive values. As x takes increasingly smaller values, $f(x)$ takes increasingly larger values. For example, $f(1) = 1$, $f(0.1) = 10$, $f(0.01) = 100$, $f(0.001) = 1{,}000$, etc. The values of f are positive and become arbitrarily large in this limit, i.e., given any positive number y we can always find a value of x for which $f(x) > y$. We express this mathematically as

$$\lim_{x \to 0^+} \frac{1}{x} = \infty.$$

The superscript '+' on 0 indicates that we are taking the limit as x approaches 0 from the right through positive values. Similarly, the values of $f(x)$ as x tends to 0 from the left are negative and become arbitrarily large in this limit, i.e., given any negative number z we can always find a value of x for which $f(x) < z$. We express this mathematically as

$$\lim_{x \to 0^-} \frac{1}{x} = -\infty.$$

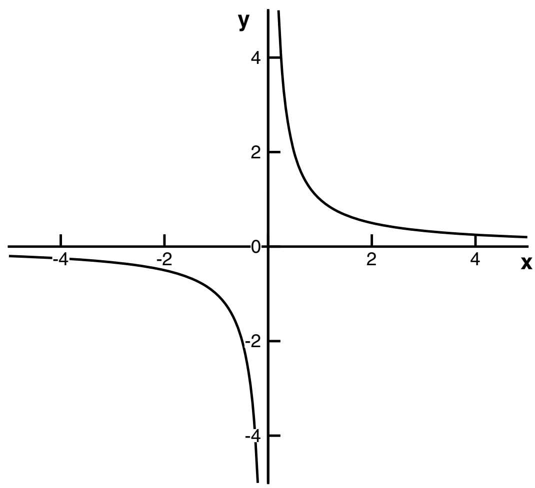

Figure 4.5 The graph of the function $f(x) = 1/x$.

The superscript '$-$' on 0 indicates that we are taking the limit as x approaches 0 from the left through negative values.

Example 4.4

The fixed costs of producing a good are 10 and the variable costs are 4 per unit. Find expressions for total cost TC and average cost AC. Sketch the graph of AC as a function of Q.

Solution. The total cost function is

$$\begin{aligned} TC &= FC + VC \times Q \\ &= 10 + 4Q. \end{aligned}$$

The average cost function, AC, is given by

$$AC = \frac{TC}{Q}.$$

Table 4.1 Tables of values of AC in Example 4.4.

Q	0.01	0.1	1	10	100
AC	1004	104	14	5	4.1

Therefore, using the above expression for TC we have

$$\begin{aligned} AC &= \frac{10+4Q}{Q} \\ &= \frac{10}{Q} + \frac{4Q}{Q} \\ &= \frac{10}{Q} + 4. \end{aligned}$$

This function is tabulated in Table 4.1 and sketched in Fig. 4.6. The dashed line in this figure corresponds to $VC = 4$. As Q tends to ∞, AC tends to 4, i.e.,

$$\lim_{Q\to\infty} AC = 4.$$

In this example, it is no coincidence that AC approaches the value of VC, i.e., 4, as Q becomes large. In fact, this result holds whenever VC is constant. To see this, let us examine the expression for AC:

$$\begin{aligned} AC &= \frac{TC}{Q} \\ &= \frac{FC + VC \times Q}{Q} \\ &= \frac{FC}{Q} + VC. \end{aligned}$$

As Q becomes large, FC/Q approaches 0. Therefore, AC tends to VC as Q tends to ∞, i.e.,

$$\lim_{Q\to\infty} AC = VC.$$

Example 4.5

The fixed costs of producing a good are 8, and the variable costs are $3 + 5Q$ per unit. Find expressions for total cost TC and average cost AC. Evaluate TC and AC when $Q = 10$. Sketch the graph of AC as a function of Q.

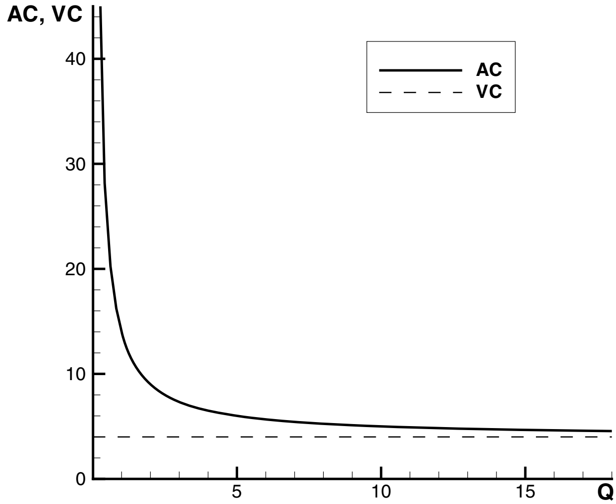

Figure 4.6 The graphs of the average cost function $AC = \frac{10}{Q} + 4$ and the variable cost per unit $VC = 4$.

Solution. The total cost function is

$$\begin{aligned} TC &= FC + VC \times Q \\ &= 8 + (3 + 5Q)Q \\ &= 8 + 3Q + 5Q^2. \end{aligned}$$

The average cost function, AC, is given by

$$AC = \frac{TC}{Q}.$$

Therefore, using the above expression for TC we have

$$\begin{aligned} AC &= \frac{8 + 3Q + 5Q^2}{Q} \\ &= \frac{8}{Q} + \frac{3Q}{Q} + \frac{5Q^2}{Q} \\ &= \frac{8}{Q} + 3 + 5Q. \end{aligned}$$

When $Q = 10$,

$$TC = 8 + 3 \times 10 + 5 \times 10^2 = 8 + 30 + 500 = 538,$$

Table 4.2 Tables of values of AC in Example 4.5.

Q	0.01	0.1	1	10	100
AC	803.05	83.5	16	53.8	503.08

and

$$AC = \frac{8}{10} + 3 + 5 \times 10 = 0.8 + 3 + 50 = 53.8.$$

This function is tabulated in Table 4.2 and sketched in Fig. 4.7. The dashed line in this figure is the straight line $AC = 3 + 5Q$. As Q tends to ∞, AC tends to VC. This is because the term $8/Q$ in the equation for AC becomes negligibly small for large values of Q. Since VC tends to ∞ as Q tends to ∞, we have

$$\lim_{Q \to \infty} AC = \infty.$$

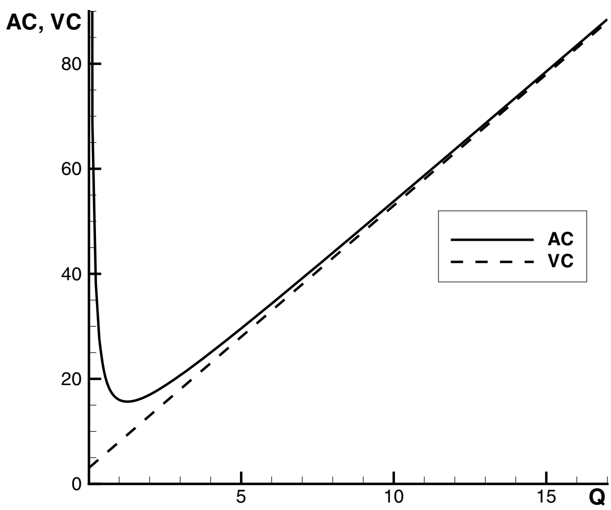

Figure 4.7 The graphs of the average cost function $AC = \frac{8}{Q} + 3 + 5Q$ and the variable cost per unit $VC = 3 + 5Q$.

Example 4.6

Suppose that the total cost to an electronics company of producing Q flat screen televisions is

$$TC = 800Q + 1{,}000{,}000.$$

Obtain an expression for the average cost function. What is the average cost of production when Q is very large?

Solution. The average cost function is given by

$$AC = \frac{TC}{Q} = \frac{800Q + 1{,}000{,}000}{Q} = 800 + \frac{1{,}000{,}000}{Q}.$$

The second term in this expression for the average cost function tends to 0 as Q tends to ∞. Therefore, in the limit of arbitrarily large Q we have

$$\lim_{Q\to\infty} AC = 800.$$

4.5 Inverse Functions

Given a function $y = f(x)$, consider the reverse process in which y becomes the input and x the output. This reverse process, under certain conditions, defines what is known as the **inverse** function of f. If we denote the inverse function by, say g, then we can write $x = g(y)$ (see Fig. 4.8). Thus, y is now the independent variable and x the dependent variable. For example, consider the determination of the inverse of the function $y = f(x) = 6x+2$. This is achieved by reversing the input and output processes of the function. The inverse of the function that multiplies the input by 6 and then adds 2 to the result is the process that subtracts 2 from the input and then divides the result by 6. This process defines the inverse of the function, i.e.,

$$x = g(y) = \frac{y-2}{6}.$$

The inverse of a one-to-one function satisfies the definition of a function and therefore is itself a function. Therefore, a necessary condition for a given function to have an inverse is that it is one-to-one. Thus every linear function $f(x) = ax + b,\ a \neq 0$, has an inverse since it is one-to-one.

Nonlinear functions may not possess an inverse function. For example, the function

$$y = f(x) = x^2,$$

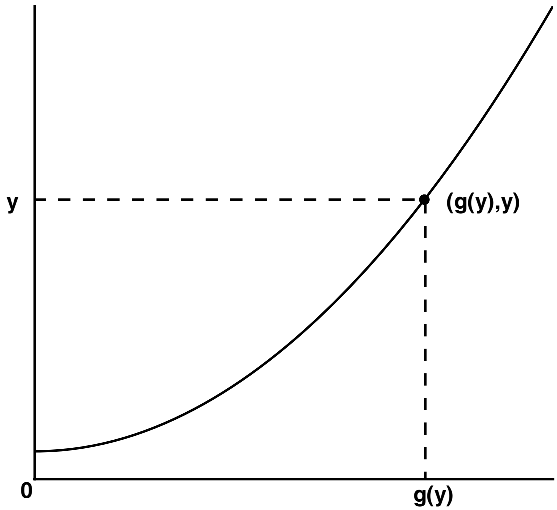

Figure 4.8 Graph of $y = f(x)$ where g is the inverse function of f.

is a many-to-one function, i.e., there are two values of x that correspond to each value of y (see Fig. 4.9 where $x = \pm 2$ both correspond to $y = 4$). If we tried to find the inverse of this many-to-one function, we would obtain a one-to-many relationship, which contravenes the definition of a function. Thus, only a one-to-one function can possess an inverse. However, if the domain of f is restricted to positive values of x, say, then f does possess an inverse defined by

$$x = g(y) = \sqrt{y}.$$

This situation is shown in Fig. 4.10.

Example 4.7

Find the inverses of the functions

1. $f(x) = 2x - 3$,
2. $f(x) = (x - 2)^2, \ 2 \leq x$.

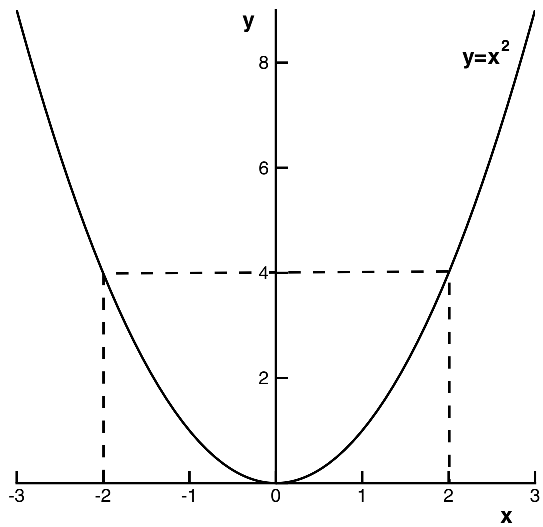

Figure 4.9 Graph showing the many-to-one function $y = x^2$.

Solution.

1. Let $y = 2x - 3$. We rearrange this equation so that x appears by itself on the left-hand side. Adding 3 to both sides, we have

$$y + 3 = 2x.$$

Finally, dividing both sides by 2 yields the inverse function

$$x = g(y) = \frac{y+3}{2}.$$

2. Let $y = (x-2)^2$. For $x \geq 2$ this function is one-to-one and therefore possesses an inverse. Taking the square root of both sides of this equation gives

$$x - 2 = \sqrt{y}.$$

Finally, adding 2 to both sides yields the inverse function

$$x = g(y) = \sqrt{y} + 2.$$

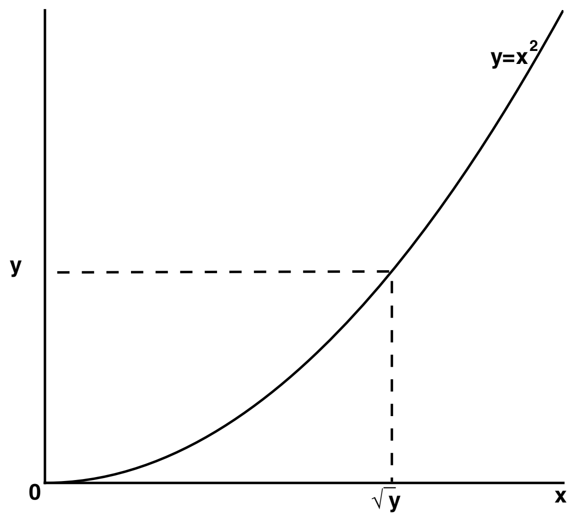

Figure 4.10 Graph showing the inverse function of $f(x) = x^2$ when the domain of f is restricted to positive values of x.

The motivation for introducing inverse functions in this book is that in economics, some functions are plotted with the dependent variable y on the horizontal axis and the independent variable x on the vertical axis. The demand function is one such example. The demand function expresses the dependence of the quantity demanded, Q, of a good on the market price, P. We may write this function as

$$Q = f(P).$$

Given a particular rule for f, it is relatively simple to determine the value of Q for a given value of P or to sketch the graph of the function. A mathematician would plot this function with the independent variable (P) on the horizontal axis and the dependent variable (Q) on the vertical axis. However, economists prefer to plot them the other way round with Q on the horizontal axis and P on the vertical axis. To facilitate this, the demand equation is rearranged so that P is expressed in terms of Q, i.e.,

$$P = g(Q),$$

for some function g. The functions f and g are said to be **inverse functions**.

Example 4.8

For the demand function $Q = f(P)$, where

$$f(P) = -\frac{P}{3} + 18$$

determine the value of Q when $P = 30$. Express P in terms of Q and hence find the value of P when $Q = 9$.

Solution.

$$Q = f(P) = -\frac{P}{3} + 18.$$

When $P = 30$,

$$Q = -\frac{30}{3} + 18 = -10 + 18 = 8.$$

To express P in terms of Q, we rearrange the terms to isolate P on the left-hand side of the equation. Multiplying both sides by 3 gives

$$3Q = -P + 54.$$

A simple rearrangement of this equation yields the following expression for P in terms of Q

$$P = -3Q + 54.$$

When $Q = 9$,

$$P = -3(9) + 54 = -27 + 54 = 27.$$

In determining P as a function of Q, we have found the inverse function of f. We may write

$$P = g(Q), \text{ where } g(Q) = -3Q + 54.$$

EXERCISES

4.1. Sketch the graph of the cubic function $f(x) = 6 + 12x + 3x^2 - 2x^3$ for $-2 \leq x \leq 3$.

4.2. Sketch the graph of the cubic function $f(x) = 8x^3 + 30x^2 + 13x - 15$ for $-4 \leq x \leq 2$.

4.3. The fixed costs of producing a good are 12 and the variable costs are 7 per unit. Find expressions for TC and AC. Evaluate TC and AC when $Q = 4$ and $Q = 12$. Sketch the graph of AC as a function of Q.

4.4. The fixed costs of producing a good are 9 and the variable costs are $4+3Q$ per unit. Find expressions for TC and AC. Evaluate TC and AC when $Q=5$ and $Q=10$. Sketch the graph of AC as a function of Q.

4.5. Suppose that the total cost to a furniture company of producing Q desks is

$$TC = 50Q + 40{,}000.$$

Obtain an expression for the average cost function, AC. What value does AC approach when Q is very large?

4.6. Find the inverses of the following functions:

a) $f(x) = -3x + 2$,

b) $f(x) = 5x + 3$,

c) $f(x) = (x-3)^2 + 2,\ 3 \leq x$.

4.7. For the demand function

$$Q = -\frac{P}{4} + 25$$

determine the value of Q when $P = 36$. Express P in terms of Q and hence find the value of P when $Q = 5$.

5
The Exponential and Logarithmic Functions

5.1 Introduction

An important class of nonlinear functions that is of particular interest in economics comprises the exponential and logarithmic functions. These functions are useful for investigating problems associated with economic growth and decay and mathematical problems in finance such as the compounding of interest on an investment or the depreciation of an asset. For example, if a person invests £3,000 in an investment bond for which there is a guaranteed annual rate of interest of 5% for two years, the evaluation of an exponential function will provide the return at the end of that period. If a credit card company charges interest on an outstanding balance, the evaluation of an exponential function will provide information on the AER (annual equivalent rate). We begin this chapter by sketching the graphs of some exponential functions and highlighting some of their important properties. Exponential functions are functions in which a constant base a is raised to a variable exponent x. The general form of an exponential function is given by

$$y = a^x, \quad \text{where } a > 0 \text{ and } a \neq 1. \tag{5.1}$$

The parameter a is known as the **base** of the exponential function. The independent variable x occurs as the **exponent** of the base.

5.2 Exponential Functions

All exponential functions of the form $f(x) = a^x$ satisfy the following properties:

Properties

1. The domain of $f(x)$ is the set of all real numbers; the range of $f(x)$ is the set of all positive real numbers.
2. For all $a > 1$, $f(x)$ is increasing; for $0 < a < 1$, $f(x)$ is decreasing.
3. For all $a > 0$ with $a \neq 1$, $f(0) = 1$.
4. For $a > 1$, $f(x)$ tends to 0 as x tends to $-\infty$; for $0 < a < 1$, $f(x)$ tends to 0 as x tends to $+\infty$.
5. For $a > 1$, $f(x)$ tends to $+\infty$ (i.e., increases without bound) as x tends to $+\infty$; for $0 < a < 1$, $f(x)$ tends to $+\infty$ as x tends to $-\infty$.

In Fig. 5.1, the graphs of $y = 2^x$ and $y = 2^{-x} = (\frac{1}{2})^x$ are sketched for $-4 \leq x \leq 4$. These graphs illustrate some of the properties of exponential functions. Clearly, the domain of both functions is the entire real line, and the range is the set of all positive real numbers. The graph of $f(x) = 2^x$ is strictly increasing and $f(x)$ tends to $+\infty$ as x tends to $+\infty$. The graph of $f(x) = 2^{-x}$ is strictly decreasing and $f(x)$ tends to 0 as x tends to $+\infty$. Note also that the graphs of $y = 2^x$ and $y = 2^{-x}$ are reflections of each other in the y-axis under the reflection $(x, y) \to (-x, y)$.

In Fig. 5.2, the graphs of two exponential functions with bases $a = 2$ and $a = 5$ are sketched. This figure shows that, for bases a_1, a_2 satisfying $a_2 > a_1 > 1$, a_2^x increases in value faster than a_1^x for $x > 0$.

An important base that is useful in many areas of mathematics as well as in applications to problems in economics is the irrational number e, whose most significant digits are given by

$$e = 2.7182818284\ldots.$$

This mathematical constant like the constant π does not have a finite decimal representation and is another example of an irrational number. Its decimal form is therefore never ending and is not a repeating decimal. It is interesting to see how this number can be defined without going into the mathematical details. Consider the function

$$f(x) = \left(1 + \frac{1}{x}\right)^x.$$

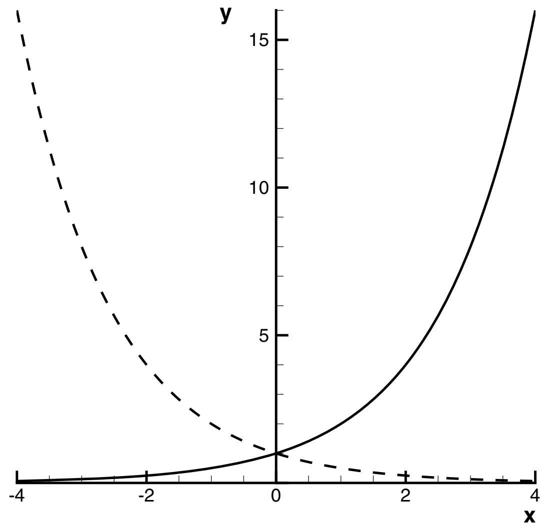

Figure 5.1 The graph of the functions $f(x) = 2^x$ (continuous curve) and $f(x) = 2^{-x}$ (dashed curve).

Let us evaluate this function for increasing values of x, for example $x = 1$, 10, 100, 1,000, and 10,000.

x	$f(x)$
1	$\left(1+\frac{1}{1}\right)^{1} = 2$
10	$\left(1+\frac{1}{10}\right)^{10} = 2.593742460$
100	$\left(1+\frac{1}{100}\right)^{100} = 2.704813829$
1,000	$\left(1+\frac{1}{1{,}000}\right)^{1000} = 2.716923932$
10,000	$\left(1+\frac{1}{10{,}000}\right)^{10000} = 2.71815$

These calculations show that as x gets larger, the value of

$$\left(1+\frac{1}{x}\right)^{x}$$

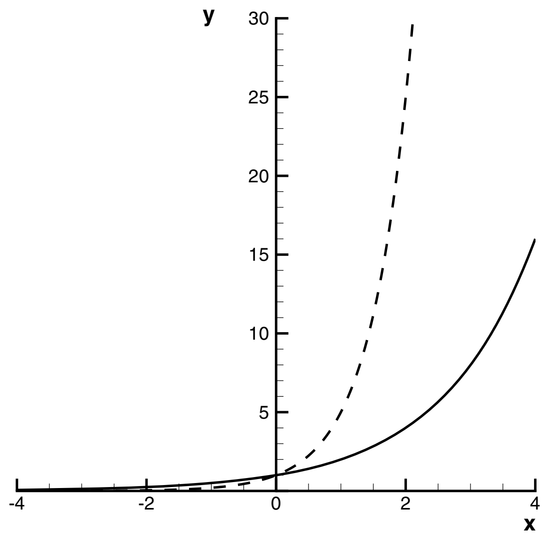

Figure 5.2 The graph of the functions $f(x) = 2^x$ (continuous curve) and $f(x) = 5^x$ (dashed curve).

increases and approaches a limiting value of 2.718281828..., which traditionally is denoted by the letter e. Mathematically, we write

$$e = \lim_{x \to \infty} \left(1 + \frac{1}{x}\right)^x,$$

i.e., as x approaches infinity, the value of the function $f(x) = \left(1 + \frac{1}{x}\right)^x$ approaches the constant e. The graph of this function is plotted in Fig. 5.3 where the dotted line corresponds to the straight line $y = e$. In this figure we see that, as x increases, $f(x)$ gradually approaches the dashed line.

5.3 Logarithmic Functions

Logarithms have inspired a feeling of dread in generations of students on their first encounter with them. Logarithmic functions are closely related to exponential functions and it is this relationship that we will exploit in our description of some of their key properties. Logarithms are useful for simplifying calculations involving economic functions. If we take the exponential function defined by

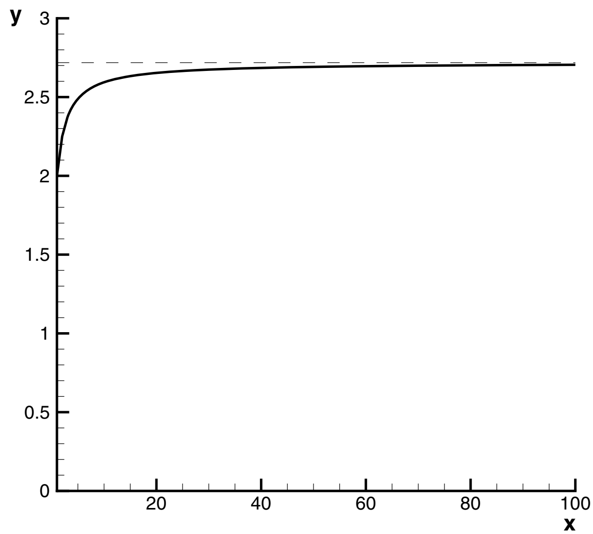

Figure 5.3 The graph of the function $f(x) = \left(1 + \frac{1}{x}\right)^x$. The dashed line corresponds to the constant function $f(x) = e$.

$y = a^x$ and interchange the dependent variable y with the independent variable x, we obtain

$$x = a^y.$$

This defines a new function $y = \log_a x$, known as the **logarithmic function** with base a, which is the exponent to which a must be raised to get x, i.e.,

$$x = a^y \quad \Leftrightarrow \quad y = \log_a x.$$

Thus, the logarithmic function $y = \log_a x$ is the inverse of the exponential function $y = a^x$. For example, if we wish to evaluate $y = \log_{10} 100$, then $100 = 10^y$. Since $100 = 10^2$, we find that $y = 2$ so that $\log_{10} 100 = 2$. The restrictions on the base are the same as for the exponential functions, i.e., $a > 0,\ a \neq 1$.

There are two important bases:

- $a = 10$ gives rise to **common logarithms**, written simply as $\log x$.
- $a = e$ where $e \approx 2.71828$ gives rise to **natural logarithms**, written as $\ln x$.

Common and natural logarithms may be evaluated numerically by pressing either the log or ln keys, respectively, on a scientific calculator. For example, to evaluate $\log 2.5$:

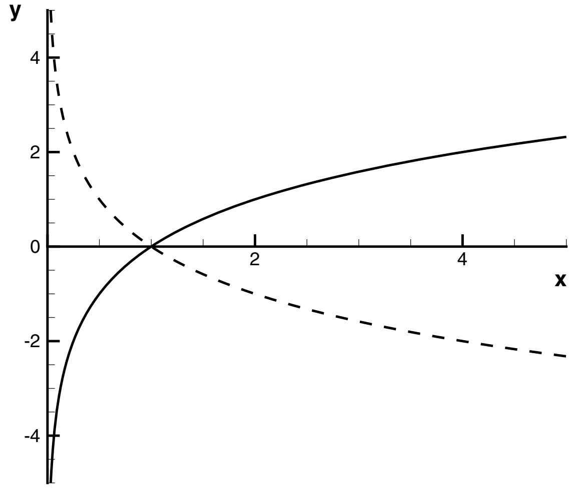

Figure 5.4 The graph of the functions $f(x) = \log_2 x$ (continuous curve) and $f(x) = \log_{1/2} x$ (dashed curve).

1. Enter 2.5
2. Press the [log] key

You should obtain the answer 0.397940009 to 9 decimal places, i.e., $\log 2.5 = 0.397940009$. Note that on some calculators you press the [log] key first, then enter the number and finally press the [=] key. Similarly, to evaluate the natural logarithm of 2.5:

1. Enter 2.5
2. Press the [ln] key

In this case, you should obtain the answer 0.916290732 to 9 decimal places, i.e., $\ln 2.5 = 0.916290732$.

Properties of the function $f(x) = \log_a x$

1. The domain of the function is the set of all positive real numbers; the range is the set of all real numbers.
2. For base $a > 1$, $f(x)$ is increasing. For $0 < a < 1$, $f(x)$ is decreasing.

Table 5.1 Table of values of $\log_2 x$ and $\log_{1/2} x$.

x	$1/4$	$1/2$	1	2	4
$\log_2 x$	-2	-1	0	1	2
$\log_{1/2} x$	2	1	0	-1	-2

3. At $x = 1$, $y = 0$ independent of the base.

The graphs of the logarithmic functions $y = \log_2 x$ and $y = \log_{1/2} x$ are shown in Fig. 5.4. These logarithmic functions may be written equivalently as $x = 2^y$ and $x = (1/2)^y$, respectively, and are tabulated in Table 5.1. Note that these graphs are reflections of the graphs of $y = 2^x$ and $y = 2^{-x}$, respectively, in the line $y = x$.

Example 5.1

Evaluate the following:

1. $\log_8 64$,
2. $\log_3(\frac{1}{81})$,
3. $\log_{16} 2$.

Solution.

1. Let $y = \log_8 64$, then $8^y = 64 = 8^2$ and so $y = 2$.
2. Let $y = \log_3 \frac{1}{81}$, then $3^y = \frac{1}{81} = \frac{1}{3^4} = 3^{-4}$ and so $y = -4$.
3. Let $y = \log_{16} 2$, then $16^y = 2$ or $(2^4)^y = 2^{4y} = 2$ and so $4y = 1$ and therefore $y = \frac{1}{4}$.

Example 5.2

Solve the following for x:

1. $\log_4 x = 3$,
2. $\log_{81} x = \frac{3}{4}$.

Solution.

1. $x = 4^3 = 64$.

2. $x = 81^{3/4} = (81^{1/4})^3 = 3^3 = 27.$

Rules of Logarithms

For a, x, y positive real numbers, and n a real number, and base $a \neq 1$:

$$\log_a(xy) = \log_a x + \log_a y, \tag{5.2}$$
$$\log_a(x/y) = \log_a x - \log_a y, \tag{5.3}$$
$$\log_a x^n = n \log_a x, \tag{5.4}$$
$$\log_a \sqrt[n]{x} = \log_a x^{1/n} = \frac{1}{n} \log_a x. \tag{5.5}$$

Note that $\log_a x^2$ means the logarithm of x^2 and not the square of $\log_a x$, which is written as $\log^2 x$.

To prove the first two rules, let $s = \log_a x$ and $t = \log_a y$. Using the relationship between the logarithmic and exponential functions, we have $x = a^s$ and $y = a^t$. Then using the product rule for exponents, we obtain

$$xy = a^s a^t = a^{s+t}.$$

So $s + t$ is the power to which the base must be raised to give xy, i.e.,

$$s + t = \log_a x + \log_a y = \log_a(xy).$$

Similarly, using the quotient rule for exponents, we have

$$\frac{x}{y} = \frac{a^s}{a^t} = a^{s-t}.$$

So $s - t$ is the power to which the base must be raised to give x/y, i.e.,

$$s - t = \log_a x - \log_a y = \log_a(x/y).$$

Example 5.3

Solve the equation $\ln(x + 4)^2 = 3$ for x.

Solution.

$$\begin{aligned} 2\ln(x+4) &= 3 \\ \ln(x+4) &= \frac{3}{2} \\ x+4 &= e^{1.5} \\ x+4 &= 4.48169 \text{ to 5 decimal places} \\ x &= 0.48169 \text{ to 5 decimal places} \end{aligned}$$

Example 5.4

Express $\log_a 3 + \log_a 4 - \log_a 6$ as a single logarithm.

Solution.

$$\begin{aligned}
\log_a 3 + \log_a 4 - \log_a 6 &= \log_a(3 \times 4) - \log_a 6 \\
&= \log_a \left(\frac{3 \times 4}{6} \right) \\
&= \log_a 2
\end{aligned}$$

Example 5.5

Find the value of x satisfying

$$\log_a x = 3 \log_a 2 + \log_a 20 - \log_a 1.6.$$

Solution.

$$\begin{aligned}
\log_a x &= 3 \log_a 2 + \log_a 20 - \log_a 1.6 \\
&= \log_a 2^3 + \log_a 20 - \log_a 1.6 \\
&= \log_a \left(\frac{8 \times 20}{1.6} \right) \\
&= \log_a 100
\end{aligned}$$

Therefore $x = 100$.

5.4 Returns to Scale of Production Functions

The output, Q, of any production process depends on a variety of inputs, known as **factors of production**. These include land, capital, labour, and enterprise. For simplicity, here we restrict our attention to capital, K, and labour, L. The dependence of Q on K and L is indicated by writing

$$Q = Q(K, L).$$

Q is called a **production function**. It is an example of a function of two variables – in this case K and L. Functions of two variables are described in more detail in Chapter 8.

If $Q(K, L) = 100K^{1/3}L^{1/2}$, then when $K = 27$ and $L = 100$ the output $Q(27, 100)$ is given by

$$\begin{aligned} Q &= 100(27)^{1/3}(100)^{1/2} \\ &= 100(3)(10) \\ &= 3{,}000 \end{aligned}$$

Of particular interest is what happens to the output when the inputs are scaled in some way. If capital and labour double, does the production level double, does it go up by more than double, or does it go up by less than double? For the above production function, we see that when K and L are replaced by $2K$ and $2L$, respectively, then using the rules of indices (see Section 1.6):

$$\begin{aligned} Q &= 100(2K)^{1/3}(2L)^{1/2} \\ &= 100(2^{1/3}K^{1/3})(2^{1/2}L^{1/2}) \\ &= (2^{1/3}2^{1/2})(100K^{1/3}L^{1/2}) \\ &= 2^{5/6}(100K^{1/3}L^{1/2}) \end{aligned}$$

The term in brackets is just the original output. So this is multiplied by $2^{5/6} \approx 1.78$ so output goes up by just less than double when capital and labour are doubled.

In general, a function

$$Q = Q(K, L)$$

is said to be **homogeneous** if

$$Q(\lambda K, \lambda L) = \lambda^n Q(K, L), \tag{5.6}$$

for some number n where λ is a general number. The power, n, is called the **degree of homogeneity**. Let us take the previous example again:

$$\begin{aligned} Q(\lambda K, \lambda L) &= 100(\lambda K)^{1/3}(\lambda L)^{1/2} \\ &= (\lambda^{1/3}\lambda^{1/2})100K^{1/3}L^{1/2} \\ &= \lambda^{5/6}Q(K, L) \end{aligned}$$

This production function is homogeneous of degree 5/6.

In general, if the degree of homogeneity, n, satisfies

1. $n < 1$ the function is said to display **decreasing returns to scale**.

2. $n = 1$ the function is said to display **constant returns to scale**.

3. $n > 1$ the function is said to display **increasing returns to scale**.

5.4.1 Cobb-Douglas Production Functions

Functions of the form

$$Q = AK^{\alpha}L^{\beta}$$

where A, α, β are constants are called **Cobb-Douglas production functions**. These are homogeneous of degree $\alpha + \beta$ since if

$$Q(K, L) = AK^{\alpha}L^{\beta}$$

then

$$\begin{aligned} Q(\lambda K, \lambda L) &= A(\lambda K)^{\alpha}(\lambda L)^{\beta} \\ &= \lambda^{\alpha+\beta}(AK^{\alpha}L^{\beta}) \\ &= \lambda^{\alpha+\beta}Q(K, L) \end{aligned}$$

Therefore, Cobb-Douglas production functions exhibit

1. decreasing returns to scale if $\alpha + \beta < 1$.
2. constant returns to scale if $\alpha + \beta = 1$.
3. increasing returns to scale if $\alpha + \beta > 1$.

Example 5.6

Show that the production function

$$Q = K^2 + 3KL,$$

is homogeneous and comment on its returns to scale.

Solution. In this example we are given that

$$Q = Q(K, L) = K^2 + 3KL.$$

If we scale or multiply both K and L by λ, then the corresponding value of output is

$$\begin{aligned} Q(\lambda K, \lambda L) &= (\lambda K)^2 + 3(\lambda K)(\lambda L) \\ &= \lambda^2 K^2 + 3\lambda^2 KL \\ &= \lambda^2(K^2 + 3KL) \\ &= \lambda^2 Q(K, L). \end{aligned}$$

Therefore, we have shown that

$$Q(\lambda K, \lambda L) = \lambda^2 Q(K, L),$$

which on comparison with (5.6) demonstrates that the production function is homogeneous of degree 2. Since the degree of homogeneity is greater than one, the function displays increasing returns to scale.

The Cobb-Douglas production function is an example of a nonlinear function. However, it may be converted to a linear function through a simple logarithmic transformation as follows. Take natural logarithms of both sides of the equation

$$Q = AK^{\alpha}L^{\beta}.$$

Then

$$\begin{aligned}\ln Q &= \ln(AK^{\alpha}L^{\beta})\\ &= \ln A + \ln K^{\alpha} + \ln L^{\beta}\\ &= \ln A + \alpha \ln K + \beta \ln L.\end{aligned}$$

This is what we call a log-linear function. If we define $\tilde{Q} = \ln Q$, $\tilde{K} = \ln K$, and $\tilde{L} = \ln L$, then

$$\tilde{Q} = \ln A + \alpha\tilde{K} + \beta\tilde{L},$$

a linear function in the variables $\tilde{K}$ and $\tilde{L}$.

5.5 Compounding of Interest

There is a plethora of investment products and loan facilities available to an individual in the financial market place. It is important for both an individual or a business to make an informed choice between the financial products on offer in order to maximize the return on their capital or to minimize the interest on their loan repayments, for example. Suppose that a person wants to borrow some capital and is offered two loan products. The first charges interest on the loan at the annual rate of 12% while the second charges interest at a monthly rate of 1%. Which product should the person go for? In this section, we show how such decisions can be made.

Suppose that an individual wishes to invest a sum of £10,000 over a period of three years and that the annual rate of interest is 5%. After one year, the interest on the investment amounts to 5% of £10,000, which is £500. If the investment is subject to **simple interest**, then the return on the investment would be £500 per year for each subsequent year. The total amount of interest earned over the five-year period in this case is 5 × £500 = £2,500. However, most financial investment products use **compound interest** as a means of enticing their customers not to withdraw the interest earned after the first and subsequent years from the accumulated value of their investment. When interest is compounded annually, the amount of interest earned in the second year is 5% of £10,500, which is the sum of the initial investment (£10,000) and the first year's interest (£500). The interest earned in the second year is therefore £525 and so the value of the investment at the end of the second year is £10,500 + £525 = £11,025. Finally, at the end of the third year the investment is worth £11,025 plus 5% of £11,025 interest giving a total of £11,571.25.

There is a formula that can be used to determine the future value of an investment. Let P_0 denote the value of the initial investment. This is sometimes known as the **principal**. Let P_t denote the value of the investment after t years. If the interest on the principal is compounded annually, at an interest rate r (written as a decimal or fraction), then after one year the investment is worth

$$P_1 = P_0 + rP_0 = P_0(1 + r). \tag{5.7}$$

Similarly, after the second year the investment is worth

$$P_2 = P_1 + rP_1 = P_1(1 + r). \tag{5.8}$$

Substituting for P_1 in (5.8) using (11.1) we have

$$P_2 = [P_0(1 + r)](1 + r) = P_0(1 + r)^2. \tag{5.9}$$

In general, one can show that

$$P_t = P_0(1 + r)^t. \tag{5.10}$$

Now suppose that the interest is compounded semi-annually (six monthly intervals). In this case, (5.10) would have to be modified to

$$P_t = P_0\left(1 + \frac{r}{2}\right)^{2t}. \tag{5.11}$$

Note that the differences between this formula and the formula (5.10) when the interest is compounded annually are that the interest rate is divided by 2 since the interest is added twice a year and t is replaced by $2t$ since this is the number of times that the interest is added during t years. Similarly, one can show that if interest is added monthly, the value of the investment after t years is

$$P_t = P_0\left(1 + \frac{r}{12}\right)^{12t}. \tag{5.12}$$

If this argument is continued and interest is compounded n times a year, then we have the formula

$$P_t = P_0\left(1 + \frac{r}{n}\right)^{nt}. \tag{5.13}$$

If n is very large then we are approaching the situation in which interest is added continuously (at every instant of time) instead of at discrete moments in time. If we make the substitution $m = n/r$ in (5.13), then we have

$$\begin{aligned} P_t &= P_0 \left(1 + \frac{1}{m}\right)^{mrt} \\ &= P_0 \left[\left(1 + \frac{1}{m}\right)^{m}\right]^{rt} \end{aligned} \tag{5.14}$$

We saw in Section 5.2 that

$$\left(1 + \frac{1}{m}\right)^{m} \to e \text{ as } m \to \infty.$$

If we allow $m \to \infty$ in (5.14) (which is equivalent to allowing $n \to \infty$ in (5.13) since r is held constant), then we obtain the formula for the continuous compounding of interest:

$$P(t) = P_0 e^{rt}. \tag{5.15}$$

In this formula, t need no longer be a positive integer. It can take any positive value.

For negative growth rates, such as depreciation or deflation, the same formulae apply but with t or r negative.

Example 5.7

Suppose that the sum of €100 is invested at an annual rate of interest of 10%. Calculate the value of the investment in five years' time if the interest is compounded (a) annually, (b) semi-annually, (c) continuously.

Solution.

1. We apply the formula (5.10) with $P_0 = 100$, $r = 10\% = 0.1$ and $t = 5$. Inserting these values into the formula gives

$$P_5 = 100(1 + 0.10)^5 = €161.05.$$

2. We apply the formula (5.11) with $P_0 = 100$, $r = 10\% = 0.1$ and $t = 5$. Inserting these values into the formula gives

$$P_5 = 100\left(1 + \frac{0.10}{2}\right)^{2\times 5} = 100(1.05)^{10} = €162.89.$$

3. We apply the formula (5.15) with $P_0 = 100$, $r = 10\% = 0.1$ and $t = 5$. Inserting these values into the formula gives

$$S = 100e^{0.10 \times 5} = 100e^{0.2} = €164.87.$$

Example 5.8

The value of an asset, currently priced at \$250,000, is expected to increase by 12% a year.

1. Find its value in ten years' time.
2. After how many years will it be worth at least 1.25 million dollars?

Solution.

1. We use the formula (5.15) with $P_0 = 250,000$, $r = 12\% = 0.12$, and $t = 10$. Inserting these values into the formula yields

$$\begin{aligned} P_{10} &= 250{,}000(1+0.12)^{10} \\ &= 250{,}000(1.12)^{10} \\ &= \$776{,}462.05 \end{aligned}$$

   Therefore, after 10 years the asset will be worth \$776,462.05.

2. In this part of the question, we use the formula (5.10) again but this time we know $P_t = 1{,}250{,}000$ and we need to determine the value of t. We need to find the value of t for which

$$\begin{aligned} 1{,}250{,}000 &= 250{,}000(1+0.12)^t \\ 5 &= (1.12)^t \end{aligned}$$

   Take natural logarithms of both sides:

$$\ln 5 = \ln(1.12)^t = t \ln 1.12.$$

   Therefore,

$$t = \ln 5 / \ln 1.12 = 14.20$$

   So after 15 years, the asset will be worth at least 1.25 million dollars.

Example 5.9

A credit card company charges interest at 2% per month. What is the annual equivalent rate correct to two decimal places?

Solution. Suppose that the balance outstanding on the credit card is B then the amount owing (loan plus interest) over one year is

$$B\left(1+\frac{2}{100}\right)^{12} = B(1.02)^{12}.$$

Here we have used the formula (5.12) but in which we have not divided r by 12 since the rate of interest is already a monthly one. Let R be the annual equivalent rate. Then if interest is charged annually, the amount owing after a year is

$$B\left(1+\frac{R}{100}\right).$$

Equating these two expressions enables us to find R:

$$B(1.02)^{12} = B\left(1+\frac{R}{100}\right).$$

Therefore,

$$(1.02)^{12} = 1+\frac{R}{100}.$$

Rearranging this equation gives

$$R = [(1.02)^{12} - 1] \times 100 = 0.2682 = 26.82\%.$$

This is the annual equivalent rate.

5.6 Applications of the Exponential Function in Economic Modelling

Example 5.10

During a recession, a firm's revenue declines continuously at an annual rate of 10% so that total revenue (measured in millions of pounds) in t years' time is modelled by

$$TR = 8e^{-0.1t}.$$

1. Calculate the current revenue and also the revenue in two years' time.
2. Sketch the graph of TR against t.
3. Rearrange the formula to get t in terms of TR.
4. After how many years will revenue decline to below £5 million?

Solution.

1. When $t = 0$, $TR = 8e^0 = 8$. Therefore, the current revenue is £8 million. When $t = 2$, $TR = 8e^{-(0.1)(2)} = 8e^{-0.2} = 6.55$. Therefore, after two years the revenue will have declined to £6.55 million.

2. The graph of the revenue function TR plotted as a function of time t is shown in Fig. 5.5.

3. The first step in the process of rearranging the formula for TR is to divide both sides by 8:
$$\frac{TR}{8} = e^{-0.1t}.$$
Then taking natural logarithms of both sides:
$$\ln\left(\frac{TR}{8}\right) = \ln e^{-0.1t} = -0.1t.$$
Finally, dividing both sides by -0.1, we obtain the formula for t in terms of TR:
$$t = \frac{1}{(-0.1)} \ln\left(\frac{TR}{8}\right) = -10 \ln\left(\frac{TR}{8}\right).$$

4. We now use this formula to determine the number of years after which the revenue will decline to £5 million. Inserting $TR = 5$ in this formula yields
$$t = -10 \ln\left(\frac{5}{8}\right) = 4.700 \text{ (to 3 decimal places)}.$$
Therefore, after 5 years TR will decline to £5 million. An estimate for this answer can be found from the graph in Fig. 5.5. The dashed line in this graph corresponds to $TR = 5$. The intersection of this straight line with the curve $TR = 8e^{-0.1t}$ provides the answer.

Example 5.11

The percentage, y, of Europeans possessing a mobile phone t years after it was introduced is modelled by

$$y = 80 - 70e^{-0.2t}.$$

1. Find the percentage of Europeans that have mobile phones
   a) at the launch of the product;
   b) after 3 years;
   c) after 10 years.

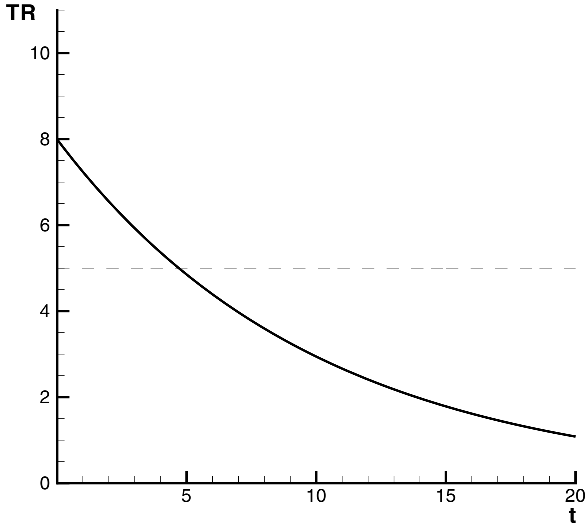

Figure 5.5 The graph of the function $TR = 8e^{-0.1t}$. The dashed line corresponds to $TR = 5$.

2. What is the market saturation level?
3. After how many years will the percentage of Europeans possessing mobile phones first reach 75%?

Solution.

1. a) The launch of the product corresponds to $t = 0$ since t measures the time from the introduction of mobile phones into the market place. So putting $t = 0$ into the expression for y gives

$$y = 80 - 70e^{0} = 80 - 70 = 10\%.$$

b) After three years $t = 3$, the percentage of Europeans possessing mobile phones is given by

$$y = 80 - 70e^{-0.2\times 3} = 80 - 70e^{-0.6} = 41.58\%.$$

c) After ten years $t = 10$, the percentage of Europeans possessing mobile phones is given by

$$y = 80 - 70e^{-0.2\times 10} = 80 - 70e^{-2} = 70.53\%.$$

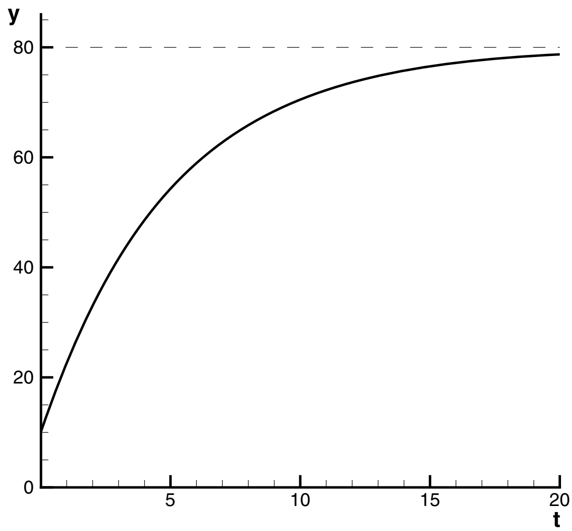

Figure 5.6 The graph of the function $y = 80 - 70e^{-0.2t}$.

2. The market saturation is the limiting value of y as t tends to ∞. Since

$$e^{-0.2t} \to 0 \text{ as } t \to \infty,$$

we have

$$y \to 80 \text{ as } t \to \infty,$$

and so the market saturation level is 80% (see Fig. 5.6).

3. To determine the time after which 75% of Europeans possess a mobile phone, we rearrange the equation and express t in terms of y and then put $y = 75$ into the resulting formula. A simple rearrangement gives

$$e^{-0.2t} = \frac{(80 - y)}{70}.$$

Taking natural logarithms of both sides yields

$$-0.2t = \ln\left(\frac{80 - y}{70}\right).$$

Finally, multiplying both sides by $1/(-0.2) = -5$ gives

$$
\begin{aligned}
t &= -5\ln\left(\frac{80-y}{70}\right) \\
&= -5\ln\left(\frac{5}{70}\right) \\
&= 13.20
\end{aligned}
$$

Therefore, after 14 years the percentage of Europeans possessing mobile phones will break through the 75% barrier.

EXERCISES

5.1. Sketch the functions $y = 2^x$ and $y = 3^x$ on the same graph for $-2 \leq x \leq 2$.

5.2. Evaluate (a) $\log_3 9$, (b) $\log_4 2$, (c) $\log_7(1/7)$.

5.3. Show that the following production functions are homogeneous and comment on their returns to scale:

a) $Q = 7KL^2$,

b) $Q = 50K^{1/4}L^{3/4}$.

5.4. Determine the annual rate of interest required for a principal of £2,000 to produce a value of £10,000 after 8 years.

5.5. Determine the annual equivalent rate (AER) corresponding to a monthly rate of 1.15%.

5.6. An economy is forecast to grow continuously at an annual rate of 3% so that the gross national product (GNP), measured in billions of euros, after t years is given by

$$GNP = 60e^{0.03t}.$$

a) Calculate the current value of GNP and its future value in four years' time.

b) After how many years is GNP forecast to be 90 billion euros?

5.7. Determine the rate of interest required for an investment that is currently worth \$1,000 to be worth \$4,000 after 10 years if the interest is compounded continuously.

5.8. Determine the annual equivalent rate (AER) corresponding to a monthly rate of 1%.

5.9. The percentage, y, of households possessing dishwashers t years after they have been introduced in a country is modelled by

$$y = 30 - 25e^{-0.2t}.$$

a) Find the percentage of households that have dishwashers

i. at their launch;

ii. after 1 year;

iii. after 10 years;

iv. after 20 years.

b) What is the market saturation level?

c) After how many years will the percentage of households possessing dishwashers first reach 15%?

5.10. A firm's turnover, y, measured in millions of pounds, after t years is given by

$$y = 8e^{0.09t}.$$

What is its turnover in its initial year of trading and after two years of trading? After how many years will its turnover have doubled since it started trading?

6
Differentiation

6.1 Introduction

Economists are interested in the effects of change. Therefore, the concept of the derivative of a function, which provides information about how a function changes in response to changes in the independent variable, is an important one in economic analysis. For example, the derivative of a production function provides information about the manner in which the output of a production process changes as the number of workers employed by the company changes. Differentiation is the mathematical tool that allows us to quantify such rates of change. As we will see in Chapter 7, differentiation is also an important tool in the determination of the maximum or minimum values of economic functions such as profit and cost.

In Chapter 2, some linear functions in economics such as linear demand and supply functions were introduced. These functions are characterized by their graphs being lines having a constant slope, i.e., the gradient is constant irrespective of the value of the independent variable. We say that the rate of change of the function is independent of the point where it is measured. Furthermore, the slope or gradient of a linear function may be determined by taking any two points on the straight line and calculating the ratio of the change in the vertical direction with respect to the change in the horizontal direction as the value of the independent variable increases. The corresponding situation for a nonlinear function is quite different, however, and the rate of change of a nonlinear function varies as one moves along the curve given by its

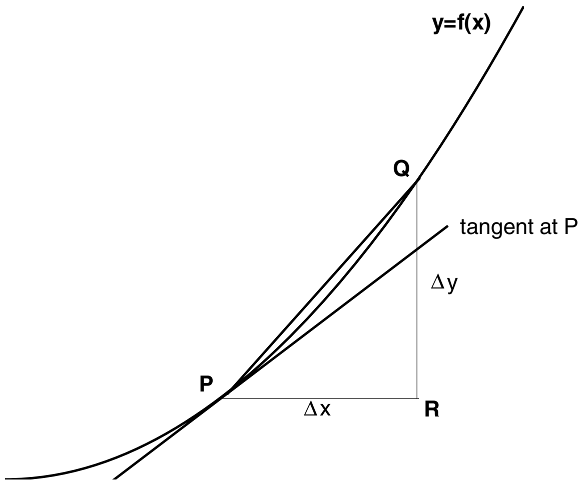

Figure 6.1 The graph of a nonlinear function in which the tangent at the point P is drawn.

graph.

In Fig. 6.1, we show part of the graph of a nonlinear function $y = f(x)$. On this graph, we have drawn the tangent to the curve at the point P. The **tangent** to a curve at a point P is the straight line that passes through P and that just touches the curve at this point. The **slope** or **gradient** of a curve $y = f(x)$ at P is then defined to be the gradient of the tangent to the curve at P. It is a measure of the prevailing rate of change of y relative to x at P. We can see from Fig. 6.1 that the gradient of a nonlinear function varies as we move along the curve.

In mathematics, we use the notation $f'(a)$ (pronounced f primed of a) to represent the slope of the tangent to the function f at $x = a$. The slope of the tangent to a function is called the **derivative** of the function – corresponding to each value of x there is a uniquely defined derivative $f'(x)$. Therefore, the derivative of a function of x is also a function of x.

If $y = f(x)$, then an alternative notation for the derivative of a function is

$$\frac{dy}{dx}.$$

This is pronounced 'dee y by dee x'. Note that this is a single entity not to be manipulated in any sense and represents the derivative of y with respect to x. If, for example, $f(x) = x^2$, then it is natural to use $f'(x)$ to represent the derivative of $f(x)$, whereas if $y = x^2$ is used then dy/dx is more appropriate.

Consider the function $y = f(x)$. The graph of this function is shown in Fig. 6.1. The slope or gradient of the function at the point $P : (x, f(x))$ is the slope of the tangent to the graph of the function at P (see Fig. 6.1). This slope can be approximated by the slope of the chord PQ where Q is the point $(x + \Delta x, f(x + \Delta x))$. (A **chord** is a straight line joining any two points on a curve.) So the horizontal distance from P to Q is Δx. If x is a variable, the notation Δx will denote a small change in x. Therefore,

$$\begin{aligned}\text{the slope of } PQ &= \frac{QR}{PR} \\ &= \frac{f(x+\Delta x) - f(x)}{(x+\Delta x) - x} \\ &= \frac{f(x+\Delta x) - f(x)}{\Delta x}\end{aligned}$$

If Q is allowed to approach P in which case Δx approaches 0, the slope of the chord PQ approaches the slope of the tangent at P, i.e.,

$$\text{Slope of the Tangent at } P = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}. \tag{6.1}$$

The value of this limit, if it exists, is known as the **derivative** of the function f at x and is written $f'(x)$ or dy/dx. Thus, we have

$$f'(x) = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x}. \tag{6.2}$$

So the derivative of a function at a point is the ratio of the change in y to the change in x between the point and a point that it infinitesimally close to it. So the derivative measures the instantaneous rate of change of the function.

If Δy denotes the change y corresponding to the change Δx in x, then

$$\Delta y = f(x + \Delta x) - f(x),$$

with $f(x + \Delta x)$ being the value of $y = f(x)$ when the value of x changes to $x + \Delta x$. Therefore,

$$f'(x) \left(\text{or } \frac{dy}{dx}\right) = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}. \tag{6.3}$$

Therefore, for a small change Δx in x and corresponding small change Δy in y, we have that

$$\frac{dy}{dx} \approx \frac{\Delta y}{\Delta x},$$

or

$$\Delta y \approx \Delta x \left(\frac{dy}{dx}\right). \tag{6.4}$$

This makes sense if dy/dx is regarded as the rate of change of y relative to x. In particular, dy/dx can be regarded as approximately the change in y resulting from a 1 unit increase in x (provided the value of x is relatively large so that 1 unit is relatively small). The approximation (6.4) is known as the **small increments formula**.

The process of finding the derivative of a function is known as **differentiation**. The definition of a function may be used to determine the derivative of a given function. This process is known as differentiation from first principles. For example, if $f(x) = x^2$, then using (6.2) we have

$$\begin{aligned}
f'(x) &= \lim_{\Delta x \to 0} \frac{(x+\Delta x)^2 - x^2}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{(x^2 + 2x\Delta x + (\Delta x)^2) - x^2}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{2x\Delta x + (\Delta x)^2}{\Delta x} \\
&= \lim_{\Delta x \to 0} \left(\frac{2x\Delta x}{\Delta x} + \frac{(\Delta x)^2}{\Delta x}\right) \\
&= \lim_{\Delta x \to 0} (2x + \Delta x) \\
&= 2x.
\end{aligned}$$

The process of determining the derivative of a function from first principles can be quite time consuming and involve lengthy mathematical calculations. Fortunately, there is a more rapid route to determining the derivative of the sorts of functions that we encounter in economics based on a number of rules, known as rules of differentiation. Some of these rules will be derived in the next section using the definition of the derivative of a function (6.2), but others will be stated simply without justification.

Example 6.1

Differentiate $y = f(x) = x^2$ and use the small increments formula to estimate the change in y if x changes from 1 to 1.01. Calculate also the actual change in y.

Solution. We have already shown that $dy/dx = f'(x) = 2x$ and so $f'(1) = 2$. If x increases from 1 to 1.01, then $\Delta x = 0.01$. Therefore, we can estimate the change in y using the small increments formula (6.4) as

$$\Delta y \approx \Delta x \times f'(1) = 0.01 \times 2 = 0.02.$$

The actual change Δy in y is $f(1.01) - f(1) = 1.0201 - 1 = 0.0201$.

6.2 Rules of Differentiation

In this section, we show how to differentiate functions without having to use the definition (6.2). A few rules are sufficient to differentiate all the functions encountered in this book.

6.2.1 Constant Functions

Consider the constant function $f(x) = k$, where k is a constant. Using the definition of a derivative (6.2), we have

$$\begin{aligned} f'(x) &= \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} \\ &= \lim_{\Delta x \to 0} \frac{k - k}{\Delta x} \\ &= \lim_{\Delta x \to 0} \frac{0}{\Delta x} \\ &= 0. \end{aligned}$$

Thus, if $f(x) = k$ then $f'(x) = 0$. For example, if $f(x) = 8$, then $f'(x) = 0$. This rule is obvious if $f'(x)$ is regarded as the rate of change of $f(x)$ relative to x. In this case, $f(x)$ is constant.

6.2.2 Linear Functions

Consider the linear function $f(x) = ax + b$, where a and b are constants. Using the definition of a derivative (6.2), we have

$$\begin{aligned} f'(x) &= \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x} \\ &= \lim_{\Delta x \to 0} \frac{(a(x + \Delta x) + b) - (ax + b)}{\Delta x} \\ &= \lim_{\Delta x \to 0} \frac{a\Delta x}{\Delta x} \\ &= \lim_{\Delta x \to 0} a \\ &= a. \end{aligned}$$

Thus, if $f(x) = ax + b$ then $f'(x) = a$. This is the linear function rule. For example, if $f(x) = 3x+2$, then $f'(x) = 3$, and if $f(x) = 5-\frac{1}{4}x$, then $f'(x) = -\frac{1}{4}$.

6.2.3 Power Functions

Consider the power function $f(x) = kx^n$, where k is a constant and n is any real number. The derivative of this power function is given by $f'(x) = knx^{n-1}$. So to obtain the derivative of a power function, we multiply it by the power and reduce the original power by one. For example, if $f(x) = 4x^3$, then $f'(x) = 4 \times 3 \times x^{3-1} = 12x^2$, and if $f(x) = x^4$, then $f'(x) = 4x^{4-1} = 4x^3$. When $k = 1$, an important special case of this rule is realized, i.e., if $f(x) = x^n$, then

$$f'(x) = nx^{n-1}. \tag{6.5}$$

This rule, known as the power function rule, is derived using the definition (6.2). Since it involves the expansion of $(x + \Delta x)^n$ and some algebra, we omit the details here.

6.2.4 Sums and Differences of Functions

The rules we have introduced so far can be used to generate the derivatives of polynomials, the terms of which are power functions. Consider the function f, which is the sum of two functions g and h, i.e., $f(x) = g(x) + h(x)$. Using the definition of a derivative (6.2) we have

$$\begin{aligned}
f'(x) &= \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{(g(x+\Delta x) + h(x+\Delta x) - (g(x) + h(x))}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{(g(x+\Delta x) - g(x)) + (h(x+\Delta x) - h(x))}{\Delta x} \\
&= \lim_{\Delta x \to 0} \frac{(g(x+\Delta x) - g(x))}{\Delta x} \\
&\quad + \lim_{\Delta x \to 0} \frac{(h(x+\Delta x) - h(x))}{\Delta x} \\
&= g'(x) + h'(x).
\end{aligned}$$

Thus, if $f(x) = g(x) + h(x)$, then

$$f'(x) = g'(x) + h'(x).$$

This is intuitively clear when derivatives are viewed as rates of change: the rate of change relative to x of two functions of x is the sum of their rates of change. Similarly, we can show that if f is the difference of two functions g and h, i.e., $f(x) = g(x) - h(x)$, then

$$f'(x) = g'(x) - h'(x).$$

Thus, the derivative of a sum of two functions is equal to the sum of the derivatives of the individual functions. Similarly, the derivative of the difference of two functions is equal to the difference of the derivatives of the two functions. For example, if $f(x) = 12x^5 - 4x^4$, then $f'(x) = 60x^4 - 16x^3$, and if $f(x) = 9x^2 + 2x - 3$, then $f'(x) = 18x + 2$.

Example 6.2

Differentiate each of the following functions:

1. $f(x) = 9x - 6$,
2. $y = -9x^{-4}$,
3. $f(x) = x^8 + 8x^6 + 11$.

Solution.

1. This is a linear function (see Section 6.2.2) with $a = 9$ and $b = -6$. Therefore, using the linear function rule, we have $f'(x) = 9$.

2. This is a power function (see Section 6.2.3) with $k = -9$ and $n = -4$. Therefore, using the power function rule, we have

$$\frac{dy}{dx} = (-9)(-4)x^{-4-1} = 36x^{-5} = \frac{36}{x^5}.$$

3. This is an example of a polynomial function comprising two power functions and a constant function. Therefore, using the rule for the sum of functions in conjunction with the power function and constant function (see Section 6.2.1) rules, we have

$$f'(x) = 8x^{8-1} + 8 \times 6x^{6-1} + 0 = 8x^7 + 48x^5.$$

Note that the linear function rule can be deduced from a combination of the rule for the differentiation of the sum of two functions and the constant function and power function rules.

6.2.5 Product of Functions

Suppose that $y = uv$ where u and v are functions of x. Let Δu, Δv, and Δy denote very small changes in u, v, and y, respectively, that correspond to a small change Δx in x. Then

$$\begin{aligned} y + \Delta y &= (u + \Delta u)(v + \Delta v) \\ &= uv + u\Delta v + v\Delta u + \Delta u\Delta v. \end{aligned}$$

Since $y = uv$,

$$\Delta y = u\Delta v + v\Delta u + \Delta u\Delta v.$$

We can ignore the term $\Delta u\Delta v$ since it is the product of two very small changes and therefore negligible. Therefore,

$$\frac{\Delta y}{\Delta x} = u\frac{\Delta v}{\Delta x} + v\frac{\Delta u}{\Delta x} + \frac{\Delta u}{\Delta x}\Delta v.$$

As $\Delta x \to 0$,

$$\frac{\Delta y}{\Delta x} \to \frac{dy}{dx},$$

and

$$u\frac{\Delta v}{\Delta x} + v\frac{\Delta u}{\Delta x} + \frac{\Delta u}{\Delta x}\Delta v \to u\frac{dv}{dx} + v\frac{du}{dx},$$

since $(\Delta u/\Delta x)\Delta v \to 0$ as $\Delta x \to 0$. Thus, we obtain the product rule for differentiation: if $y = uv$, then

$$\frac{dy}{dx} = u\frac{dv}{dx} + v\frac{du}{dx}. \tag{6.6}$$

6.2.6 Quotient of Functions

Suppose that $y = u/v$ where u and v are functions of x. Let Δu, Δv, and Δy denote small changes in u, v, and y, respectively, that correspond to a very small change Δx in x. Thus $\Delta u \to 0$ and $\Delta v \to 0$ as $\Delta x \to 0$. Then

$$y + \Delta y = \frac{u + \Delta u}{v + \Delta v}.$$

Subtracting $y = u/v$ from both sides of this equation yields

$$\Delta y = \frac{u + \Delta u}{v + \Delta v} - \frac{u}{v}.$$

Simplifying the fraction on the right-hand side of this equation gives

$$\begin{aligned} \Delta y &= \frac{v(u + \Delta u) - u(v + \Delta v)}{v(v + \Delta v)} \\ &= \frac{v\Delta u - u\Delta v}{v(v + \Delta v)}. \end{aligned}$$

Therefore,

$$\frac{\Delta y}{\Delta x} = \frac{v\frac{\Delta u}{\Delta x} - u\frac{\Delta v}{\Delta x}}{v(v + \Delta v)}.$$

Finally, letting $\Delta x \to 0$, we obtain the quotient rule for differentiation: if $y = u/v$ then

$$\frac{dy}{dx} = \frac{v\frac{du}{dx} - u\frac{dv}{dx}}{v^2}. \tag{6.7}$$

6.2.7 The Chain Rule

Suppose that y is a function of u, i.e., $y = f(u)$, and that u in turn is a function of x, i.e., $u = g(x)$. We say that y is a function of a function and to express y as a function of x we write

$$y = f(g(x)).$$

If Δy and Δu denote changes in y and u, respectively, that correspond to a small change Δx in x, then

$$\frac{\Delta y}{\Delta x} = \frac{\Delta y}{\Delta u}\frac{\Delta u}{\Delta x}.$$

Then, as $\Delta x \to 0$, we obtain the so-called **chain rule**:

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx} = f'(u)g'(x). \tag{6.8}$$

We may also write this in terms of derivatives of f and g and then express the result solely in terms of a function of x, i.e.,

$$\frac{dy}{dx} = f'(g(x))g'(x). \tag{6.9}$$

As an illustration of the use of the chain rule to obtain the derivative of a function, consider $y = (x^2 + 3x + 2)^5$. If we let $u = x^2 + 3x + 2$, then $y = u^5$. Differentiating u with respect to x and y with respect to u, we obtain

$$\frac{dy}{du} = 5u^4, \quad \frac{du}{dx} = 2x + 3.$$

Then using the chain rule yields

$$\begin{aligned} \frac{dy}{dx} &= \frac{dy}{du}\frac{du}{dx} \\ &= 5u^4(2x+3) \\ &= 5(x^2+3x+2)^4(2x+3). \end{aligned}$$

If we put $x = y$ in (6.8), we obtain

$$\frac{dy}{dy} = 1 = \frac{dy}{du}\frac{du}{dy}.$$

It follows that

$$\frac{du}{dy} = \frac{1}{\frac{dy}{du}}. \tag{6.10}$$

Example 6.3

Find the derivative of each of the following functions:

1. $f(x) = (2x^3 + 1)(x^2 - 3x)$ and evaluate $f'(1)$,
2. $f(x) = \dfrac{5x^2 + 3}{x^2 + 1}$ and evaluate $f'(0)$,
3. $y = (7x^4 + 2)^6$ and evaluate dy/dx when $x = 0$.

Solution.

1. To differentiate this function, we use the product rule with $u = 2x^3 + 1$ and $v = x^2 - 3x$. Now $du/dx = 6x^2$ and $dv/dx = 2x - 3$. Therefore, using the product rule (6.6), we have

$$f'(x) = (2x^3 + 1)(2x - 3) + 6x^2(x^2 - 3x),$$

which, after some simplification, gives

$$f'(x) = 10x^4 - 24x^3 + 2x - 3.$$

Finally, $f'(1) = 10 - 24 + 2 - 3 = -15$.

2. To differentiate this function, we use the quotient rule with $u = 5x^2 + 3$ and $v = x^2 + 1$. Now $du/dx = 10x$ and $dv/dx = 2x$. Therefore, using the quotient rule (6.7), we have

$$\begin{aligned} f'(x) &= \frac{(10x)(x^2+1) - (5x^2+3)(2x)}{(x^2+1)^2} \\ &= \frac{10x^3 + 10x - 10x^3 - 6x}{(x^2+1)^2} \\ &= \frac{4x}{(x^2+1)^2}. \end{aligned}$$

Evaluating this derivative when $x = 0$ gives $f'(0) = 0$.

3. To differentiate $y = (7x^4 + 2)^6$, we use the chain rule (6.8). Let $u = 7x^4 + 2$ then $y = u^6$. Now

$$\frac{dy}{du} = 6u^5, \quad \frac{du}{dx} = 28x^3.$$

Therefore,

$$\frac{dy}{dx} = (6u^5)(28x^3) = 168(7x^4 + 2)^5 x^3.$$

When $x = 0$, $dy/dx = 0$.

6.3 Exponential and Logarithmic Functions

Let f be the exponential function $f(x) = e^{g(x)}$, where $g(x)$ is some function of x. Then the derivative of f is

$$f'(x) = g'(x)\, e^{g(x)}. \tag{6.11}$$

For example, if $f(x) = e^{x^2}$, then $f'(x) = 2xe^{x^2}$ since $g(x) = x^2$ and $g'(x) = 2x$. When $g(x) = 1$, we have the important result that the derivative of the exponential function e^x is itself, i.e., e^x, since $g'(x) = 1$. More generally, if $f(x) = e^{kx}$, where k is a constant, then

$$f'(x) = ke^{kx}.$$

For example, if $f(x) = e^{-2x}$, then $f'(x) = -2e^{-2x}$.

Let f be the natural logarithmic function $f(x) = \ln g(x)$, then the derivative of f is

$$f'(x) = \frac{g'(x)}{g(x)}. \tag{6.12}$$

For example, if $f(x) = \ln 6x^2$ then

$$f'(x) = \frac{12x}{6x^2} = \frac{2}{x}$$

since $g(x) = 6x^2$ and $g'(x) = 12x$. When $g(x) = x$ we have the result that the derivative of $\ln x$ is $1/x$ since $g'(x) = 1$.

We display these rules in Table 6.1.

Example 6.4

Find the derivative of each of the following functions:

1. $f(x) = 3e^{7-2x}$,
2. $f(x) = \ln(x^2 + 6x + 2)$.

Solution.

1. If $f(x) = 3e^{7-2x}$, then $g(x) = 7 - 2x$. Since $g'(x) = -2$, then

$$f'(x) = 3g'(x)e^{g(x)} = -6e^{7-2x}.$$

Table 6.1 Derivatives of the exponential and logarithmic functions.

$f(x)$	$f'(x)$
$e^{g(x)}$	$g'(x)e^{g(x)}$
e^x	e^x
$\ln g(x)$	$\frac{g'(x)}{g(x)}$
$\ln x$	$\frac{1}{x}$

2. If $f(x) = \ln(x^2 + 6x + 2)$, then $g(x) = x^2 + 6x + 2$. Since $g'(x) = 2x + 6$, then
$$f'(x) = \frac{g'(x)}{g(x)} = \frac{2x+6}{x^2+6x+2}.$$

6.4 Marginal Functions in Economics

6.4.1 Marginal Revenue and Marginal Cost

Sometimes in economics, we are interested in the effect on total revenue, TR, of a change in the value of Q. To do this, the concept of **marginal revenue** is introduced. The marginal revenue of a good, MR, is defined by

$$MR = \frac{d(TR)}{dQ}.$$

The marginal revenue function measures the instantaneous rate of change in total revenue, TR, compared with demand, Q. For example, the marginal revenue function, MR, corresponding to

$$TR = 100Q - 2Q^2$$

is given by

$$MR = \frac{d(TR)}{dQ} = 100 - 4Q.$$

If the current demand is 15, say, then

$$MR = 100 - 4 \times 15 = 40.$$

This means that when demand is changed slightly from its current value of 15, the corresponding change in total revenue is 40 times as large. However, if the demand is 20, then

$$MR = 100 - 4 \times 20 = 20,$$

which means that when demand is changed slightly from $Q = 20$, the corresponding change in total revenue is only 20 times as large.

Economists say that MR is approximately the change in TR resulting from a one unit increase in demand Q. In general,

$$\Delta(TR) \approx MR \times \Delta Q.$$

(This is just a consequence of the small increments formula (6.4).) This approximation is a good one provided the quantities of Q involved are very large so that one unit is relatively very small. An analogous statement can be made regarding marginal cost, MC.

The marginal cost function, MC, is defined by

$$MC = \frac{d(TC)}{dQ}. \tag{6.13}$$

The average cost function, AC, is defined by

$$AC = \frac{TC}{Q}. \tag{6.14}$$

Example 6.5

If the average cost function for a good is

$$AC = \frac{24}{Q} + 15 + 3Q,$$

find an expression for the total cost function. What are the fixed costs in this case? Write down an expression for the marginal cost function.

Solution. To find an expression for TC, we use the formula for AC given by (6.14). Hence

$$\begin{aligned} TC &= AC \times Q \\ &= \left(\frac{24}{Q} + 15 + 3Q\right) Q \\ &= 24 + 15Q + 3Q^2. \end{aligned}$$

Since $TC = FC + (VC)Q$, the fixed cost element of the total cost function is independent of Q. Therefore, in this example the fixed costs are 24. Finally, an expression for the marginal cost function is obtained by differentiating TC with respect to Q. Therefore,

$$\begin{aligned} MC &= \frac{d(TC)}{dQ} \\ &= 15 + 6Q. \end{aligned}$$

Note that the fixed costs have no influence on the marginal cost function since the derivative of a constant is zero.

6.4.2 Marginal Propensities

The relationship between consumption C and national income Y is sometimes of the form

$$C = f(Y),$$

where f is some appropriate consumption function. Of interest is the effect on C due to variations in Y, i.e., if national income rises by a certain amount, what effect does this have on the spending patterns of the population. This is analyzed using the concept of **marginal propensity to consume**, MPC, defined by

$$MPC = \frac{dC}{dY},$$

i.e., the marginal propensity to consume is the derivative of consumption with respect to income. For example, if the consumption function is

$$C = 0.01Y^2 + 0.2Y + 50$$

to calculate MPC when $Y = 30$, we have

$$MPC = \frac{dC}{dY} = 0.02Y + 0.2.$$

When $Y = 30$, $MPC = (0.02)(30) + 0.2 = 0.8$.

Economists say that MPC is approximately the change in consumption due to a one unit increase in national income Y. More generally, if national income increases by a small amount ΔY, then the corresponding small change ΔC in consumption is approximately $MPC \times \Delta Y$, i.e.,

$$\Delta C \approx MPC \times \Delta Y.$$

If national income is used up only in consumption and savings, then

$$Y = C + S.$$

If we differentiate both sides of this equation with respect to Y:

$$\frac{dY}{dY} = \frac{dC}{dY} + \frac{dS}{dY},$$

i.e.,

$$1 = MPC + MPS,$$

where

$$MPS = \frac{dS}{dY}$$

is the **marginal propensity to save**. Economists say that MPS is approximately the change in savings due to a one unit increase in national income Y.

More generally, we can show, using the small increments formula again, that if national income increases by a small amount ΔY, then the corresponding small change ΔS in savings is given by

$$\Delta S \approx MPS \times \Delta Y.$$

Thus if we know MPC, we can easily determine MPS. In the above example the value of MPS when $Y = 30$ is given by

$$1 = 0.8 + MPS$$

i.e.,

$$MPS = 0.2.$$

This indicates that when income increases by one unit (from its current level of 30), consumption rises by approximately 0.8 units, whereas savings rise by approximately 0.2 units. At this level of income, the nation has a greater propensity to consume than it has to save.

Example 6.6

If the consumption function is

$$C = 0.005Y^2 + 0.3Y + 20,$$

calculate the marginal propensities to consume and save when $Y = 10$ and give an interpretation of the results.

Solution. The marginal propensity to consume is defined by

$$MPC = \frac{dC}{dY} = 0.01Y + 0.3.$$

When $Y = 10$,

$$MPC = 0.01 \times 10 + 0.3 = 0.1 + 0.3 = 0.4.$$

If national income is used up in consumption and savings only, then

$$MPC + MPS = 1.$$

When $Y = 10$, the marginal propensity to save is

$$MPS = 1 - MPC = 1 - 0.4 = 0.6.$$

Therefore, at this level of national income, the nation has a greater propensity to save than it has to consume.

6.5 Approximation to Marginal Functions

The exact value of MR at the point Q_0 is given by

$$\frac{d(TR)}{dQ},$$

and so is given by the slope of the tangent to the total revenue function at A (see Fig. 6.2). The point B also lies on the curve – it corresponds to a one unit increase in Q, i.e., $\Delta Q = 1$. The vertical distance from A to B therefore equals the change in TR when Q increases by one unit. The slope of the chord joining A to B is

$$\frac{\Delta(TR)}{\Delta Q} = \frac{\Delta(TR)}{1} = \Delta(TR).$$

Note that the slope of the tangent is approximately the same as that of the chord joining A and B. Therefore, the latter produces a reasonable approximation to MR in many cases.

This approximation holds for any value of ΔQ. Therefore, as we have seen in Section 6.4.1

$$MR \approx \frac{\Delta(TR)}{\Delta Q}, \tag{6.15}$$

or

$$\Delta(TR) \approx MR \times \Delta Q, \tag{6.16}$$

i.e., change in total revenue ≈ marginal revenue × change in demand. Note that if the total revenue function is linear, then we have equality: $\Delta(TR) = MR \times \Delta Q$.

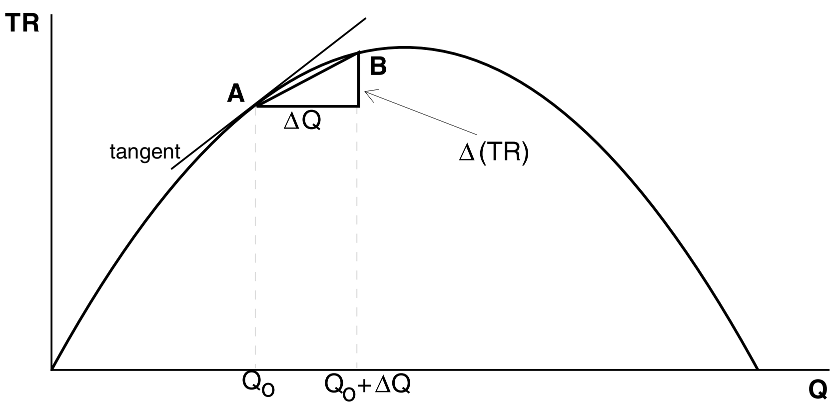

Figure 6.2 Approximation to marginal revenue.

Example 6.7

If the total revenue function of a good is given by

$$100Q - Q^2,$$

write down an expression for the marginal revenue function. If the current demand is 60, estimate the change in the value of TR due to a two unit increase in Q.

Solution. To determine the marginal revenue function, we differentiate the total revenue function. Therefore,

$$MR = \frac{d(TR)}{dQ} = 100 - 2Q.$$

When $Q = 60$,

$$MR = 100 - 2 \times 60 = 100 - 120 = -20.$$

When there is a two unit increase in Q, i.e., $\Delta Q = 2$, then the estimated change in TR is given by

$$\Delta(TR) \approx MR \times \Delta Q = -20 \times 2 = -40,$$

i.e., there is an estimated 40 unit reduction in TR.

A similar approximation to (6.15), using the small increments formula (6.4), holds for the marginal cost function:

$$MC \approx \frac{\Delta(TC)}{\Delta Q}, \tag{6.17}$$

or

$$\Delta(TC) \approx MC \times \Delta Q, \tag{6.18}$$

i.e., change in total cost $\approx$ marginal cost $\times$ change in demand. Note that we have equality if the total cost function is linear, then $\Delta(TC) = MC \times \Delta Q$.

Example 6.8

Find the marginal cost function given the average cost function

$$AC = \frac{100}{Q} + 2.$$

Deduce that a one unit increase in Q will always result in a two unit increase in TC, irrespective of the current level of output.

Solution. To determine the marginal cost function, it is first necessary to find an expression for the total cost function, TC. Now TC and AC are related by

$$AC = \frac{TC}{Q},$$

and therefore

$$TC = AC \times Q = \left(\frac{100}{Q} + 2\right) Q = 100 + 2Q.$$

The corresponding marginal cost function is

$$MC = \frac{d(TC)}{dQ} = 2.$$

Since TC is a linear function, we have $\Delta(TC) = MC \times \Delta Q$. Therefore, if output increases by one unit, i.e., $\Delta Q = 1$, then

$$\Delta(TC) = 2,$$

irrespective of the current level of output.

6.6 Higher Order Derivatives

We have already seen that the derivative of a function of x is itself a function of x. This suggests the possibility of differentiating a second time to get the 'slope of the slope of a function'. This is written as $f''(x)$ or d^2y/dx^2. This function is known as the second order derivative of $f(x)$. Higher order derivatives are found by applying the rules of differentiation to lower order derivatives. The third order derivative $f'''(x)$ or d^3y/dx^3 measures the slope and rate of change of the second order derivative, etc. Thus, if

$$f(x) = 2x^4 + 5x^3 + 3x^2,$$

we have

$$f'(x) = 8x^3 + 15x^2 + 6x$$
$$f''(x) = 24x^2 + 30x + 6$$
$$f'''(x) = 48x + 30$$
$$f^{(4)}(x) = 48$$
$$f^{(5)}(x) = 0$$

Example 6.9

For each of the following functions, find the second derivative and evaluate it at $x = 2$.

1. $f(x) = x^6 + 3x^4 + x$,
2. $y = 2x^2 + 38x - 6$,
3. $y = (8x - 4)^6$.

Solution.

1. To differentiate this polynomial function, we use a combination of the rule for the sum of functions and the power function rule. So

$$\begin{aligned} f'(x) &= 6x^{6-1} + 3 \times 4x^{4-1} + 1 \\ &= 6x^5 + 12x^3 + 1. \end{aligned}$$

Differentiating a second time gives

$$\begin{aligned} f''(x) &= 6 \times 5x^{5-1} + 12 \times 3x^{2-1} + 0 \\ &= 30x^4 + 36x^2 \end{aligned}$$

Evaluating the second derivative when $x = 2$, we have

$$f''(2) = 30(2^4) + 36(2^2) = 624$$

2. To differentiate this quadratic function, we use a combination of the rule for the sum of functions and the power function rule. So

$$\begin{aligned} \frac{dy}{dx} &= 2 \times 2x^{2-1} + 38 \\ &= 4x + 38. \end{aligned}$$

Differentiating a second time gives

$$\frac{d^2y}{dx^2} = 4$$

At $x = 2$, $\frac{d^2y}{dx^2} = 4$.

3. To differentiate this function, we use the chain rule. Let $u = 8x - 4$, then $y = u^6$. Since

$$\frac{dy}{du} = 6u^5, \quad \text{and} \quad \frac{du}{dx} = 8,$$

we have, using the chain rule

$$\frac{dy}{dx} = \frac{dy}{du}\frac{du}{dx} = 6u^5 \times 8 = 48(8x - 4)^5.$$

Applying the chain rule a second time gives

$$\frac{d^2y}{dx^2} = 48 \times 5u^4 \times 8 = 1920(8x - 4)^4.$$

Evaluating the second derivative when $x = 2$ gives

$$\frac{d^2y}{dx^2} = 39813120.$$

6.7 Production Functions

In one of the simplest models for production, the quantity of output produced, Q, is assumed to be a function of capital, K, and labour, L. However, in the short run K can be assumed to be fixed and so Q is then a function of L alone. In this instance, Q is referred to as the short run production function. The independent variable L is usually measured in terms of the number of workers or the number of worker hours. The derivative of the production function with respect to L, known as the **marginal product of labour** (MP_L), measures the rate at which output changes as the number of workers increases. Thus, we have

$$MP_L = \frac{dQ}{dL}. \tag{6.19}$$

Economists say that MP_L is approximately the change in output resulting from a one unit increase in labour.

Example 6.10

For the production function

$$Q = 8\sqrt{L},$$

find the marginal product of labour. Determine the output and the marginal product of labour when

1. $L = 1$,
2. $L = 4$,
3. $L = 100$.

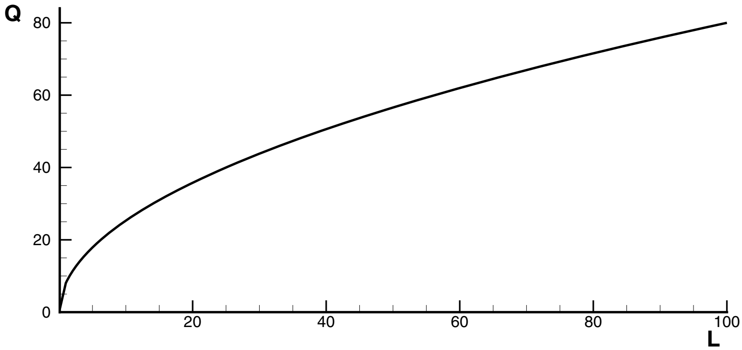

Figure 6.3 Graph of the production function $Q = 8L^{1/2}$.

Solution. The marginal product of labour is found by differentiating $Q = 8L^{1/2}$. This gives, using the power function rule,

$$MP_L = \frac{dQ}{dL} = 8 \times \frac{1}{2}L^{1/2-1} = 4L^{-1/2} = \frac{4}{L^{1/2}}.$$

1. When $L = 1$, $Q = 8$ and $MP_L = 4$.
2. When $L = 4$, $Q = 16$ and $MP_L = 2$.
3. When $L = 100$, $Q = 80$ and $MP_L = 0.4$.

As L increases from 0, so does output (see Fig. 6.3). However, MP_L decreases and therefore although output increases, it does so at a decreasing rate. In this situation, we say that there are diminishing returns to labour.

Example 6.11

Consider the production function is

$$Q = 120\sqrt{L} - 5L,$$

where Q denotes output and L denotes the size of the workforce. Calculate the value of MP_L when

1. $L = 1$,
2. $L = 16$,
3. $L = 100$,
4. $L = 900$,

and discuss the implication of these results.

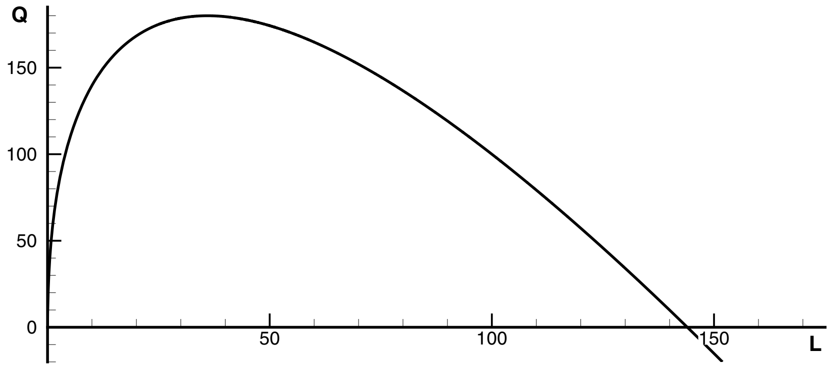

Figure 6.4 Graph of the production function $Q = 60\sqrt{L} - 5L$.

Solution. The marginal product of labour, MP_L, is found by differentiating the production function with respect to L. This gives

$$MP_L = \frac{dQ}{dL} = 120 \times \frac{1}{2}L^{1/2-1} - 5 = 60L^{-1/2} - 5 = \frac{60}{L^{1/2}} - 5.$$

1. When $L = 1$, $MP_L = 55$.
2. When $L = 16$, $MP_L = 10$.
3. When $L = 100$, $MP_L = 1$,
4. When $L = 900$, $MP_L = -3$.

In the last part of this example, we see that a size of workforce is reached that, if exceeded, actually results in a decrease in output. This may seem counterintuitive at first sight. However, this situation can occur in production processes where productivity is diminished due to problems of overcrowding on the shop floor or the need to create an elaborate administration to organize the larger workforce. The graph of this production function is sketched in Fig. 6.4.

The production function in the last example satisfies what is known as the **law of diminishing marginal productivity**. This law, also known as the **law of diminishing returns**, states that the increase in output due to a one unit increase in labour will eventually decline. A typical production function that satisfies this law is shown in Fig. 6.5. The graph of the corresponding marginal product of labour, MP_L, is shown in Fig. 6.6. Note that the maximum value of MP_L is attained when $L = L_0$ and $MP_L = 0$ at the value of L corresponding to maximum production.

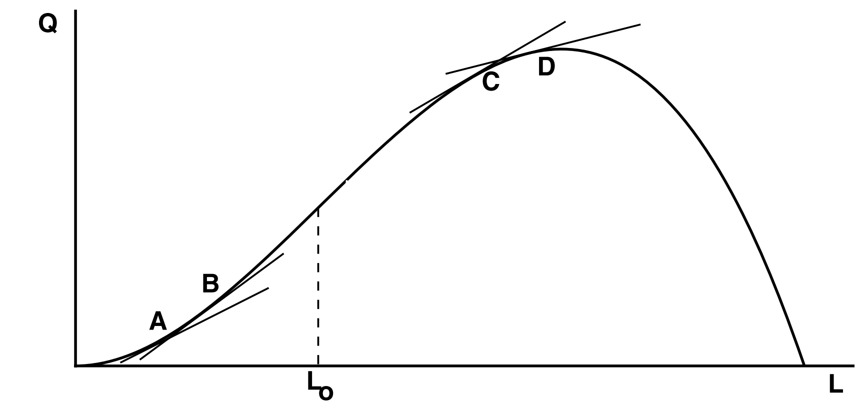

Figure 6.5 Graph illustrating a production function that satisfies the law of diminishing marginal productivity.

Between $L = 0$ and $L = L_0$, the curve bends upwards, becoming progressively steeper and so the slope, MP_L, of the production function increases. Mathematically speaking,

$$\frac{d(MP_L)}{dL} > 0,$$

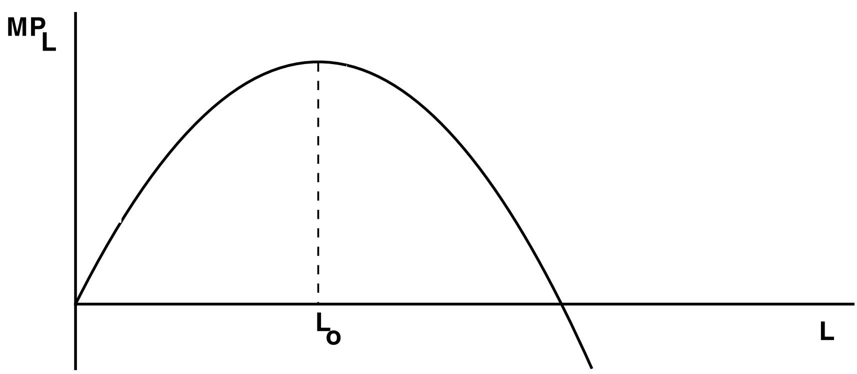

Figure 6.6 Graph of marginal product of labour corresponding to the production function shown in Fig. 6.5.

or, since $MP_L = dQ/dL$,

$$\frac{d^2Q}{dL^2} > 0,$$

i.e., if we take any two points A: $(L_1, Q(L_1))$ and B: $(L_2, Q(L_2))$ on the curve between the points $(0, 0)$ and $(L_0, Q(L_0))$ with $L_1 < L_2$, then the slope of the tangent at B is greater than that at A (see Fig. 6.5). Similarly, for $L > L_0$ the curve of the production function bends downwards and the slope of the slope function decreases and is negative, i.e.,

$$\frac{d^2Q}{dL^2} < 0,$$

i.e., if we take any two points C: $(L_3, Q(L_3))$ and D: $(L_4, Q(L_4))$ on the curve beyond the point $(L_0, Q(L_0))$ with $L_3 < L_4$, then the slope of the tangent at C is greater than that at D (see Fig. 6.5). The law of diminishing returns states that this must happen eventually, i.e.,

$$\frac{d^2Q}{dL^2} < 0,$$

for $L > L_0$.

Example 6.12

Show that the law of diminishing marginal productivity holds for the production function

$$Q = 15L^2 - 0.2L^3.$$

Solution. Differentiating the production function gives

$$MP_L = \frac{dQ}{dL} = 30L - 0.6L^2.$$

Differentiating a second times gives

$$\frac{d^2Q}{dL^2} = 30 - 1.2L.$$

The expression defining the second derivative, i.e., $30 - 1.2L$ becomes negative when $30 - 1.2L < 0$, i.e., when

$$L > \frac{30}{1.2} = 25.$$

Therefore, the law of diminishing marginal productivity holds for this production function for $L > 25$, i.e., $L_0 = 25$.

EXERCISES

6.1. Find the derivatives of the following functions:

a) $f(x) = 4$,

b) $f(x) = 4x^3$,

c) $f(x) = x^8$,

d) $f(x) = 2x^{3/2}$,

e) $f(x) = 3x + 7$.

6.2. Find dy/dx for each of the following:

a) $y = 5 + 2x - 3x^2$,

b) $y = x^3 + 3x^2 + 5$,

c) $y = x^2 + 5$,

d) $y = x^4 - 3x^2 + 1$.

6.3. Find the first and second derivatives of the following functions:

a) $y = e^{4x}$,

b) $y = 3e^{-2x}$.

Evaluate these derivatives when $x = 0$.

6.4. Find the first and second derivatives of the following functions:

a) $y = \ln 4x$,

b) $y = 2 \ln 7x$.

Evaluate these derivatives when $x = 1$.

6.5. If $TC = 3Q^2 + 7Q + 12$, find expressions for the marginal and average cost functions. Evaluate them when $Q = 3$ and $Q = 5$.

6.6. For each of the following demand functions, find expressions for TR and MR and evaluate them when $Q = 4$ and $Q = 10$.

a) $Q = 36 - 2P$,

b) $44 - 4P - Q = 0$.

6.7. Find the first and second derivatives of the function

$$y = 6x^3 - 20x^2 - 9x + 12.$$

Evaluate these derivatives when $x = 1$.

6.8. Find the marginal cost function for the average cost function given by

$$AC = \tfrac{3}{2}Q + 4 + \tfrac{46}{Q}.$$

6.9. The fixed costs of producing a good are 50 and the variable costs are $2 + \frac{1}{4}Q$ per unit.

a) Find expressions for TC and MC.

b) Evaluate TC and MC when $Q = 20$. Hence estimate the change in TC brought about by a 2 unit increase in output from the current level of 20 units.

6.10. If the consumption function is

$$C = 0.03Y^2 + 0.1Y + 30$$

calculate MPC and MPS when $Y = 4$ and give an interpretation of the results.

6.11. Show that the law of diminishing marginal productivity holds for the production function

$$Q = 18L^2 - 0.6L^3.$$

7
Maxima and Minima

7.1 Introduction

In this book, the concept of the derivative of a function has been introduced, and its application in economics has been described. However, the primary use of the derivative in economic analysis is related to the process of optimization. Optimization is defined to be the process of determining the local or relative maximum or minimum of a function.

In this chapter, the process of determining and classifying the relative or local extrema of a given function is described from a mathematical perspective by appealing to the local properties of the function near the extrema. The application of this theory to a range of functions that arise in economics is described in some detail together with an interpretation of the results. Optimization is important and useful for solving a range of problems in micro and macro economics. For example, in the theory of production, the firm wishes to maximize the output. In microeconomics, a business wishes to maximize profit. In macroeconomics, a government may wish to maximize revenue from taxation. The determination of the maxima and minima of a function also provides invaluable information for the purpose of sketching its graph.

7.2 Local Properties of Functions

In this section, some local properties of functions are introduced that will be useful in identifying and characterising the local maxima and minima of a given function.

7.2.1 Increasing and Decreasing Functions

A function $f(x)$ is said to be **increasing** on the domain $a \leq x \leq b$ if, for any two points x_1, x_2, where $a \leq x_1 < x_2 \leq b$, then $f(x_1) < f(x_2)$ (see Fig. 7.1(a)). That is, f increases as x increases. A function $f(x)$ is said to be decreasing on the domain $a \leq x \leq b$ if, for any two points x_1, x_2, where $a \leq x_1 < x_2 \leq b$, then $f(x_1) > f(x_2)$ (see Fig. 7.1 (b)).

Since the first derivative of a function measures the slope of a function, a function that is increasing on some domain is characterised by a positive first derivative. That is, $f(x)$ increases as x takes increasing values in the domain. More precisely, if $f'(x) > 0$ for all x belonging to some domain $a \leq x \leq b$, then the function f is said to be **increasing** for values of x satisfying $a \leq x \leq b$. Similarly, a function that is decreasing over some domain is characterised by a negative first derivative. More precisely, if $f'(x) < 0$ for all x belonging to some domain $a \leq x \leq b$, then the function f is said to be **decreasing** for values of x satisfying $a \leq x \leq b$. For example, the function $f(x) = x^2$ (see Fig. 7.2) is a decreasing function for $x < 0$ since $f'(x) = 2x < 0$ for $x < 0$ and an increasing function for $x > 0$ since $f'(x) > 0$ for $x > 0$. The function $f(x) = 4x - x^2$ (see Fig. 7.3) is an increasing function for $-1 \leq x \leq 2$ since $f'(x) = 4 - 2x > 0$ for $-1 \leq x \leq 2$ and a decreasing function for $2 \leq x \leq 5$ since $f'(x) < 0$ for $2 \leq x \leq 5$.

7.2.2 Concave and Convex Functions

Consider a function $f(x)$ defined on some domain. If the tangents to the graph of this function at each point on this domain are such that the graph lies above them, then the function is said to be convex on the domain. If the tangents to the graph of this function at each point on this domain are such that the curve lies below them, then the function is said to be convex on the domain. These two situations are shown in Fig. 7.4. In the case of the convex function shown in Fig. 7.4(a), we observe that the slope of the function increases as one moves from the point x_1 to the point x_2. In this particular example, the slope of the function is negative at $x = x_1$ and gradually increases to take a positive value

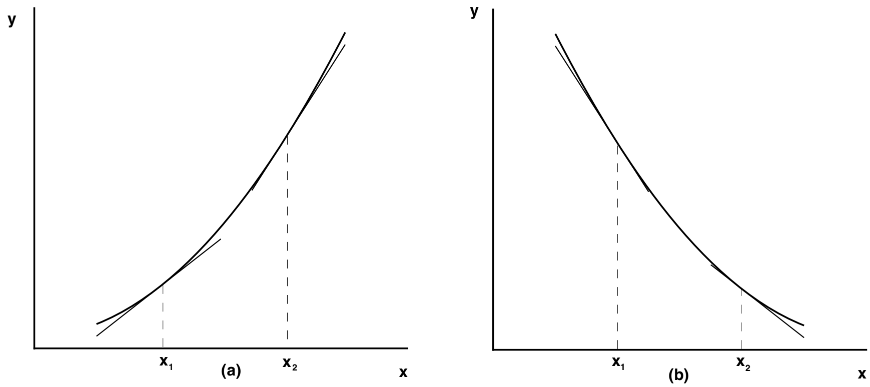

Figure 7.1 Examples of graphs of (a) an increasing function; (b) a decreasing function.

at $x = x_2$. Thus, a function that is **convex** on a domain is characterised by the condition $f''(x) > 0$ on the domain. Similarly, a function that is **concave** on a domain is characterised by the condition $f''(x) < 0$ on the domain (see Fig.7.4(b)).

For example, the function $f(x) = x^2$ (see Fig. 7.2) is convex on the domain $-2 \leq x \leq 2$. In fact, it is convex on any domain $a \leq x \leq b$ since $f''(x) = 2 > 0$. The function $f(x) = 4x - x^2$ (see Fig. 7.3) is concave on the domain $-1 \leq x \leq 5$. In fact, it is concave on any domain $a \leq x \leq b$ since $f''(x) = -2 < 0$.

7.3 Local or Relative Extrema

A function of x possesses a local maximum or minimum at $x = a$ if the function is neither increasing nor decreasing at $x = a$. That is, the rate of change of y relative to x is 0 when $x = a$. A **local or relative extremum** of a function is a point at which the function attains a local maximum or minimum. This means that the tangent to the curve $y = f(x)$ is 'horizontal' at a local or relative extremum and therefore has zero slope. Equivalently, since the slope is given by the first derivative of the function, that derivative must be zero at $x = a$. A point where $f'(x) = 0$ is known as a **critical point** or value. It is also known as a **stationary point**.

So the stationary or critical points of a function $f(x) = 0$ are the values of x for which $f'(x) = 0$. It remains to classify them as maxima or minima. This is done by calculating the second derivative of the function and evaluating it

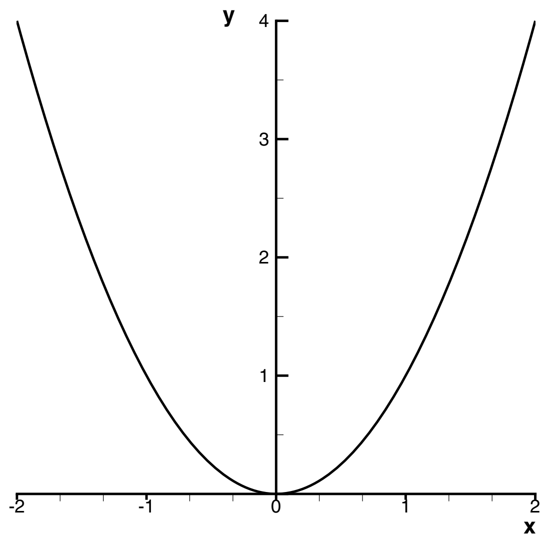

Figure 7.2 The graph of the function $f(x) = x^2$.

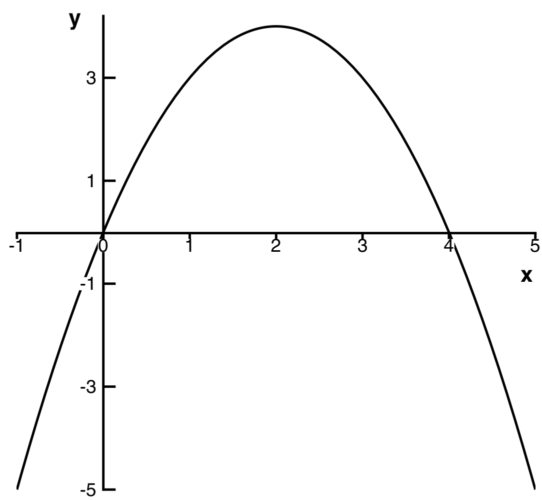

Figure 7.3 The graph of the function $f(x) = 4x - x^2$.

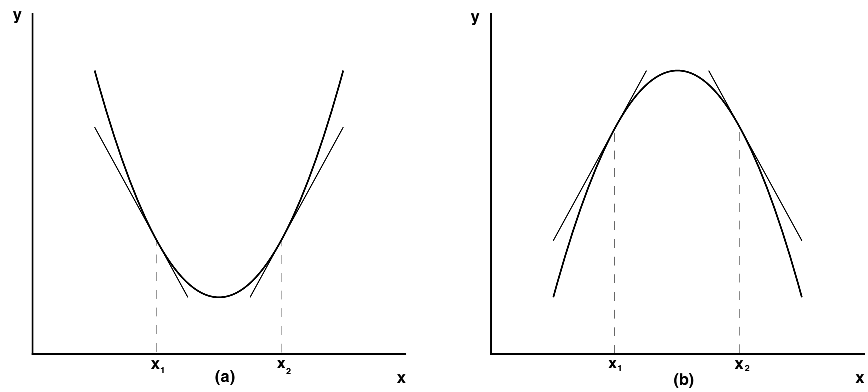

Figure 7.4 Examples of graphs of (a) a convex function; (b) a concave function.

at $x = a$.

To distinguish between a relative maximum and a relative minimum, it is necessary to inspect the behaviour of the second derivative and, in particular, to determine the sign of $f''(a)$ where $f'(a) = 0$. For points just to the left of a local maxima at $x = a$, the slope of the tangent is positive, and for points just to the right, the slope of the tangent is negative. So in the neighbourhood of a local maxima, the first derivative of $f(x)$ is a decreasing function of x, i.e., $f''(x) < 0$ and, in particular, $f''(a) < 0$. Therefore, if $f''(a) < 0$, which means that the function is concave and the curve lies below the tangent at $x = a$, then the function has a local maximum at $x = a$. For points just to the left of a local minima at $x = a$, the slope of the tangent is negative, and for points just to the right, the slope of the tangent is positive. So in the neighbourhood of a local minima, the first derivative of $f(x)$ is an increasing function of x (i.e., $f''(x) > 0$) and, in particular, $f''(a) > 0$. Therefore, if $f''(a) > 0$, which means that the function is convex and the curve lies above the tangent at $x = a$, then the function has a local minimum at $x = a$.

We now summarize the steps involved in finding and classifying the stationary points of a function $f(x)$:

Second Derivative Test

Step 1.

Solve the equation

$$f'(x) = 0$$

to find the stationary point(s).

Step 2.

Suppose $x = a$ gives a stationary point (i.e., $f'(a) = 0$).

If $f''(a) > 0$	then the function has a local minimum at $x = a$.
If $f''(a) < 0$	then the function has a local maximum at $x = a$.
If $f''(a) = 0$	then the test is inconclusive.

A knowledge of the stationary points of a function is essential when sketching the graph of a nonlinear function since it provides information about its general shape. The graph of a function can be sketched using a similar process to that used to determine and classify the stationary points of a function. Once the stationary points of a function have been determined and classified, the graph of the function can be sketched by drawing a smooth curve through these points. A more accurate representation of the graph may be obtained by evaluating the function at a greater number of points and drawing a smooth curve through them.

Example 7.1

Find and classify the stationary points of the following functions:

1. $f(x) = x^2 - 4x + 5$,
2. $f(x) = 2x^3 + 3x^2 - 12x + 4$.

Solution.

1. We need to calculate the first and second order derivatives of $f(x) = x^2 - 4x + 5$.

$$f'(x) = 2x - 4$$
$$f''(x) = 2$$

Step 1. The stationary points are the solutions of the equation

$$f'(x) = 0,$$

i.e.,

$$2x - 4 = 0.$$

Therefore $x = 2$ is a stationary point.

Step 2. To classify this point, we need to evaluate $f''(2)$. In this case, $f''(2) = 2 > 0$ so the function has a minimum at $x = 2$. The graph of this function is shown in Fig. 7.5.

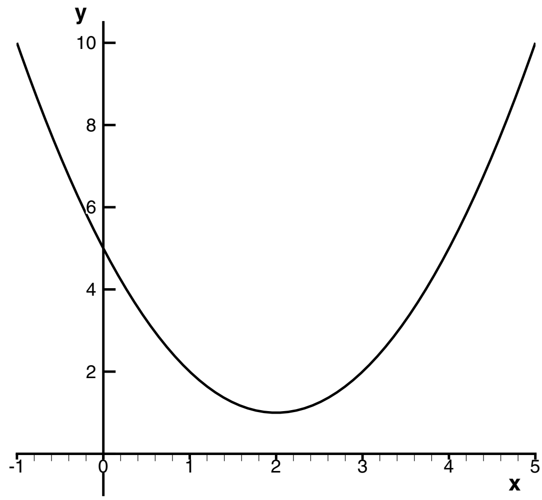

Figure 7.5 The graph of the function $f(x) = x^2 - 4x + 5$ plotted for values of x lying between -1 and 5.

2. In this example, we have $f(x) = 2x^3 + 3x^2 - 12x + 4$

$$f'(x) = 6x^2 + 6x - 12$$

$$f''(x) = 12x + 6$$

Step 1. The stationary points are the solutions of the equation $f'(x) = 0$, i.e.,

$$6(x^2 + x - 2) = 0$$

$$6(x + 2)(x - 1) = 0$$

Therefore, the stationary points are $x = -2$ and $x = 1$.

Step 2. To classify these points, we need to evaluate $f''(x)$ at $x = -2$ and $x = 1$. Now,

$$f''(-2) = -24 + 6 = -18 < 0,$$

and so the function has a maximum at $x = -2$, and

$$f''(1) = 12 + 6 = 18 > 0,$$

so the function has a minimum at $x = 1$. The graph of this function is shown in Fig. 7.6.

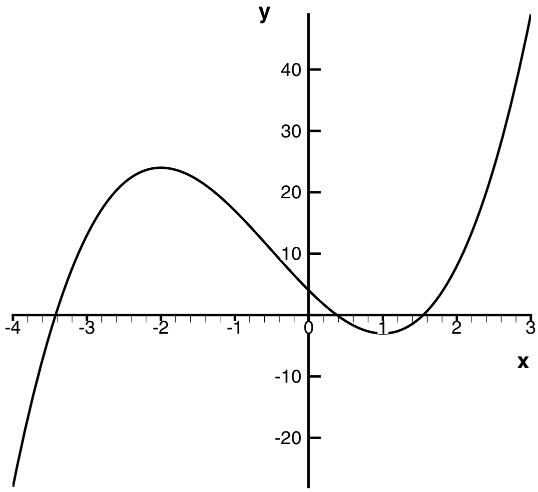

Figure 7.6 The graph of the function $f(x) = 2x^3 + 3x^2 - 12x + 4$ plotted for values of x lying between -4 and 3.

7.4 Global or Absolute Extrema

The functions we have investigated so far have either possessed a single stationary point (see Figs. 7.2 and 7.3) or two stationary points (see Fig. 7.6). In general, however, we may encounter functions that possess several local extrema of the same type. For example, the function $f(x) = \frac{2}{5}x^5 + \frac{3}{4}x^4 - 8x^3 - 3x^2 + 12x$ (see Fig. 7.7) has local maxima at $x = -4$ and $x = 1/2$ and local minima at $x = -1$ and $x = 3$. The higher of the two local maxima occurs at $x = -4$. However, the largest value of $f(x)$ for values of x lying between -4 and 3 occurs at the end point $x = 3$. We say that this point is a **global or absolute maximum**. The lower of the two local minima occurs at $x = 3$. However, the smallest value of $f(x)$ for values of x lying between -4 and 3 occurs at the other end point $x = -4$. We say that this point is a **global or absolute minimum**. Note that the absolute maximum and absolute minimum values of this function are not stationary points since the slope of the function is not zero at either $x = -4$ or $x = 3$. This example demonstrates that the absolute maximum or absolute minimum values of a function defined in a given interval may not occur at local extrema.

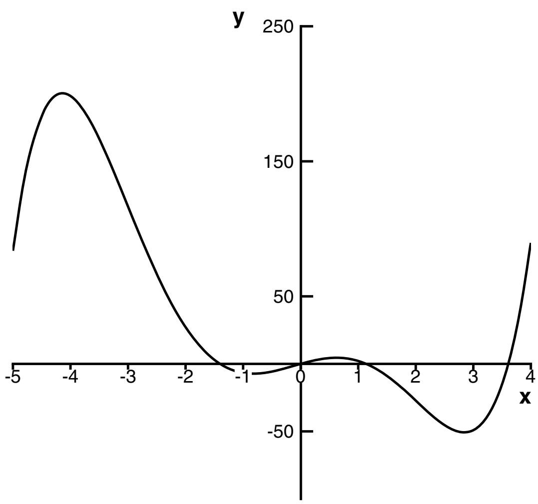

Figure 7.7 The graph of the function $f(x) = \frac{2}{5}x^5 + \frac{3}{4}x^4 - 8x^3 - 3x^2 + 12x$ plotted for values of x lying between -5 and 4.

There is no method for determining global extrema other than to evaluate the function at all local extrema and the end points and to determine from these calculations the values of x that generate the global extrema. For most of the examples we encounter in economics, the local extrema will coincide with the global extrema.

7.5 Points of Inflection

The local extrema of a function $f(x)$ have been characterised by the solutions $x = a$ of the equation $f'(x) = 0$ and classified as being local maxima or local minima depending on whether $f''(a) < 0$ or $f''(a) > 0$, respectively. So far, we have not asked what happens if $f''(a) = 0$. In this situation, the second derivative test is inconclusive and the stationary point $x = a$ is either a local maxima, a local minima, or a point of inflection. At a stationary point of inflection, the function is neither convex nor concave. The function crosses its tangent at this point and changes from concave to convex or vice versa.

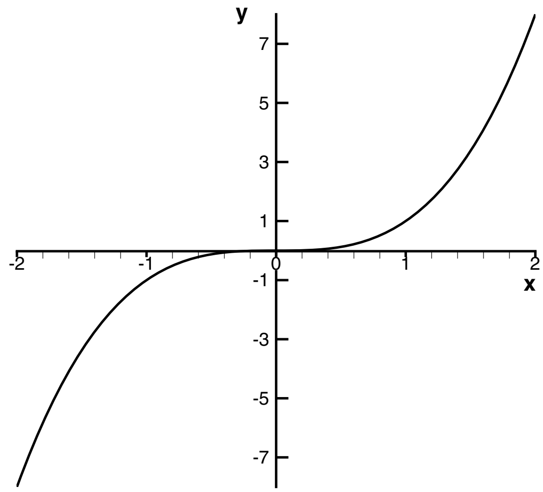

Figure 7.8 The graph of the function $f(x) = x^3$ plotted for values of x lying between -2 and 2.

For example, the function $f(x) = x^3$ (see Fig. 7.8) has a stationary point of inflection at $x = 0$. For this function, we have $f'(0) = f''(0) = 0$. In addition, the function is increasing for all values of x, convex for $x < 0$ and concave for $x > 0$. The function changes from being convex to concave at the point $x = 0$.

It is possible to have a point of inflection that is not a stationary point. For example, the function $f(x) = x^3 - 3x^2 + 2x$ (see Fig. 7.9) has a point of inflection at $x = 1$. At this point, we have $f''(1) = 0$ but $f'(1) = -1 \neq 0$.

7.6 Optimization of Production Functions

Production functions were introduced in Chapter 5. Production depends on a number of factors including capital and labour. However, in the short run we can assume that a firm's production depends solely on labour with all other factors of production, including capital, constant. We can express this symbolically by writing

$$Q = Q(L).$$

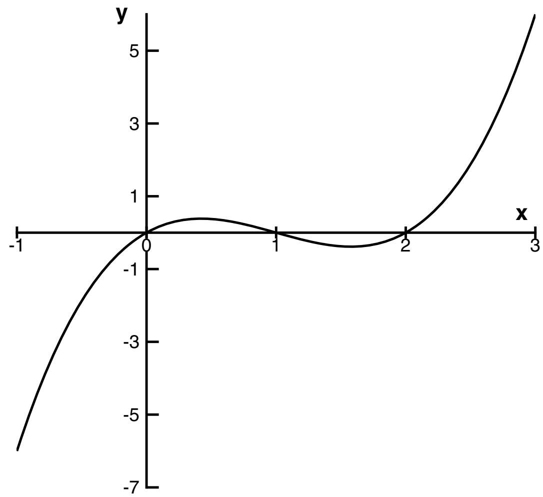

Figure 7.9 The graph of the function $f(x) = x^3 - 3x^2 + 2x$ plotted for values of x lying between -1 and 3.

The marginal product of labour, MP_L, is the derivative of the output with respect to labour and is defined by

$$MP_L = \frac{dQ}{dL} = Q'(L). \tag{7.1}$$

Under the assumption that production depends on labour alone, it is possible to calculate the size of the workforce that maximizes production. The following example illustrates this process.

Example 7.2

A firm's short run production function is given by

$$Q = 6L^2 - 0.2L^3$$

where L denotes the number of workers.

1. Find the size of the workforce that maximizes output and hence sketch a graph of this production function.
2. Find the size of the workforce that maximizes the average product of labour. Calculate MP_L and AP_L at this value of L. What do you observe?

Solution.

1. To solve the first part of this example, it is necessary to determine and classify the stationary points of the production function.

   Step 1. At a stationary point of the production function

   $$\frac{dQ}{dL} = 12L - 0.6L^2 = 0.$$

   Therefore

   $$L(12 - 0.6L) = 0$$

   and so $L = 0$ or $L = 12/0.6 = 20$.

   Step 2. It is obvious on economic grounds that $L = 0$ gives the minimum $Q = 0$. We can, of course, check this by differentiating a second time to get

   $$\frac{d^2Q}{dL^2} = 12 - 1.2L.$$

   When $L = 0$,

   $$\frac{d^2Q}{dL^2} = 12 > 0,$$

   which confirms that $L = 0$ gives a minimum for Q.
   When $L = 20$,

   $$\frac{d^2Q}{dL^2} = 12 - 24 = -12 < 0,$$

   thus $L = 20$ gives a maximum for Q.

   The firm should therefore employ 20 workers to achieve a maximum output

   $$Q = 6(20)^2 - 0.2(20)^3 = 800.$$

   The graph of this production function is sketched in Fig. 7.10.

2. To solve the second part of the problem, it is necessary to determine and classify the stationary point of the **average product of labour**, AP_L, which is defined by

   $$AP_L = \frac{Q}{L}. \tag{7.2}$$

   This is sometimes called **labour productivity** since it measures the average output per worker. In this example,

   $$AP_L = \frac{6L^2 - 0.2L^3}{L} = 6L - 0.2L^2.$$

   Step 1. At a stationary point

   $$\frac{d(AP_L)}{dL} = 0,$$

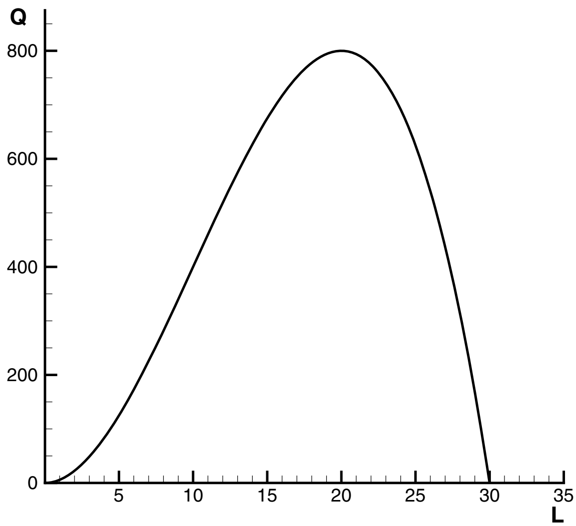

Figure 7.10 The graph of the production function $Q = 6L^2 - 0.2L^3$.

i.e.,

$$6 - 0.4L = 0,$$

and therefore $L = 6/0.4 = 15$.

Step 2. To classify this stationary point, we differentiate a second time to get

$$\frac{d^2(AP_L)}{dL^2} = -0.4 < 0$$

which shows that it is a maximum. The labour productivity is therefore greatest when the firm employs 15 workers. The corresponding labour productivity is

$$AP_L = 6(15) - 0.2(15)^2 = 45.$$

So the largest number of goods produced per worker is 45.

To find an expression for MP_L, we need to differentiate Q with respect to L, which we have already done in the first part of the problem. We have

$$MP_L = \frac{dQ}{dL} = 12L - 0.6L^2.$$

When $L = 15$,

$$MP_L = 12(15) - 0.6(15)^2 = 45.$$

We observe that at $L = 15$, the values of MP_L and AP_L are equal.

In this example, we have discovered that:

> At the point of maximum average product of labour,
> **marginal product of labour = average product of labour,**
> i.e., $MP_L = AP_L$.

In fact, this result holds for any production function $Q = Q(L)$ as we shall demonstrate. If we differentiate the expression (7.2) defining the average product of labour using the quotient rule, we obtain

$$\begin{aligned}\frac{d(AP_L)}{dL} &= \frac{d(Q/L)}{dL}\\ &= \frac{L\frac{dQ}{dL} - Q\frac{dL}{dL}}{L^2}\\ &= \frac{Q'(L) - Q(L)/L}{L}\\ &= \frac{MP_L - AP_L}{L}. \qquad (7.3)\end{aligned}$$

At a stationary point for the average product of labour, we have

$$\frac{d(AP_L)}{dL} = 0. \qquad (7.4)$$

This means that $MP_L = AP_L$, as required. This result shows that at a stationary point of the average product of labour, the marginal product of labour is equal to the marginal product of labour. The analysis has shown that this result is true for any function AP_L and is not restricted to certain choices.

At a stationary point of the average product of labour, we can obtain a simple expression for the second derivative of AP_L with respect to L as follows:

$$\begin{aligned}\frac{d^2(AP_L)}{dL^2} &= \frac{d}{dL}\left(\frac{MP_L - AP_L}{L}\right)\\ &= \frac{L\left(\frac{d(MP_L)}{dL} - \frac{d(AP_L)}{dL}\right) - (MP_L - AP_L)\frac{dL}{dL}}{L^2}.\end{aligned}$$

At a stationary point, we know $MP_L = AP_L$ and also $d(AP_L)/dL = 0$. Therefore,

$$\begin{aligned}\frac{d^2(AP_L)}{dL^2} &= \frac{1}{L}\frac{d(MP_L)}{dL}\\ &= \frac{1}{L}\frac{d^2Q}{dL^2},\end{aligned}$$

since $MP_L = dQ/dL$. So at a stationary point of the average product of labour

$$\boxed{\frac{d^2(AP_L)}{dL^2} = \frac{Q''(L)}{L}.}$$

7.7 Optimization of Profit Functions

We turn our attention to the problem of determining the maximum profit for a given firm. The profit function, π, which is the difference between the total revenue and total cost functions, is expressed as a function of the output Q. The function is then optimized with respect to Q.

Example 7.3

Maximize the profit for a firm, given that its total revenue function is given by $TR = 4{,}000Q - 33Q^2$ and its total cost function by $TC = 2Q^3 - 3Q^2 + 400Q + 5{,}000$, assuming $Q > 0$.

Solution. The profit function is given by

$$\begin{aligned} \pi &= TR - TC \\ &= 4{,}000Q - 33Q^2 - (2Q^3 - 3Q^2 + 400Q + 5{,}000) \\ &= -2Q^3 - 30Q^2 + 3{,}600Q - 5{,}000 \end{aligned}$$

Step 1. At a stationary point of the profit function,

$$\frac{d\pi}{dQ} = 0.$$

Now

$$\begin{aligned} \frac{d\pi}{dQ} &= -6Q^2 - 60Q + 3{,}600 \\ &= -6(Q^2 + 10Q - 600) \\ &= -6(Q + 30)(Q - 20). \end{aligned}$$

Therefore, the stationary points of the profit function are $Q = -30$ or $Q = 20$. (As an alternative to factorization, the equation

$$\frac{d\pi}{dQ} = 0$$

can also be solved using the quadratic formula (3.8).)

Step 2. The stationary point $Q = -30$ has no economic significance since π is only defined for $Q > 0$. Therefore, we can ignore it. To classify the second stationary point, we differentiate a second time

$$\frac{d^2\pi}{dQ^2} = -12Q - 60.$$

When $Q = 20$,

$$\frac{d^2\pi}{dQ^2} = -240 - 60 = -300 < 0,$$

which shows that π has a local maximum when $Q = 20$. Therefore, the profit is maximized when $Q = 20$ and the maximum profit is given by

$$\begin{aligned} \pi &= -2(20)^3 - 30(20)^2 + 3{,}600(20) - 5{,}000 \\ &= -16{,}000 - 12{,}000 + 72{,}000 - 5{,}000 \\ &= 39{,}000. \end{aligned}$$

Example 7.4

The demand equation for a good is

$$P + Q = 30$$

and the total cost function is

$$TC = \frac{1}{2}Q^2 + 6Q + 7.$$

1. Find the level of output that maximizes total revenue.
2. Find the level of output that maximizes profit. Calculate MR and MC at this value of Q. What do you observe?

Solution.

1. The total revenue function is defined by $TR = P \times Q$. Now $P = 30 - Q$ by rearranging the demand equation. Therefore,

   $$TR = (30 - Q)Q = 30Q - Q^2.$$

   Therefore,

   $$\frac{d(TR)}{dQ} = 30 - 2Q.$$

   Step 1. At a stationary point of the total revenue function

   $$\frac{d(TR)}{dQ} = 0,$$

so

$$30 - 2Q = 0.$$

Therefore $Q = 15$.

Step 2. To classify this point, we differentiate a second time to get

$$\frac{d^2(TR)}{dQ^2} = -2 < 0,$$

so TR has a local maximum when $Q = 15$.

2. The profit function is defined by

$$\begin{aligned}\pi &= TR - TC \\ &= (30Q - Q^2) - (\frac{1}{2}Q^2 + 6Q + 7) \\ &= -\frac{3}{2}Q^2 + 24Q - 7.\end{aligned}$$

Therefore,

$$\frac{d\pi}{dQ} = -3Q + 24.$$

Step 1. At a stationary point of the profit function

$$\frac{d\pi}{dQ} = 0,$$

so

$$-3Q + 24 = 0,$$

which has the solution $Q = 8$.

Step 2. To classify the stationary point, we differentiate the profit function a second time to get

$$\frac{d^2\pi}{dQ^2} = -3 < 0,$$

so π has a local maximum at $Q = 8$. Now

$$MR = \frac{d(TR)}{dQ} = 30 - 2Q,$$

and

$$MC = \frac{d(TC)}{dQ} = Q + 6.$$

Therefore, when $Q = 8$, then $MR = 14$ and $MC = 14$. So then

At the value of Q that maximizes profit,
marginal revenue = marginal cost

This result is true for any profit function irrespective of the market conditions under which the firm operates since at a stationary point for the profit function, we have

$$\frac{d\pi}{dQ} = \frac{d(TR)}{dQ} - \frac{d(TC)}{dQ} = MR - MC = 0.$$

Therefore, $MR = MC$ at a stationary point for the profit function.

7.8 Other Examples

Example 7.5

The cost of building an office block, x floors high, comprises three components:

1. £18 million for the land,
2. £200,000 per floor,
3. specialized costs of £20,000x per floor. (Thus if there are to be 4 floors, the specialized cost per floor will be £80,000.)

How many floors should the block contain if the average cost per floor is to be minimized?

Solution. First of all, we need to derive an expression for the total cost of construction of the office block. Suppose that the building has x floors. Then the £18 million is a fixed cost because it is independent of the number of floors. The total cost involved in the second component is £200,000x. In addition, there are specialized costs of £20,000x per floor. So if there are x floors, the specialized costs will be

$$(20{,}000x)x \;=\; 20{,}000x^2.$$

The total cost of construction in terms of monetary units of £1,000 is therefore

$$TC \;=\; 18{,}000 + 200x + 20x^2.$$

The average cost per floor, AC, is formed by dividing the total cost by the number of floors, i.e.,

$$\begin{aligned} AC &= \frac{TC}{x} \\ &= \frac{18{,}000 + 200x + 20x^2}{x} \\ &= \frac{18{,}000}{x} + 200 + 20x \end{aligned}$$

Step 1. At a stationary point

$$\frac{d(AC)}{dx} = 0.$$

Now

$$\frac{d(AC)}{dx} = -18{,}000x^{-2} + 20.$$

So we need to solve the equation

$$-18{,}000x^{-2} + 20 = 0.$$

Therefore,

$$\begin{aligned} 20 &= 18{,}000x^{-2} = \frac{1{,}800}{x^2} \\ 1 &= \frac{900}{x^2} \text{ (Divide both sides by 20.)} \\ x^2 &= \frac{900}{x^2}x^2 \text{ (Multiply both sides by } x^2\text{.)} \\ x^2 &= 900 \end{aligned}$$

Therefore $x^2 = 900$ and so $x = \pm\sqrt{900} = \pm 30$.
Step 2. To confirm that $x = 30$ yields a minimum, we need to differentiate a second time.

$$\frac{d^2(AC)}{dx^2} = 36{,}000x^{-3}$$

So obviously when $x = 30$,

$$\frac{d^2(AC)}{dx^2} > 0.$$

Thus $x = 30$ gives a minimum for AC.

Therefore an office block 30 floors high produces the lowest average cost per floor.

Example 7.6

The supply and demand equations of a good are

$$\begin{aligned} P &= Q_s + 8, \\ P &= -3Q_d + 80, \end{aligned}$$

respectively. The government decides to impose a tax, €t, per unit of good. Find the value of t that maximizes the government's total tax revenue on the assumption that equilibrium conditions prevail in the market.

Solution. To account for the imposition of tax, we replace P by $P - t$ in the supply equation. This is because the price that the supplier actually receives is the price P that the consumer pays, less the tax t deducted by the government. The new supply equation is then

$$P - t = Q_s + 8,$$

so that

$$P = Q_s + 8 + t.$$

In equilibrium,

$$Q_s = Q_d.$$

If this common value is denoted by Q then the demand and supply equations are

$$\begin{aligned} P &= -3Q + 80, \\ P &= Q + 8 + t. \end{aligned}$$

Hence,

$$Q + 8 + t = -3Q + 80.$$

Therefore

$$4Q = 72 - t,$$

and so

$$Q = 18 - \frac{1}{4}t.$$

Now if the number of goods sold is Q and the government raises t per good, then the total tax revenue is given by

$$\begin{aligned} T &= tQ \\ &= t(18 - \frac{1}{4}t) \\ &= 18t - \frac{1}{4}t^2. \end{aligned}$$

This is the function we wish to maximize.
Step 1. At a stationary point,

$$\frac{dT}{dt} = 0,$$

so

$$\frac{dT}{dt} = 18 - \frac{1}{2}t = 0,$$

which has the solution $t = 36$.
Step 2. To classify the stationary point, we differentiate a second time to get

$$\frac{d^2T}{dt^2} = -\frac{1}{2} < 0,$$

which confirms that it is a maximum. Hence the government should impose a tax of €36 on each good.

EXERCISES

7.1. For the following functions, determine the stationary point(s) and classify them. Use this information to sketch graphs of these functions.

a) $f(x) = 2x^2 - x + 6.$

b) $f(x) = x^2 - 4x + 3.$

c) $f(x) = x^3 - 3x + 3.$

d) $f(x) = 1 - 9x - 6x^2 - x^3.$

7.2. The demand equation for a good is given by

$$P + 4Q = 96,$$

and the total cost function TC is

$$TC = Q^3 - 13Q^2 + 48Q + 17.$$

a) Find the level of output that maximizes total revenue.

b) Find the maximum profit and the level of output for which it is achieved.

c) Sketch the graph of profit against Q, for $Q \geq 0$.

7.3. The demand equation for a good is given by

$$P + 2Q = 20,$$

and the total cost function TC is

$$TC = Q^3 - 8Q^2 + 20Q + 2.$$

a) Find the level of output that maximizes total revenue.

b) Find the maximum profit and the level of output for which it is achieved. Verify that, for this value of Q, $MR = MC$.

7.4. The prevailing market price for a good is 30. The total cost function is

$$TC = 100 + 44Q - 5Q^2 + \frac{1}{2}Q^3.$$

What is the level of output that maximizes the profit?

7.5. Find the first and second order derivatives, with respect to L, of the short run production function

$$Q = 15L^2 - 2L^3.$$

Hence, determine and classify the stationary points of this function.

7.6. A firm's short run production function is given by

$$Q = 12L^2 - \frac{1}{2}L^3,$$

where L denotes the number of workers.

a) Find the size of the workforce that maximizes output and hence sketch a graph of this production function.

b) Find the size of the workforce that maximizes the average product of labour, AP_L. Calculate MP_L and AP_L at this value of L. What do you observe?

7.7. The cost of building an office block, x floors high, is made up of three components:

a) \$20.16 million for the land,

b) \$175,000 per floor,

c) specialized costs of \$35,000x per floor.

How many floors should the block contain if the average cost per floor is to be minimized?

7.8. The supply and demand equations for a good are

$$Q_d = 500 - 9P,$$

and

$$Q_s = -100 + 6P,$$

respectively. The government decides to impose a tax, t per unit of good. Find the value of t that maximizes the government's total tax revenue on the assumption that equilibrium conditions prevail in the market. For this level of tax find:

a) the equilibrium price,

b) the equilibrium quantity,

c) the total tax raised.

8
Partial Differentiation

8.1 Introduction

Economic models that we have encountered so far have assumed that a quantity under consideration depends only on the value of one variable; i.e., the quantity is a function of one variable. For example, $Q = 100 - 5P$, the demand equation (or demand function) for some good describes a model where the demand Q depends only on the price P of the good. In practice, Q will depend on other variables such as consumer income or the price of a substitutable good. To take into account all variables affecting the value of Q would make an economic model too difficult to analyse or use. Useful models should lend themselves readily to analysis, perhaps with the aid of computers, while at the same time give a reasonably accurate model of the real situation.

The profit function of a firm producing only one good is of the form $y = f(x)$, where x is the output of the good. The derivative $\frac{dy}{dx}$ gives a measure of the rate of change of profit y relative to output and is itself a function of x. The function f can be visualised geometrically as a graph and the slope of the tangent at any point on the graph is the value of $\frac{dy}{dx}$ at that point.

Suppose now that the firm produces two goods G_1 and G_2 with outputs x_1 and x_2, respectively. We expect the profit function f now to depend on x_1 and x_2. This is an example of a function of two variables. The analysis of the rate of change of f relative to x_1 and x_2 is done by considering the rate of change of f relative to one variable while the other is assumed constant and vice versa. This is the concept of a partial derivative.

The utility of a consumer of the two goods G_1 and G_2 will also be a function of x_1 and x_2. Partial derivatives enable us to analyse the marginal effects of keeping, say, the output x_1 of G_1 at some fixed level while changing slightly the output x_2 of G_2.

8.2 Functions of Two or More Variables

An expression such as $f(x,y) = 3x^2+xy+y+1$ is a function of the two variables x and y. The function notation is a natural extension of that for functions of one variable. Thus, $f(\alpha, \beta)$ denotes the value of the function when $x = \alpha$ and $y = \beta$. For example, for the function defined above:

$$f(1,1) = 3+1+1+1 = 6, \qquad f(0,0) = 0+0+0+1 = 1$$

and

$$f(2,-1) = 3 \times 4 + 2 \times (-1) + (-1) + 1 = 10.$$

We can simply write f for $f(x,y)$ if there is no need to mention x and y.

These ideas extend in a natural way to functions of more than two variables. For example, $g(x,y,z) = x^2 + yz - 2z^2 + 8$ is a function of the three variables x, y, and z. The value of g when say $x = 5$, $y = -2$ and $z = 3$ is $5^2 + (-2) \times 3 - 2 \times 3^2 + 8 = 25 - 6 - 18 + 8 = 9$. This is more succinctly expressed in function notation by $g(5,-2,3) = 9$.

If we let $z = f(x,y)$, then the two variables x, y are said to be the **independent variables** and z is the **dependent variable** as its value depends on the values of x and y. Thus z may be the value of some measurement or observation depending on the values of x and y; for instance, production Q depends on the values of capital K and labour L in a simple production model (though in more advanced models, Q will depend on additional input variables).

A useful notational device is the short-hand way of expressing, for example, that 'Q is regarded as a function of K and L' by writing simply: '$Q(K,L)$' for Q. Then if say $Q = 5K^{\frac{1}{2}}L^{\frac{1}{3}}$, writing $Q(9,8) = 30$ is a quick way of saying '$Q = 30$ when $K = 9$ and $L = 8$'. In this notation, for instance, we have: $Q(4,1) = 10,\ Q(4,27) = 30$ and so on.

8.3 Partial Derivatives

Let $f(x,y)$ be any function of the two variables x and y. Then f may be regarded as a function of one variable x if we were to treat y as a constant. In this case, its derivative with respect to x is called the *partial derivative* of f with respect to x, denoted by

$$\frac{\partial f}{\partial x} \text{ or } f_x.$$

Similarly the partial derivative

$$\frac{\partial f}{\partial y} \text{ or } f_y$$

is obtained by differentiating f with respect to y, treating x as constant.

More generally, if f is a function of two or more variables and x is any one of these variables, then the partial derivative f_x is obtained by differentiating f with respect to x, treating *all* the other variables as constants.

Example 8.1

Determine the partial derivatives of the following functions:

1. $f(x, y) = 2x + 5y - 3$,
2. $g(u, v) = 3u^2v$,
3. $z = x^2 + 3xy^2 + 5$,
4. $Q = 4K^{\frac{1}{2}}L^{\frac{1}{3}}$,
5. $f(x, y, z) = xy^2 - 3x^2 + 4yz - 5z^2 + 8$,
6. $z = xe^{2y}$.

Solution.

1. The partial derivatives of the function $f(x, y) = 2x + 5y - 3$ are

   $$\frac{\partial f}{\partial x} = 2 \text{ and } \frac{\partial f}{\partial y} = 5.$$

   To see why, for example, $\frac{\partial f}{\partial y} = 5$, note that when we partially differentiate with respect to y, then $2x$ and -3 are both constants and therefore their derivatives are zero.

2. For this example, recall that constant factors can be taken outside the differentiation operator. For instance, if y is a function of x, then

   $$\frac{d}{dx}(ky) = k\frac{dy}{dx}$$

if k is a constant. In particular,

$$\frac{d}{dx}(kx) = k\frac{dx}{dx} = k.$$

Thus

$$\frac{\partial g}{\partial u} = 3v \times \frac{\partial u^2}{\partial u} = 3v \times 2u = 6vu$$

since when we differentiate partially with respect to u, the factor $3v$ in $3u^2v$ is constant and $\frac{\partial u^2}{\partial u} = 2u$. Partially differentiating with respect to v, we have

$$\frac{\partial g}{\partial v} = 3u^2 \times 1 = 3u^2$$

since now $3u^2$ is constant and $\frac{\partial v}{\partial v} = 1$.

Note that we could write g_u for $\frac{\partial g}{\partial u}$ and g_v for $\frac{\partial g}{\partial v}$ in this example. The advantage with writing g_u is that it allows us to use function notation. For instance, $g_u(2,5) = 60$ expresses succinctly that the value of g_u is 60 when $u = 2$ and $v = 5$.

3. Here

$$z_x = 2x + 3y^2 \times 1 + 0 = 2x + 3y^2$$

and

$$z_y = 0 + 3x \times 2y + 0 = 6xy.$$

4. The production model described by this Cobb-Douglas function gives the output Q as a function of only K (capital input) and L (labour input).

$$\frac{\partial Q}{\partial K} = 4L^{\frac{1}{3}} \times \frac{1}{2}K^{\frac{1}{2}-1} = 2K^{-\frac{1}{2}}L^{\frac{1}{3}}$$

and

$$\frac{\partial Q}{\partial L} = 4K^{\frac{1}{2}} \times \frac{1}{3}L^{\frac{1}{3}-1} = \frac{4}{3}K^{\frac{1}{2}}L^{-\frac{2}{3}}.$$

In Example 4, we could also have used the alternative notation Q_K for $\frac{\partial Q}{\partial K}$ for instance.

If we wish to use function notation in, for instance, Example 4, then for the output Q, we can write $Q(K,L)$ initially and then replace this simply by Q whenever there's no need to specify K or L. This also extends to the partial derivatives Q_K and Q_L, so for instance,

$$Q_K(9,8) = 2 \times 9^{-\frac{1}{2}} \times 8^{\frac{1}{3}} = 2 \times \frac{1}{3} \times 2 = \frac{4}{3}.$$

5. This example is of a function of three variables.

   $f_x = y^2 - 3 \times 2x + 0 - 0 + 0 = y^2 - 6x$

   $f_y = x \times 2y - 0 + 4z - 0 + 0 = 2xy + 4z$

   $f_z = 0 - 0 + 4y - 5 \times 2z + 0 = 4y - 10z$

6. $z_x = e^{2y}$ (since e^{2y} is treated as constant here) and

   $z_y = x \times 2e^{2y} = 2xe^{2y}$.

   (Recall that the derivative of e^{ky} with respect to y is ke^{ky} for any constant k.)

8.4 Higher Order Partial Derivatives

The discussion that follows is for functions of two variables. This is for the sake of simplicity as the ideas extend to functions of more than two variables in a natural way.

Consider any function $f(x, y)$ in the variables x and y. The partial derivatives f_x and f_y are themselves functions of x and y.

We say that f_x and f_y are the **first order** partial derivatives of f. Their partial derivatives are the **second order** partial derivatives of f. They are

$$\begin{aligned}
\frac{\partial}{\partial x}(f_x) &= \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) \text{ denoted by } \frac{\partial^2 f}{\partial x^2} \text{ or } f_{xx} \\
\frac{\partial}{\partial x}(f_y) &= \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) \text{ denoted by } \frac{\partial^2 f}{\partial x \partial y} \text{ or } f_{xy} \\
\frac{\partial}{\partial y}(f_y) &= \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right) \text{ denoted by } \frac{\partial^2 f}{\partial y^2} \text{ or } f_{yy} \\
\frac{\partial}{\partial y}(f_x) &= \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right) \text{ denoted by } \frac{\partial^2 f}{\partial y \partial x} \text{ or } f_{yx}
\end{aligned}$$

The second order partial derivatives f_{xy} and f_{yx} of f are known as **cross-derivatives**.

For all functions of two variables that we shall consider, these two cross-derivatives are always equal. That is

$$f_{xy} = f_{yx}.$$

Example 8.2

Determine the first and second order partial derivatives of $z = x^2 + 3xy^2 + 5$ (see Example 8.1.3).

Solution. In Example 8.1.3, we determined

$$z_x = \frac{\partial z}{\partial x} = 2x + 3y^2 \text{ and } z_y = \frac{\partial z}{\partial y} = 6xy.$$

Therefore

$$\begin{aligned} z_{xx} &= \frac{\partial}{\partial x}(z_x) = \frac{\partial}{\partial x}(2x + 3y^2) = 2 \\ z_{xy} &= \frac{\partial}{\partial x}(z_y) = \frac{\partial}{\partial x}(6xy) = 6y \\ z_{yy} &= \frac{\partial}{\partial y}(z_y) = \frac{\partial}{\partial y}(6xy) = 6x \\ z_{yx} &= \frac{\partial}{\partial y}(z_x) = \frac{\partial}{\partial y}(2x + 3y^2) = 3 \times 2y = 6y \end{aligned}$$

Observe that the cross-derivatives z_{xy} and z_{yx} are equal.

Example 8.3

Determine the first and second order partial derivatives of $f(x, y) = 8x^2y^3$.

Solution. The first and second order partial derivatives are

$$\begin{aligned} f_x &= 8 \times 2x \times y^3 = 16xy^3, \\ f_y &= 8x^2 \times 3y^2 = 24x^2y^2, \\ f_{xx} &= 16y^3, \\ f_{yy} &= 24x^2 \times 2y = 48x^2y, \\ f_{yx} &= 48x \times y^2 = 48xy^2 = f_{xy}. \end{aligned}$$

Example 8.4

A firm's profit is given by the function

$$\pi = 800 - 4Q^2 - 5Q + 3QY - 5Y^2 + 40Y$$

where Q denotes output and Y advertising expenditure. Determine the first and second order partial derivatives of π.

Solution. The first and second order partial derivatives of π, considered as functions of Q and Y, are

$$\pi_Q = -8Q - 5 + 3Y, \quad \pi_Y = 3Q - 10Y + 40,$$
$$\pi_{QQ} = -8, \quad \pi_{YY} = -10, \quad \pi_{YQ} = 3 = \pi_{QY}.$$

8.5 Partial Rate of Change

In the case of a function $y = f(x)$ of one variable x, the derivative $\frac{dy}{dx}$ or $f'(x)$ can be thought of as the rate of change of y relative to x. (This rate of change is itself a function of x and therefore can vary with x.) Therefore, if x changes by a small increment $\triangle x$, then the corresponding change $\triangle y$ in y is, approximately,

$$\frac{dy}{dx} \times \triangle x.$$

There is an analogous situation for partial derivatives. We consider the case of a function of two variables as this extends to more variables in an obvious way.

Suppose a quantity z is a function of two variables x and y. Then $\frac{\partial z}{\partial x}$ can be regarded as the rate of change of z relative to x. That is, if x changes by a small increment $\triangle x$ and we assume all the other variables remain fixed, then the resulting change $\triangle z$ in z is approximately $\frac{\partial z}{\partial x} \times \triangle x$. So $\frac{\triangle z}{\triangle x}$ is close to $\frac{\partial z}{\partial x}$ and gets closer to this partial derivative the smaller that $\triangle x$ becomes.

Similarly, $\frac{\partial z}{\partial y} \times \triangle y$ is the approximate change in z caused by a small change $\triangle y$ in y.

If x and y change by small increments $\triangle x$ and $\triangle y$, respectively, the resulting change in z can be estimated by the following formula.

The Small Increments Formula (SIF)
$\triangle z = \frac{\partial z}{\partial x} \times \triangle x + \frac{\partial z}{\partial y} \times \triangle y$
$= z_x \triangle x + z_y \triangle y$

Notes

1. Although we have written an equals sign in the SIF, the formula is in fact an approximation. The reason is that as x changes so will $\dfrac{\partial z}{\partial x}$, in general. Similarly for y. That is, $\dfrac{\partial z}{\partial x}$ and $\dfrac{\partial z}{\partial y}$ are functions of x and y and therefore, in general, they vary as x, y vary.
2. An increment, say $\triangle x$, is negative if x decreases. So for example $\triangle x = -0.05$ means that the current value of x is reduced by 0.05.

Example 8.5

Evaluate $z = x^2 + 3y$ when $x = 5$ and $y = 8$. Using the SIF, estimate the change in z if x increases to 5.01 and y decreases to 7.98.

Solution. Since $z = x^2 + 3y$, then

$$z_x = 2x \text{ and } z_y = 3.$$

When $x = 5$ and $y = 8$ the value of z is $25 + 24 = 49$.

Now suppose x increases to 5.01 and y decreases to 7.98. That is, $\triangle x = 0.01$ and $\triangle y = -0.02$. The SIF gives an estimate for the change in z as

$$\triangle z = z_x \times (0.01) + z_y \times (-0.02).$$

Here the partial derivatives z_x and z_y are evaluated at the initial values $x = 5$ and $y = 8$. So $z_x = 2x = 10$ and $z_y = 3$.

Therefore

$$\triangle z = 10 \times (0.01) - 3 \times (0.02) = 0.04.$$

This means z increases from its initial value of 49 to approximately the value $49 + 0.04 = 49.04$ when x and y change as described. (The actual new value of z is 49.041, which is quite close to the estimate.)

Example 8.6

The profit of a company producing two goods is given by

$$Y = 80A - (0.2)A^2 + 150B - (0.1)B^2 - 200$$

where Y is profit, A is the output of good 1, and B is the output of good 2. Evaluate the profit when $A = 10$ and $B = 6$. Estimate the profit when the output of good 1 increases by 2% and that of good 2 by 3%.

Solution.

$$\frac{\partial Y}{\partial A} = 80 - (0.2) \times 2A = 80 - (0.4)A$$
$$\frac{\partial Y}{\partial B} = 150 - (0.1) \times 2B = 150 - (0.2)B$$

When $A = 10$ and $B = 6$ units of product, we have, after substituting these values and simplifying, that

$$Y = 1476.4, \ \frac{\partial Y}{\partial A} = 76 \text{ and } \frac{\partial Y}{\partial B} = 148.8.$$

If we increase production of good 1 by 2% and production of good 2 by 3%, then $\triangle A = \frac{2}{100} \times 10 = 0.2$ and $\triangle B = \frac{3}{100} \times 6 = 0.18$. Using the SIF, we can estimate $\triangle Y$ by

$$\triangle Y = \frac{\partial Y}{\partial A} \times \triangle A + \frac{\partial Y}{\partial B} \times \triangle B.$$

That is

$$\triangle Y = 76 \times 0.2 + 148.8 \times 0.18.$$

Therefore

$$\triangle Y = 41.984.$$

So Y increases by about 2.8%. (The actual change in Y is 41.973 to 3 decimal places.)

Example 8.7

In a production model, output $Q(K, L)$ is given by

$$Q = 5K^{\frac{1}{2}} L^{\frac{2}{3}}$$

where K denotes labour and L labour costs. Evaluate output Q and the marginal costs of capital and labour when $K = 4$ and $L = 8$. Estimate output if K increases to 4.1 and L decreases to 7.95.

Solution. The first order partial derivatives of Q are

$$\frac{\partial Q}{\partial K} = \frac{5}{2} K^{-\frac{1}{2}} L^{\frac{2}{3}}$$
$$\frac{\partial Q}{\partial L} = 5K^{\frac{1}{2}} \times \frac{2}{3} L^{-\frac{1}{3}} = \frac{10}{3} K^{\frac{1}{2}} L^{-\frac{1}{3}}.$$

When $K = 4$ and $L = 8$, the output $Q = 5 \times 2 \times 4 = 40$, and

$$\frac{\partial Q}{\partial K} = \frac{5}{2} \times \frac{1}{2} \times 4 = 5 \text{ and } \frac{\partial Q}{\partial L} = \frac{10}{3} \times 2 \times \frac{1}{2} = \frac{10}{3}.$$

To estimate Q when $K = 4.1$ and $L = 7.95$, use the SIF with $\triangle K = 0.1$ and $\triangle L = -0.05$ to compute

$$\triangle Q = Q_K \times \triangle K + Q_L \times \triangle L = 5 \times (0.1) + \frac{10}{3} \times (-0.05) = \frac{1}{3}.$$

Therefore, $Q(4.1, 7.95)$ is approximately $40\frac{1}{3}$. (To see how close an approximation this is, computing $Q(4.1, 7.95) = 5 \times (4.1)^{\frac{1}{2}}(7.95)^{\frac{2}{3}}$ using a calculator gives 40.328 to 3 decimal places.)

8.6 The Chain Rule and Total Derivatives

Suppose y is a function of a single variable x and that x is a function of a single variable t. Then y may be regarded as a function of the single variable t since any value to t determines x, which in turn determines y. The chain rule for functions of one variable (see (6.8)) is

$$\frac{dy}{dt} = \frac{dy}{dx}\frac{dx}{dt}.$$

For example, let $y = x^3$ and $x = t^2$. Then

$$\frac{dy}{dt} = \frac{dy}{dx} \cdot \frac{dx}{dt} = 3x^2 \times 2t = 6x^2 t.$$

In this case, one could have easily substituted for x in y to get $y = (t^2)^3 = t^6$ giving y explicitly as a function of t. Then $\frac{dy}{dt} = 6t^5$, which is the same as $6x^2t$, noting that $x^2 = t^4$. There is an analogous formula for functions of two variables.

Suppose z is a function of two variables x and y. Further, suppose that each of x and y is a function of a single variable t. Then z can be regarded as a function of the single variable t. The derivative $\frac{dz}{dt}$ is given by

Total Derivative Formula

$$\frac{dz}{dt} = \frac{\partial z}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial z}{\partial y} \cdot \frac{dy}{dt}$$

(Notice that when we write $\frac{dx}{dt}$ and not $\frac{\partial x}{\partial t}$, x is being regarded as a function of just one variable, t. Similarly for y and z.)

The derivative $\frac{dz}{dt}$ is known as the **total derivative** of z with respect to t.

The total derivative formula follows from the SIF quite easily. We have

$$\triangle z = \frac{\partial z}{\partial x} \cdot \triangle x + \frac{\partial z}{\partial y} \cdot \triangle y$$

for the incremental changes $\triangle x$, $\triangle y$, $\triangle z$ caused by a change $\triangle t$ in t. Therefore

$$\frac{\triangle z}{\triangle t} = \frac{\partial z}{\partial x} \cdot \frac{\triangle x}{\triangle t} + \frac{\partial z}{\partial y} \cdot \frac{\triangle y}{\triangle t}.$$

As $\triangle t$ tends to 0, the ratios $\frac{\triangle x}{\triangle t}$, $\frac{\triangle y}{\triangle t}$, and $\frac{\triangle z}{\triangle t}$ tend, respectively to $\frac{dx}{dt}$, $\frac{dy}{dt}$, and $\frac{dz}{dt}$ to give the total derivative formula.

A special case of the total derivative equation is when $t = x$. Then putting $t = x$ in the total derivative formula and noting that $\frac{dx}{dx} = 1$, we have

> The Total Derivative when z is a function of y and y is a function of x
> $$\frac{dz}{dx} = \frac{\partial z}{\partial x} + \frac{\partial z}{\partial y} \cdot \frac{dy}{dx}$$

Note the appearance of both $\frac{dz}{dx}$ and $\frac{\partial z}{\partial x}$ in the equation. The partial derivative $\frac{\partial z}{\partial x}$ is the rate of change of z relative to x when z is considered a function of two variables x and y; so it is implicit that y is constant when computing this partial derivative. Thus $\frac{\partial z}{\partial x}$ is the direct contribution of x to this rate of change, while $\frac{\partial z}{\partial y} \cdot \frac{dy}{dx}$ is the indirect contribution of x through y (which is a function of x).

Example 8.8

Given that

$$z = xy + 3y - 7x + 5$$

where $x = t^2$ and $y = 2t + 3$, find

1. $\frac{dz}{dt}$,
2. the value of $\frac{dz}{dt}$ when $t = 5$.

Solution.

1. Substituting for x and y in terms of z, we can express z explicitly as a function of t only. Then dz/dt is easily computed. However, it is not always easy to substitute. The total derivative equation allows us to compute dz/dt without substitution.

   We have
   $$\frac{\partial z}{\partial x} = y - 7, \frac{\partial z}{\partial y} = x + 3, \frac{dx}{dt} = 2t \text{ and } \frac{dy}{dt} = 2.$$
   Therefore
   $$\frac{dz}{dt} = \frac{\partial z}{\partial x} \cdot \frac{dx}{dt} + \frac{\partial z}{\partial y} \cdot \frac{dy}{dt} = (y-7) \times 2t + (x+3) \times 2.$$
   Finally, we obtain
   $$\frac{dz}{dt} = 2ty - 14t + 2x + 6.$$

2. When $t = 5$, then $x = 25$ and $y = 13$. Then
   $$\frac{dz}{dt} = 2 \times 5 \times 13 - 14 \times 5 + 2 \times 25 + 6 \equiv 116.$$

Example 8.9

Find dz/dx and the values of x for which $dz/dx = 0$, given that $z = 4x^2y^3$, where $y = x^2 + 3$.

Solution. Here $\dfrac{\partial z}{\partial x} = 4y^3 \times 2x = 8xy^3$, $\dfrac{\partial z}{\partial y} = 4x^2 \times 3y^2 = 12x^2y^2$, and $\dfrac{dy}{dx} = 2x$. Therefore,

$$\begin{aligned}\frac{dz}{dx} &= \frac{\partial z}{\partial x} + \frac{\partial z}{\partial y} \cdot \frac{dy}{dx} \\ &= 8xy^3 + 12x^2y^2 \times 2x \\ &= 8xy^3 + 24x^3y^2 = 8xy^2(y + 3x^2).\end{aligned}$$

Now $\dfrac{dz}{dx} = 0$ when $x = 0,\ y = 0$, or $y = -3x^2$.

Since $y = x^2 + 3$, the last two cases are, respectively, equivalent to $x^2 + 3 = 0$ and $4x^2 + 3 = 0$, neither of which is possible since each of $x^2 + 3$ and $4x^2 + 3$ is at least 3 for any value of x.

Example 8.10

A monopolist's total revenue TR is given by the formula $TR = PQ$, where P is the price of the good being supplied and Q the quantity supplied. Determine the marginal revenue MR if $P = 120 - 6Q$ is the demand function.

Solution. The total derivative formula gives

$$MR = \frac{d(TR)}{dQ} = \frac{\partial(TR)}{\partial Q} + \frac{\partial(TR)}{\partial P} \cdot \frac{dP}{dQ}.$$

Since $TR = PQ$, we have $\dfrac{\partial(TR)}{\partial Q} = P$ and $\dfrac{\partial(TR)}{\partial P} = Q$, and since $P = 120-6Q$, we have $\dfrac{dP}{dQ} = -6$.

Substituting in the expression for MR gives

$$MR = P + Q(-6) = P - 6Q.$$

Of course we could in this example easily express TR solely in terms of Q by substituting $P = 120 - 6Q$ in $TR = PQ$ to get

$$TR = (120 - 6Q)Q = 120Q - 6Q^2.$$

Then $MR = \dfrac{d(TR)}{dQ} = 120 - 6 \times 2Q = 120 - 12Q$. This is the same answer as before since $P - 6Q = (120 - 6Q) - 6Q = 120 - 12Q$.

8.7 Some Applications of Partial Derivatives

8.7.1 Implicit Differentiation

The equation $y = x^2 + 3x - 5$ gives y explicitly as a function of x. That is, y is presented as an expression in terms of x. However, an equation of the form

$$x^2y^3 - xy + 3x^3 = 5$$

relates x and y but not explicitly as it is not possible by rearranging the equation to express y just in terms of x. In this case, we say that y is implicitly a function of x.

In the first equation when y is explicitly given in terms of x, it is easy to determine the derivative of y with respect to x and it is given by $\dfrac{dy}{dx} = 2x + 3$;

but not in the second equation where y is implicitly given in terms of x. The second equation is of the form

$$f(x, y) = K$$

where $f(x, y)$ is a function of x, y and K is a constant. Using the total derivative formula, we have

$$\begin{aligned} \frac{df}{dx} &= \frac{\partial f}{\partial x} \cdot \frac{dx}{dx} + \frac{\partial f}{\partial y} \cdot \frac{dy}{dx} \\ &= \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y} \cdot \frac{dy}{dx} \\ &= f_x + f_y \cdot \frac{dy}{dx}. \end{aligned}$$

Since the value of f is constant $(= K)$, then $\frac{df}{dx} = 0$ and so $f_x + f_y \frac{dy}{dx} = 0$. Therefore

$$\boxed{\frac{dy}{dx} = \frac{-f_x}{f_y}}$$

This is the implicit differentiation formula. Observe that $\frac{dy}{dx}$ is independent of the value of the constant K.

Example 8.11

Use implicit differentiation to find dy/dx for the following:

1. $x^2 y = 3$,
2. $x^2 y^3 - xy + 3x^3 = 5$,
3. $xy = y^2 + 3x^2 + 1$.

Solution.

1. Here $f(x, y) = x^2 y$, $f_x = 2xy$, and $f_y = x^2$, so

$$\frac{dy}{dx} = -\frac{2xy}{x^2} = -\frac{2y}{x}.$$

In this case, y could have been expressed explicitly in terms of x, since $y = 3/x^2 = 3x^{-2}$.

Therefore $\frac{dy}{dx} = 3 \times (-2)x^{-3} = -6x^{-3}$, which is equal to $-2y/x$, since $y = 3x^{-2}$.

2. In this case, $f(x,y) = x^2y^3 - xy + 3x^3$, $f_x = 2xy^3 - y + 9x^2$, and $f_y = 3x^2y^2 - x$. So
$$\frac{dy}{dx} = -\frac{(2xy^3 - y + 9x^2)}{(3x^2y^2 - x)}.$$

3. Write this as $f(x,y) = y^2 + 3x^2 - xy = -1$. Then $f_x = 6x - y$ and $f_y = 2y - x$. Therefore
$$\frac{dy}{dx} = -\frac{(6x-y)}{2y-x} = \frac{y-6x}{2y-x}.$$

8.7.2 Elasticity of Demand

This is an economic model for one good and an alternative (or related) good: the demand Q for a particular good depends on its *price* P, the price P_A of the *alternative good*, and the *income* Y of consumers. Thus Q is regarded as a function of the variables P, P_A, and Y.

Examples are

1. The price of new cars is related to the price of fuel and the income of the driving population.
2. The price of domestic gas and the price of domestic electricity are related.
3. The prices of DVD players and DVDs are related.

The **own price** (or **direct price**) **elasticity of demand** E_P measures the relative percentage changes of Q and P (with P_A and Y assumed fixed).

If $\triangle Q$ is the change in Q following a change $\triangle P$ in P, the relative percentage change of Q to that of P is
$$\frac{\frac{\triangle Q}{Q} \times 100}{\frac{\triangle P}{P} \times 100} = \frac{P}{Q}\frac{\triangle Q}{\triangle P}.$$

As $\triangle P$ approaches 0, then $\dfrac{\triangle Q}{\triangle P}$ approaches $\dfrac{\partial Q}{\partial P}$; so the relative percentage change approaches $\dfrac{P}{Q}\dfrac{\partial Q}{\partial P}$.

This is therefore approximately the percentage change in Q resulting from a one percent increase in P. Since $\dfrac{\partial Q}{\partial P}$ will, in practice, be negative (demand normally decreases with increases in price), then in order to have a positive number for the own price elasticity of demand, we define:
$$E_P = -\frac{P}{Q}\frac{\partial Q}{\partial P}.$$

Therefore:

The elasticity E_P is approximately the percentage change in the demand Q resulting from a 1% *decrease* in P.

The elasticity E_P measures the sensitivity of the good to its own price.

Similarly, we can define the **cross-price elasticity of demand** E_{P_A} by

$$E_{P_A} = \frac{P_A}{Q} \cdot \frac{\partial Q}{\partial P_A}.$$

The elasticity E_{P_A} is approximately the percentage change in the demand Q for the good following a 1% *increase* in the price P_A of the alternative good.

The elasticity E_{P_A} measures the sensitivity of the good to the price of the alternative good (all else fixed).

Finally, the **income elasticity of demand** E_Y is defined by

$$E_Y = \frac{Y}{Q} \cdot \frac{\partial Q}{\partial Y}.$$

The elasticity E_Y is approximately the percentage change in the demand Q following a 1% *increase* in the income Y of consumers.

The elasticity E_Y measures the sensitivity of demand for the good to changes in the income of consumers (assuming P, P_A are fixed).

If Q increases as P_A increases, the alternative good is **substitutable** (e.g., beef and lamb). Equivalently, $\frac{\partial Q}{\partial P_A} > 0$ or $E_{P_A} > 0$. (Think of partial derivatives as rates of change.)

If Q decreases as P_A increases (equivalently $\frac{\partial Q}{\partial P_A} < 0$ or $E_{P_A} < 0$), the alternative good is **complementary**.

If $\frac{\partial Q}{\partial P_A} = 0$, the goods are unrelated (essentially Q is constant relative to P_A).

For example, computers and printers are complementary. This is because consumers who buy one will also buy the other. Therefore, the price of the pair, as a whole, becomes more expensive. It is reasonable to consider cars and pharmaceuticals as unrelated goods.

If Q increases when Y increases (equivalently $\frac{\partial Q}{\partial Y} > 0$ or $E_Y > 0$), the good is **superior** (to the alternative good).

If Q decreases when Y increases, the good is **inferior**. (Equivalently, $\frac{\partial Q}{\partial Y} < 0$ or $E_Y < 0$.)

In the two examples below, the alternative good is substitutable but the good is superior to the alternative good.

Example 8.12

Given the demand function $Q = 220 - 4P + 2P_A + \frac{Y}{50}$, find the own price, cross-price, and income elasticities of demand. Evaluate these elasticities when $P = 5, P_A = 6, Y = 1900$. What happens to demand when:

1. P decreases by 0.25%
2. P_A increases by 2%
3. Y increases by 10%?

Solution. The first order partial derivatives of Q are

$$\frac{\partial Q}{\partial P} = -4, \quad \frac{\partial Q}{\partial P_A} = 2, \quad \frac{\partial Q}{\partial Y} = \frac{1}{50}.$$

Therefore, the elasticities of demand are given by

$$E_P = \frac{4P}{Q}, \quad E_{P_A} = \frac{2P_A}{Q}, \quad E_Y = \frac{Y}{50Q}.$$

Consider the case $P = 5$, $P_A = 6$, $Y = 1{,}900$. Then $Q = 250$ and $E_P = 0.08$, $E_{P_A} = 0.048$, and $E_Y = 0.152$.

1. If price P drops by 0.25%, then demand Q rises by

$$E_P \times 0.25\% = 0.08 \times 0.25\% = 0.02\%.$$

2. If P_A increases by 8%, then Q rises by

$$E_{P_A} \times 8 = 0.048 \times 8 = 0.384\%.$$

3. If Y increases by 10%, then Q increases by

$$E_Y \times 10 = 1.52\%$$

(assuming the other variables, in this case P, P_A, are fixed – similarly for the previous two cases).

Example 8.13

For the demand function $Q = 100 - 4P^2 + 3P_A + 0.04Y^{1/2}$, find the own price, cross-price and income elasticities of demand and evaluate them when $P = 3$, $P_A = 1$, and $Y = 2{,}500$. What happens to demand when:

1. P falls or rises by 3%
2. P_A rises by 2%
3. Y rises by 10%.

Solution. We have

$$\frac{\partial Q}{\partial P} = -8P, \ \frac{\partial Q}{\partial P_A} = 3, \ \frac{\partial Q}{\partial Y} = 0.02Y^{-1/2}$$

and therefore

$$E_P = \frac{8P^2}{Q}, \ E_{P_A} = \frac{3P_A}{Q}, \ E_Y = \frac{0.02Y^{1/2}}{Q}.$$

When $P = 3$, $P_A = 1$, $Y = 2500$, then $Q = 69$. Therefore

$$E_P = \frac{24}{23}, \ E_{P_A} = \frac{1}{23}, \ E_Y = \frac{1}{69}.$$

1. If P falls or rises by 3%, then Q rises or falls by about $3 \times \dfrac{24}{23} = 3.130$ (correct to 3 decimal places).
2. If P_A rises by 2%, then Q rises by about (in fact exactly) $\dfrac{2}{23} = 0.087$ (correct to 3 decimal places).
3. If Y rises by 10%, then Q rises by about $\dfrac{10}{69} = 0.145$ (correct to 3 decimal places).

Note that in, for example, (1), the other variables, in this case P_A and Y, have not changed and are at their initial values. Also note that the percentage change in Q in (2) is actually (not approximately) $\dfrac{2}{23}$%. The other two cases are good approximations. For instance, the reader can easily verify that in (1) and (3) the actual percentage changes in Q are, respectively, 3.1774% and 0.1415% (correct to 4 decimal places).

8.7.3 Utility

Utility attempts to model a consumer's satisfaction or benefit when buying various combinations of quantities of two goods G_X and G_Y. A **utility function** $U(x, y)$ measures a consumer's satisfaction if x units of G_X and y of G_Y are consumed.

Consider the utility function:

$$U(x, y) = 3x^{1/2}y^{1/3}.$$

We have

$$\begin{aligned} U(4, 8) &= 3 \times 2 \times 2 = 12 \\ U(9, 3) &= 3 \times 3 \times 3^{1/3} \approx 12.9803 \\ U(8, 4) &= 3 \times 8^{1/2} \times 4^{1/3} \approx 13.4695 \end{aligned}$$

The consumer is more satisfied buying 8 units of G_X and 4 of G_Y than buying 9 of G_X and 3 of G_Y; or 4 of G_X and 8 of G_Y.

The **marginal utility** of good G_X is $\dfrac{\partial U}{\partial x}$ (or U_x). It is the rate of change of U relative to x (y fixed). The marginal utility $\dfrac{\partial U}{\partial x}$ is approximately the change in U when x is increased by 1 unit and y is constant. It is a measure of the additional utility that results due to the consumer buying one more unit of the good G_X. Similarly, we may define the marginal utility $\dfrac{\partial U}{\partial y}$ of good G_Y.

The SIF gives the *approximate* change ΔU in U if x and y both change, by amounts Δx, Δy, respectively, as

$$\Delta U = \frac{\partial U}{\partial x}\Delta x + \frac{\partial U}{\partial y}\Delta y.$$

Suppose we wish to keep the utility value U at some constant value, say k. Then x, y satisfy the equation

$$U(x, y) = k.$$

This defines y implicitly as a function of x. Then by using implicit differentiation, we can compute the derivative $\dfrac{dy}{dx}$, the rate of change of y relative to x. Specifically

$$\frac{dy}{dx} = -\frac{\partial U}{\partial x} \Bigg/ \frac{\partial U}{\partial y}$$

if $U(x, y) = k$, where k is any constant. (Observe that $\frac{dy}{dx}$ does not depend on the value of k.)

The **marginal rate of commodity substitution** ($MRCS$) is defined by

$$MRCS = -\frac{dy}{dx} = \frac{\partial U}{\partial x} \Bigg/ \frac{\partial U}{\partial y}.$$

The minus sign is introduced to ensure that, in general, $MRCS$ is positive. Therefore

$$\boxed{MRCS = \frac{\text{Marginal utility of } x}{\text{Marginal utility of } y}}$$

Since $MRCS = -\dfrac{dy}{dx}$, where $U(x, y) = k$, where k is some constant, it follows that, approximately

> $MRCS$ is the change in y that maintains the value of $U(x, y)$ following a unit *decrease* in x.

The $MRCS$ reflects how much a consumer is willing to give up of good G_X in exchange for more of G_Y and be as satisfied as before.

Example 8.14

For the utility function

$$U(x, y) = 3x^{1/2}y^{1/3}$$

find the marginal utilities and a simplified expression for $MRCS$ in terms of x and y. Evaluate U and $MRCS$ when $x = 100$ and $y = 27$. Hence estimate the increase in y required to maintain the current level of utility when x decreases by 1.5 units.

Solution. The marginal utilities are

$$\frac{\partial U}{\partial x} = 3 \times \frac{1}{2}x^{-1/2}y^{1/3} = (1.5)x^{-1/2}y^{1/3}$$

and

$$\frac{\partial U}{\partial y} = 3x^{1/2} \times \frac{1}{3}y^{-2/3} = x^{1/2}y^{-2/3}.$$

Therefore,

$$\begin{aligned} \text{MRCS} &= \frac{(1.5)x^{-1/2}y^{1/3}}{x^{1/2}y^{-2/3}} \\ &= (1.5)x^{-1}y \end{aligned}$$

(using the quotient rule for indices (1.12).

When for example $x = 100$ and $y = 27$, then $U = U(100, 27) = 3 \times 10 \times 3 = 90$ and $MRCS = 1.5 \times 100^{-1} \times 27 = 0.405$.

If x decreases by 1.5 (so the 'new' x is 98.5), then, to maintain the value of U at 90, y must change approximately by $1.5 \times MRCS = 1.5 \times 0.405 = 0.6075$. So the 'new' $y = 27.6075$ to maintain $U = 90$.

The reader can check how accurate this is by computing $U(98.5, 27.6075)$, which equals 90.0558 (correct to 4 decimal places).

The **law of diminishing marginal utility** states that eventually the marginal utility of a good G_X decreases as x increases. This means that for G_X, for instance, eventually $\frac{\partial^2 U}{\partial x^2} < 0$, as x increases (since $\frac{\partial^2 U}{\partial x^2}$ is the rate of change of $\frac{\partial U}{\partial x}$ relative to x).

In the above example, for the good G_X we have: $\frac{\partial U}{\partial x} = (1.5)x^{-1/2}y^{1/3}$ and so

$$\frac{\partial^2 U}{\partial x^2} = (1.5) \times (-\tfrac{1}{2})x^{-3/2}y^{1/3} = -(0.75)x^{-3/2}y^{1/3}$$

which is negative for all positive x and y.

8.7.4 Production

In an economic model for production, the output $Q(K, L)$ is considered a function only of capital K (buildings, tools, machinery, etc.) and labour L (paid work for the production process) costs. The function Q is known as the production function. The mathematical analysis of Q is similar to that for utility.

The **marginal product of capital** MP_K is defined as $\dfrac{\partial Q}{\partial K}$, while the **marginal product of labour** MP_L is defined as $\dfrac{\partial Q}{\partial L}$. Thus MP_K may be regarded as the rate of change of output relative to capital, assuming labour costs remain constant. Approximately, MP_K is the change of output Q if K increases by 1 unit and L is fixed. (The bigger K is relative to one unit of capital, the better the approximation.) Similarly for MP_L.

Suppose output Q is required to remain fixed at a constant level c. Then the equation $Q(K, L) = c$ defines K as an implicit function of L. As in the analysis of utility, using implicit differentiation we have

$$\frac{dK}{dL} = -\frac{\partial Q}{\partial L} \bigg/ \frac{\partial Q}{\partial K} = -\frac{MP_L}{MP_K}.$$

This gives the rate of change of K relative to L. That is, if L changes by a small amount ΔL, the corresponding change in K is

$$\Delta K \approx \frac{dK}{dL} \times \Delta L.$$

In practice for the type of function Q used to model productivity, a decrease in labour ($\Delta L < 0$) requires an increase in capital ($\Delta K > 0$) to maintain the value of Q at a constant level. Therefore $\dfrac{dK}{dL}$ is normally negative. The **marginal rate of technical substitution** ($MRTS$) is defined as $-\dfrac{dK}{dL}$, so that it is, in general, positive.

To sum up

> The marginal rate of technical substitution,
> $$MRTS = -\frac{dK}{dL} = \frac{\partial Q}{\partial L} \bigg/ \frac{\partial Q}{\partial K} = MP_L \,/\, MP_K,$$
> is the approximate change in K needed to maintain the value of Q if L *decreases* by one unit.

Example 8.15

Given the production function $Q(K, L) = K^2 + 2K + 3L^2$, evaluate MP_K and MP_L for $K = 3$, $L = 1.5$. Hence,

1. Write down the value of $MRTS$;
2. Estimate the increase in capital needed to maintain the current level of output given a 0.08 of a unit decrease (or increase) in labour.

Solution.

1. We have

$$MP_L = \frac{\partial Q}{\partial L} = 6L,$$
$$MP_K = \frac{\partial Q}{\partial K} = 2K + 2,$$
$$MRTS = \frac{MP_L}{MP_K} = \frac{6L}{2K+2} = \frac{3L}{K+1}.$$

   When $K = 3$ and $L = 1.5$, then $Q(3, 1.5) = 9 + 6 + 3 \times 2.25 = 21.75$ and $MP_L = 9$, $MP_K = 8$, $MRTS = 9/8$.

2. If L is decreased (increased) by 0.08 units, then, to maintain value of Q at 21.75, K must increase (decrease) by $0.08 \times MRTS = 0.09$, approximately. The reader is left to evaluate the values $Q(3.09, 1.42)$ and $Q(2.92, 1.58)$ to see how close they are to 21.75.

   (Note that if L increases by 0.08 units and the value of K stays at 3, then Q would increase in value by approximately $MP_L \times \Delta L = 9 \times 0.08 = 0.72$. The decrease of 0.09 for K computed using the $MRTS$ is approximately that needed to decrease Q by the same amount 0.72. That is, $MP_K \times \Delta K = 8 \times (-0.09) = -0.72 = \Delta Q$.)

Example 8.16

Given the production function $Q(K, L) = 5K^{1/3}L^{1/2}$, evaluate Q, MP_K, MP_L, and $MRTS$ for the case $K = 8$, $L = 9$. Estimate the value of K that will maintain the current output if L is decreased by 1 unit.

Solution. We have

$$MP_L = \frac{\partial Q}{\partial L} = 5K^{1/3} \times \frac{1}{2}L^{-1/2} = \frac{5}{2}K^{1/3}L^{-1/2}$$
$$MP_K = \frac{\partial Q}{\partial K} = 5 \times \frac{1}{3}K^{-2/3}L^{1/2} = \frac{5}{3}K^{-2/3}L^{1/2}$$
$$MRTS = MP_L \;/\; MP_K = 3KL^{-1}$$

When $K = 8$ and $L = 9$, then $Q = 30$, $MP_L = 5/3$, $MP_K = 5/4$, and $MRTS = MP_L \;/\; MP_K = 4/3$.

If L is decreased by 1 unit so that $L = 8$, then to maintain $Q = 30$, the value of K must increase by $1 \times MRTS = 4/3$ units. That is, $K = 9\frac{1}{3}$, approximately. (Of course, we could calculate the exact K from the equation $30 = 5K^{1/3}L^{1/2}$ with $L = 8$. That is $30 = 5K^{1/3} \times 8^{1/2}$, so that $K^{1/3} = 6 \div 8^{1/2}$ and therefore $K = 6^3 8^{-3/2} \approx 9.5459$.)

8.7.5 Graphical Representations

If $z = f(x, y)$ is a function of two variables x and y, then as in the case of functions of one variable, we can plot the graph in three dimensions using three axes Ox, Oy, Oz at right angles to each other. The result in general is a surface.

Drawing or visualising such surfaces is not always easy. A simpler approach is to consider the plane sections of this surface perpendicular to the Oz axis. To explain this, consider a fixed particular value of z, say k. That is, $f(x, y) = k$. This equation defines y implicitly as a function of x.

We can plot in two dimensions the graph of this function, consisting of all points (x, y) satisfying $f(x, y) = k$. As k varies, the curves $f(x, y) = k$ give a series of parallel nested curves as in Fig. 8.1 for a production function. Each curve has an equation of the form $Q = k$, where k is constant (for that curve).

In the case of utility functions, these curves are called **indifference curves** because all combinations of x and y on a particular curve $U = k$ result in the same value for U, namely k. If for example $U(x, y) = 2x^{1/2}y^{1/3}$, then the points (9,8), (36,1) lie on the same indifference curve $U = 12$. So the consumer is indifferent whether he or she were to buy 9 units of good G_x and 8 of good G_y, or 36 units of good G_x and 1 unit of good G_y. That is, it is assumed the consumer is satisfied with either combination of goods, or indeed with any combination on the same indifference curve.

For a given indifference curve with equation $U = k$, implicit differentiation, as we saw earlier, gives $dy/dx = -U_x/U_y = -MRCS$. This gives the slope of the tangent at any point on the indifference curve. Note that the $MRCS$ is minus dy/dx. The minus is to make $MRCS$ positive, since normally the tangent slope is negative for utility functions.

For a production function $Q(K, L)$, the curves $Q = k$ for different values of k are called **isoquants**. At a point (x, y) of an isoquant $Q = k$, the tangent slope gives $-MRTS$ for those value of x and y. In Figs. 8.1 and 8.2, we show isoquants for the production functions of Examples 8.15 and 8.16, respectively.

Graphical representations are discussed further in Section 9.4.

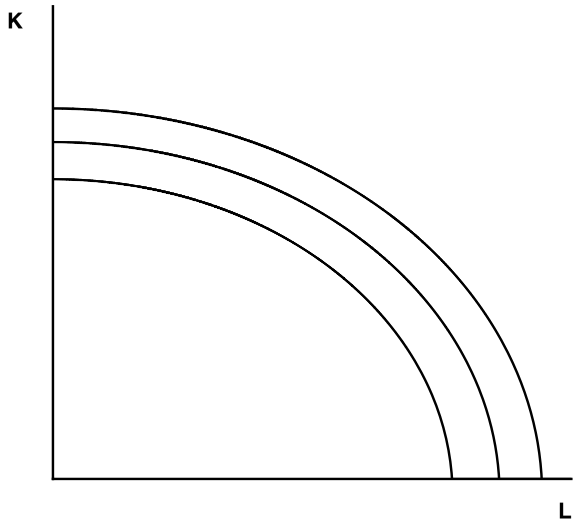

Figure 8.1 Isoquants of the production function $Q = K^2 + 2K + 3L^2$ of Example 8.15.

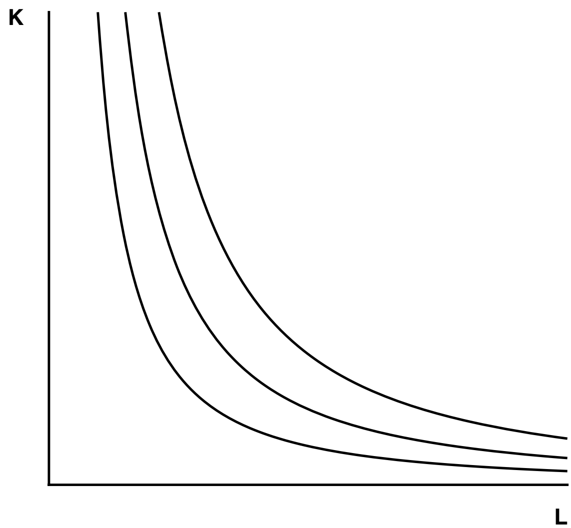

Figure 8.2 Isoquants of the production function $Q = 5K^{1/3}L^{1/2}$ of Example 8.16.

EXERCISES

8.1. Find all first and second order partial derivatives of the functions

a) $2x + y$

b) $\dfrac{x}{y}$

c) $3x^2 + y$

d) $\sqrt{xy}$

e) $3x^2 + y^2$

f) xy^2

g) xe^y

h) e^{x+y}

i) $e^x + e^y$

j) $y\ln(2x)$

k) $\ln(5xy)$.

8.2. Let $f(x, y) = 10x^{2/5}y^{1/2}$.

a) Find f_x and f_y.

b) Evaluate $f(32, 9)$, $f_x(32, 9)$, $f_y(32, 9)$.

c) Estimate the value of $f(32.1, 8.95)$.

d) Compute the actual value of $f(32.1, 8.95)$.

8.3. Let $z = x^2 - 4xy + 5$.

Evaluate z when $x = 1.5$, $y = 1$.

Estimate the percentage change in z if x is increased by 10% and y *decreased* by 5%.

8.4. If $f(x, y) = x^2 - 4xy + 3y^2 - y + 8$, determine x and y if $f_x = f_y = 0$.

8.5. Given that $3x^2 - 2y^2 + 4xy + 7y + 2 = 0$, use implicit differentiation to find $\dfrac{dy}{dx}$.

8.6. (a) Find the total derivative $\dfrac{dz}{dt}$ when $z = xy^2$, $y = 3t^2$, $x = t^2 + 3$.

Evaluate $\dfrac{dz}{dt}$ when $t = 2$.

(b) Find $\frac{dz}{dx}$ if $z = x^2 + 2xy + 3y + 5$ and $y = x^2$. Evaluate $\frac{dz}{dx}$ when $x = 1$.

8.7. Find the price, cross-price, and income elasticities of demand when $P = 10$, $P_A = 15$, and $Y = 2{,}000$, given the demand function is

$$Q = 80 - 3P + 2P_A + (0.2)Y.$$

What is the percentage increase in demand if

(a) income rises by 10%;

(b) the price P drops by 20%?

8.8. Determine (as functions of x, y) the marginal utilities and $MRCS$ of the utility function $U = 5x^{1/4}y^{2/3}$.

(a) Evaluate: U, the marginal utilities and $MRCS$, when $x = 16$, $y = 8$.

(b) *Estimate* U when $x = 16.1$ and $y = 7.5$ using the small increments formula.

(c) *Estimate* the value of y that would maintain the value of U computed in (a) if x is decreased to 14. Evaluate U for this value of y and $x = 14$ to check your answer.

9
Optimization

9.1 Introduction

Optimization is a concept of prime importance in economic analysis. Companies endeavour to maximize profit and minimize costs. Governments hope to minimize unemployment and inflation while maximizing tax revenue. Consumers are assumed to want to obtain maximum utility (satisfaction or benefit) from their consumption of particular products.

In simplified models, the optimization is unconstrained. This can be of theoretical interest, but, in practice, optimization is constrained. For instance, a firm tries to maximize profit subject to constraints on costs. A government may try to minimize interest rates while trying to keep inflation at a certain level. A pilot may fly an aircraft so as to cover the maximum possible air miles when the total fuel cost is stipulated. Consumers try to maximize utility subject to a given budget.

In this chapter, we will describe techniques of optimization when there are no constraints specified (unconstrained optimization) and subject to a constraint (constrained optimization).

9.2 Unconstrained Optimization

The optimization of functions of one variable was discussed in Chapter 7. Optimization there meant finding the **stationary** (or **critical** or **turning**) points of the function and then testing to see whether they were maxima or minima. Maxima (or minima) are points where the function changes from being increasing to decreasing (or vice versa).

For functions of two or more variables, the tests are more complicated, but the case of two variables is the easiest to handle. It is mostly with this case that we shall be concerned in this chapter.

A function $f(x, y)$ of two variables x, y is said to have a stationary (or critical) point where $x = x_0$ and $y = y_0$ if the first order partial derivatives of f are both zero for these values. That is

$$f_x(x_0, y_0) = f_y(x_0, y_0) = 0.$$

More generally, a function of two or more variables has a stationary point where all its first order partial derivatives are zero.

Example 9.1

Find the stationary point(s) of the following function:

$$f(x, y) = 3x^2 + y^2 + 4x - 4y + 7.$$

Solution. The first order partial derivatives of f are

$$f_x = 6x + 4, \qquad f_y = 2y - 4.$$

The stationary points are where $f_x = f_y = 0$. That is, $x = -\frac{4}{6} = -\frac{2}{3}$ and $y = 2$. There is therefore only one stationary point: it is given by $x = -\frac{2}{3}, y = 2$.

The use of the word 'point' as in 'stationary point' suggests that this concept can be viewed geometrically. We can think of the function f as represented by its three-dimensional graph consisting of all the points (x, y, z), where $z = f(x, y)$. The graph is, in general, a surface at each point of which there is a tangent plane, which is 'horizontal' at a stationary point. That is, it is at right angles to the z-axis and is therefore parallel to the plane containing the x and y axes.

Our main interest is in the local minimum and maximum points. These are collectively known as extrema and occur where the geometric surface representing f is, respectively, the top or the bottom of a bowl-like section of the surface (see Figs. 9.1 and 9.2, respectively). At such points, all first order partial derivatives (f_x, f_y) are necessarily zero (This is an extension of the case of a

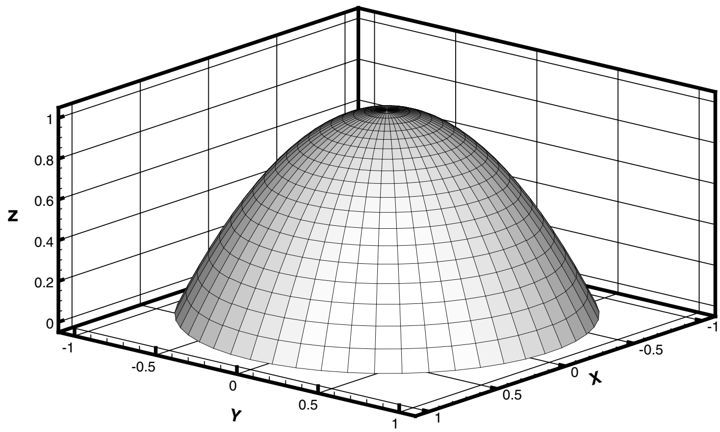

Figure 9.1 The graph of the function $z = 1 - x^2 - y^2$. The function has a local maximum at the point $x = y = 0$, $z = 1$.

function of one variable where the requirement is that its first order derivative is zero).

Thus the extrema of f are stationary points. However, not all stationary points are extrema (i.e., maximum or minimum points). Any stationary point that is not an extremum is called a **saddle point** (see Fig. 9.3).

To test whether a stationary point is a maximum or a minimum, we need the idea of the **discriminant** of a function. The discriminant D of a function of two variables x, y, $f(x, y)$ is the function:

$$D = f_{xx} f_{yy} - (f_{xy})^2.$$

For example, consider the function $f(x, y) = x^3 + 5xy + y^3 + 4$. Then

$$f_x = 3x^2 + 5y, \quad f_y = 5x + 3y^2, \quad f_{xx} = 6x, \quad f_{yy} = 6y \text{ and } f_{xy} = 5.$$

Therefore, the discriminant of f is the function

$$(6x)(6y) - 5^2 = 36xy - 25.$$

To find the stationary points of a function $f(x, y)$ and to determine their nature (whether a maximum, minimum or a saddle point), follow these steps:

1. Find the stationary points of f. That is, find the pairs of values x_0, y_0 for x, y for which simultaneously $f_x(x_0, y_0) = 0$ and $f_y(x_0, y_0) = 0$.

   The remaining steps are to determine the nature of each stationary point.

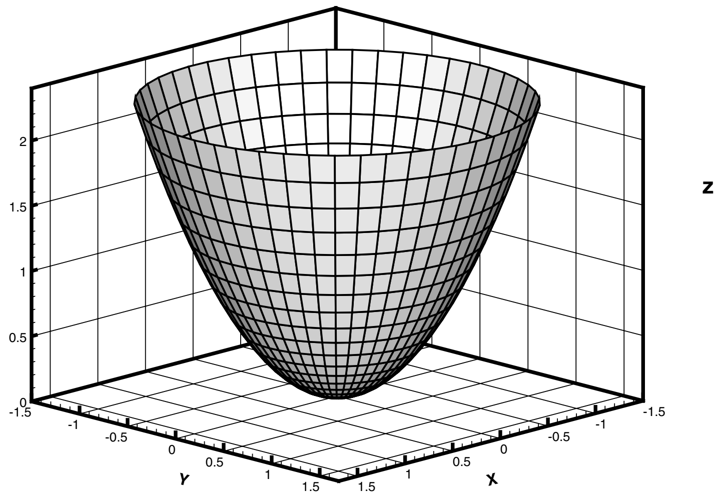

Figure 9.2 The graph of the function $z = x^2 + y^2$. The function has a local minimum at the origin $x = y = z = 0$.

2. Determine all second order partial derivatives f_{xx}, f_{yy}, f_{xy} of f.

3. Now consider a particular stationary point: say $x = x_0, y = y_0$.

   Evaluate the second order partial derivatives for $x = x_0, y = y_0$ and then evaluate the discriminant $D = f_{xx}f_{yy} - (f_{xy})^2$.

4. a) If $D < 0$, the stationary point is a saddle point.

   b) If $D = 0$, the test is inconclusive.

   c) If $D > 0$, the stationary point is an extremum of f. It is a

   $$\textit{maximum} \text{ point if } f_{xx} < 0 \text{ and } f_{yy} < 0$$

   and a

   $$\textit{minimum} \text{ point if } f_{xx} > 0 \text{ and } f_{yy} > 0.$$

(If the discriminant D is zero, the nature of the stationary point can be determined by more advanced techniques, but this will not concern us here.)

If f_{xx} and f_{yy} have different signs (one positive, the other negative), then it is easy to see that $D < 0$ and the stationary point must be a saddle point. It follows that when $D > 0$, the partial derivatives f_{xx}, f_{yy} must have the same sign.

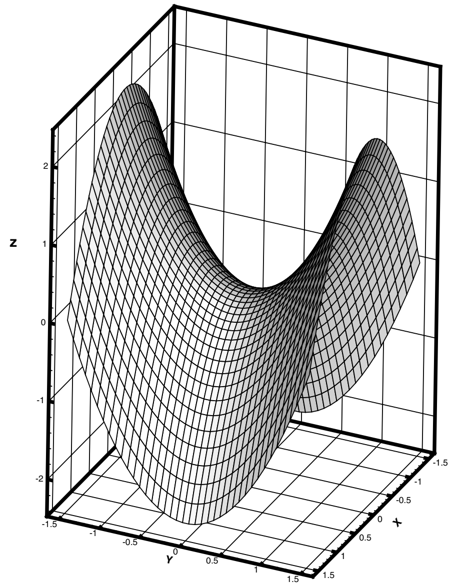

Figure 9.3 The graph of the function $z = y^2 - x^2$. The function has a saddle point at the origin $x = y = z = 0$.

Example 9.2

Determine and classify the stationary point(s) of the function

$$f(x, y) = x^2 + 3y^2 - 2xy + 1.$$

Solution. Here $f_x = 2x - 2y, f_y = 6y - 2x, f_{xx} = 2, f_{yy} = 6$, and $f_{xy} = -2$.

The stationary points occur where $f_x = f_y = 0$. That is, when $2x - 2y = 0$ and $6y - 2x = 0$. The first equation gives $x = y$ while the second gives $3y = x$. Since both equations hold, then $3y = x = y$ and therefore $3y = y$. This means $y = 0$ and so $x = 3y = 0$. Therefore, there is only one stationary point: namely $x = 0, y = 0$.

To test whether this is a maximum or a minimum point for f, note that the discriminant $D = f_{xx}f_{yy} - (f_{xy})^2 = 2 \times 6 - (-2)^2 = 12 - 4 = 8$ is positive. Since f_{xx} and f_{yy} are both positive, the stationary point is a minimum. The value of f is $f(0, 0) = 1$ at this point.

Example 9.3

Determine and classify the stationary point(s) of the function

$$g(x, y) = x^3 + 2y^2 - 3x - 8y.$$

Solution. We have

$$g_x = 3x^2 - 3, \quad g_y = 4y - 8,$$

$$g_{xx} = 6x, \quad g_{yy} = 4, \quad g_{xy} = 0.$$

To find the stationary points, solve the simultaneous equations:

$$g_x = 0, g_y = 0.$$

That is, solve $3x^2 - 3 = 0$ and $4y - 8 = 0$. Equivalently, $x^2 = 1$ and $y = 2$. Therefore, g has two stationary points:

1. $x = 1, y = 2$;
2. $x = -1, y = 2$.

In case (1), $g_{xx} = 6, g_{yy} = 4, g_{xy} = 0$, so that the discriminant is

$$6 \times 4 - 0 = 24 > 0.$$

Since $g_{xx} > 0$ and $g_{yy} > 0$, this stationary point is a minimum. Then $g(1, 2) = -10$ is a minimum value for g.

In case (2), $g_{xx} = -6, g_{yy} = 4$. Since g_{xx}, g_{yy} have different signs, the stationary point is a saddle point.

Example 9.4

X and Y represent the outputs of two goods. The total cost function is

$$TC = 2 + 3X^2 + 2Y^2 - (0.5)XY. \tag{9.1}$$

The market prices for X and Y are 10 and 15 per unit of good, respectively. Determine the outputs that give the maximum profit.

Solution. The total revenue is $TR = 10X + 15Y$ and the profit π is this amount less TC. Therefore

$$\pi = 10X + 15Y - 2 - 3X^2 - 2Y^2 + (0.5)XY, \tag{9.2}$$

which is a function of X and Y.

Now, $\pi_X = 10 - 6X + (0.5)Y$, $\pi_Y = 15 - 4Y + (0.5)X$, $\pi_{XX} = -6$, $\pi_{YY} = -4$, and $\pi_{XY} = 0.5$.

Stationary points are determined by solving $\pi_X = 0, \pi_Y = 0$; that is, the simultaneous equations

$$10 - 6X + (0.5)Y = 0, \tag{9.3}$$

$$15 + (0.5)X - 4Y = 0. \tag{9.4}$$

Multiplying equation (9.3) by 8 gives

$$80 - 48X + 4Y = 0$$

and adding this to equation (9.4) gives

$$95 - (47.5)X = 0.$$

Therefore, $X = 2$, and then substituting into equation (9.4) gives $Y = 4$. So $X = 2, Y = 4$ is the only stationary point of π. Since the discriminant is $(-6) \times (-4) - (0.5)^2 > 0$ and π_{XX}, π_{YY} are both negative, then π has a maximum when $X = 2, Y = 4$. Then the maximum profit 38 is obtained by substituting $X = 2, Y = 4$ into the expression (9.2) for π.

Example 9.5

A company wins a contract to produce rectangular open top boxes using material costing the company £5 a square metre. The contract stipulates that the boxes must all have volume 0.5 cubic metres. What dimensions should each box have so that the cost of the material used is a minimum?

Solution. Suppose the base of the box is an x by y metres rectangle and its height is z. Then the volume of the box is $xyz = 0.5$. The surface area of the base of the box is xy, and the total surface area of the 4 sides is $2xz + 2yz$. So the total area A of material used to make one box is

$$A = xy + 2xz + 2yz.$$

This appears to be a function of three variables but because $xyz = 0.5$, we can express A as a function of only x and y, as $xz = \frac{(0.5)}{y}$ and $yz = \frac{(0.5)}{x}$. Therefore

$$A = xy + 2\frac{(0.5)}{y} + 2\frac{(0.5)}{x}$$

or

$$A = xy + \frac{1}{y} + \frac{1}{x}.$$

Next, find the first and second order partial derivatives of A:

$$\frac{\partial A}{\partial x} = y - x^{-2}, \quad \frac{\partial A}{\partial y} = x - y^{-2},$$

$$\frac{\partial^2 A}{\partial x^2} = 2x^{-3}, \quad \frac{\partial^2 A}{\partial y^2} = 2y^{-3}, \quad \frac{\partial^2 A}{\partial x \partial y} = 1.$$

For a stationary point of A, we require $\frac{\partial A}{\partial x} = \frac{\partial A}{\partial y} = 0$. That is, $y = x^{-2}$ and $x = y^{-2}$. This means $y = x^{-2} = (y^{-2})^{-2} = y^{(-2)\times(-2)} = y^4$. Therefore, either $y = 0$ or $1 = y^3$. Clearly $y \neq 0$ (otherwise we would have a flat box). So $y^3 = 1$, which means $y = 1$ and also $x = y^{-2} = 1$. Since $xyz = 0.5$, then $z = 0.5$ in this case.

Now check whether this gives a minimum for A. At $x = 1, y = 1$,

$$\frac{\partial^2 A}{\partial x^2} = 2 > 0 \text{ and } \frac{\partial^2 A}{\partial y^2} = 2 > 0,$$

while the discriminant of A is given by

$$D = 2 \times 2 - 1^2 = 3 > 0.$$

So we have a minimum. Therefore $x = y = 1, z = 0.5$ are the dimensions of a box that is cheapest to produce with respect to the material used.

Notes

1. The test we described to determine the nature of a stationary point is for functions of two variables. In the case of a function of more than two variables, the situation is more complicated. However in this case, it is still true that any maximum or minimum of the function will occur at a stationary point.

2. It is important to note that we have used the terms maximum or minimum for a function rather loosely. Strictly, we should say a *local* (or *relative*) maximum or minimum: for they may not give the overall maximum or minimum of a function. Within the locality of a maximum point $x = x_0$, $y = y_0$ (that is, for any x and y sufficiently close to these values), the value $f(x, y)$ of f attains a maximum when $x = x_0$ and $y = y_0$. Similarly for a minimum point.

3. The problem stated in Example 9.5 above was initially a constrained optimization problem in three variables. But, upon substitution for z, it became an unconstrained optimization problem in the remaining two variables, x and y.

9.3 Constrained Optimization

Optimization of a quantity in economic models, or indeed in many practical situations, is rarely unconstrained. Usually there are constraints involving some or all of the variables. For instance, in considering ways to maximize, say, output, there will be constraints due to costs or of the available labour.

In this section, we shall look at two methods for optimizing subject to constraints. We'll restrict the discussion to functions of two variables, though in both cases there is a generalization to functions of more than two variables.

The general problem is this. We have a function $f(x, y)$ and want to find its maximum or minimum values subject to a constraint. That is, we want to **optimize** $f(x, y)$ subject to a constraint expressed in the form of an equation $g(x, y) = k$, where k is a constant and g is a function of x and y. We call f the **objective function**, g the **constraint function**, k the **constraint constant**, and $g(x, y) = k$ the **constraint equation** (or simply the **constraint**).

There are various methods used for constrained optimization. We will consider two important techniques: the substitution method and the Lagrange Multiplier method.

9.3.1 Substitution Method

If the constraint equation allows one of the variables, say x, to be expressed explicitly as a function of the other variables, then substitute for x in the objective function. The optimization with constraint problem now reduces to unconstrained optimization of a function of the other variables. Consider the following simple illustrative examples.

Example 9.6

A developer wants to protect as much of his land as possible and has only one kilometre of fencing available. What is the largest rectangular area that can be enclosed?

Solution. Here the objective function is

$$A = xy,$$

where x and y are the length and breadth of the paddock in kilometres. The constraint equation is

$$2x + 2y = 1,$$

since the perimeter of the fence is 1 km. Therefore $x+y = 0.5$ and so $x = 0.5-y$, giving x explicitly as a function of y. Substitute for x in A to obtain

$$A = (0.5 - y)y = 0.5y - y^2.$$

So A is now a function of one variable, namely y. Then

$$\frac{dA}{dy} = 0.5 - 2y \text{ and } \frac{d^2A}{dy^2} = -2.$$

Since $dA/dy = 0$ when $y = 0.25$ and since $d^2A/dy^2 < 0$, then $y = 0.25$ gives a maximum. In this case, $x = 0.5 - y = 0.5 - 0.25 = 0.25$, and $A = (0.25)^2 = 0.0625$ square kilometres (= 62,500 square metres) is the maximum rectangular area that can be enclosed.

Example 9.7

A firm's production function is $Q = 8K^{\frac{1}{4}}L^{\frac{1}{2}}$, where K and L are respectively, capital and labour costs. Unit capital and labour costs are 2 and 1, respectively. What is the minimum total of input costs (that is, costs due to capital and labour) if output Q is to be 240 units?

Solution. Denote the total input costs TC by C (fixed costs are not important here because they are fixed). Then $C = 2 \times K + 1 \times L = 2K + L$. This is, the objective function. So C is a function of K and L. The constraint equation is $Q = 240$. That is $8K^{\frac{1}{4}}L^{\frac{1}{2}} = 240$ or $K^{\frac{1}{4}}L^{\frac{1}{2}} = 30$. Therefore

$$L^{\frac{1}{2}} = \frac{30}{K^{\frac{1}{4}}} = 30K^{-\frac{1}{4}}.$$

Then

$$L = (L^{\frac{1}{2}})^2 = 30^2(K^{-\frac{1}{4}})^2 = 900K^{-\frac{1}{2}}$$

(using the rule (1.13)). Therefore

$$C = 2K + 900K^{-\frac{1}{2}}.$$

This expresses C as a function of the single variable K. Its derivatives are

$$\frac{dC}{dK} = 2 + 900(-\frac{1}{2})K^{-\frac{1}{2}-1} = 2 - 450K^{-\frac{3}{2}},$$
$$\frac{d^2C}{dK^2} = -450(-\frac{3}{2})K^{-\frac{3}{2}-1} = 675K^{-\frac{5}{2}}.$$

We have $dC/dK = 0$ when $2 = 450K^{-\frac{3}{2}} = 450/K^{\frac{3}{2}}$. Therefore, $K^{\frac{3}{2}} = 450/2 = 225$ and so $(K^{\frac{3}{2}})^{\frac{2}{3}} = 225^{\frac{2}{3}}$. That is, $K^1 = K = 225^{\frac{2}{3}}$, which is approximately 37.

Since d^2C/dK^2 is positive in this case, then $K = 37$ gives a minimum for C. When $K = 37$, $L = 900K^{-\frac{1}{2}}$, which is approximately 148, and $C = 2K + L$ is approximately 222.

In the previous example, we optimized total input costs while constraining output. The next problem optimizes output while restraining input costs.

Example 9.8

A firm's unit capital and labour costs are respectively 2 and 4. The production function is $Q = 6KL + 2L^2$.

1. If the total input costs are 200 units, what is the maximum possible output Q?
2. If the output is fixed at 1,200, what are the minimum input costs?

Solution.

1. The objective function is

   $$Q = 6KL + 2L^2$$

   and the constraint is

   $$2K + 4L = 200$$

   which simplifies to

   $$K + 2L = 100.$$

   Therefore, $K = 100 - 2L$ and we can express Q as a function of the single variable L by substituting for K in the formula for Q. Therefore,

   $$Q = 6(100 - 2L)L + 2L^2 = 600L - 12L^2 + 2L^2 = 600L - 10L^2.$$

   Then $dQ/dL = 600 - 20L$ and $d^2Q/dL^2 = -20$. Since $dQ/dL = 0$ when $L = 30$ and since the second derivative is negative, then $L = 30$ gives a maximum for Q. In this case, $K = 100 - 2L = 100 - 60 = 40$. Then $Q = 9{,}000$ is the maximum output when total input costs are 200 units.

2. The objective function is now the total input costs

$$C = 2K + 4L,$$

while the constraint is

$$Q = 6KL + 2L^2 = 1{,}200,$$

which implies $6KL = 1{,}200 - 2L^2$ and so

$$K = \frac{1{,}200 - 2L^2}{6L} = \frac{1200}{6L} - \frac{2L^2}{6L}.$$

Therefore,

$$K = 200L^{-1} - \frac{L}{3}$$

and so we can express $C = 2K + 4L$ as a function of the single variable L thus:

$$C = 400L^{-1} - \frac{2L}{3} + 4L = 400L^{-1} + \frac{10L}{3}.$$

Then

$$\frac{dC}{dL} = -400L^{-2} + \frac{10}{3}$$

and

$$\frac{d^2C}{dL^2} = -400 \times (-2L^{-3}) = 800L^{-3}.$$

The stationary points of C (as a function of L) are given by solving

$$\frac{dC}{dL} = 0.$$

Equivalently,

$$\frac{10}{3} = 400L^{-2} = \frac{400}{L^2}.$$

Therefore

$$L^2 = \frac{3 \times 400}{10} = 120$$

which means $L = \pm\sqrt{120}$. Ignore the negative labour costs (this would mean the labour force pays to work!) then $L = \sqrt{120} = 10.95$ (to 2 decimal places) gives a minimum for C since the second derivative is positive. In this case, $K = 200L^{-1} - \frac{L}{3} = 14.61$ and $C = 2K + 4L = 73.03$ are the minimum total input costs when production is constant at 1,200 units.

9.3.2 Lagrange Multipliers

The method of Lagrange multipliers for constrained optimization can be applied generally; unlike the substitution method. The latter requires that one variable can be expressed explicitly in terms of the others, using the constraint equation.

For simplicity, we consider functions of only two variables, although the Lagrange multiplier method easily extends to more general situations.

As before, we wish to optimize the objective function $f(x, y)$, subject to a constraint equation $k = g(x, y)$. Here k is the constraint constant (constant relative to x and y), known also as the constraint limitation.

The Lagrangian multiplier method introduces a new variable λ and introduces a new function F, known as the **Lagrangian**, defined by:

$$F(x, y, \lambda) = f(x, y) + \lambda(k - g(x, y)).$$

The parameter λ is known as the Lagrange multiplier.

It may appear that by turning a two variable problem into a three variable one makes the problem harder. However, the method transforms a problem of constrained optimization to one of unconstrained optimization. We state this as follows:

The pairs of values of x, y that optimize the function $f(x, y)$, subject to the constraint $k = g(x, y)$, occur where the Lagrangian function

$$F = f(x, y) + \lambda(k - g(x, y))$$

in the three variables x, y, λ has its optimum values.

In other words, suppose f has an optimum value when, say, $x = x_0$, $y = y_0$. Then the function F has a stationary point for $x = x_0$, $y = y_0$ and some value λ_0 of λ. So, by finding the stationary points of F, we obtain all the possible pairs x, y that optimize f subject to the given constraint. However, the methods covered in this book do not allow us to determine which of these pairs give maxima or minima. Unlike the substitution method, for the Lagrange multiplier method we rely on intuition or the particular nature of the problem to say whether a stationary point gives a maximum or minimum.

Note. The Lagrange multiplier λ multiplies the expression $k - g(x, y)$. This expression is clearly 0 when the constraint equation holds, and then $F = f$. This gives an intuitive idea why it is that, when the constraint holds, the optimum values (maxima or minima) of f are those of F.

To sum up, here is how to apply the Lagrange multiplier method to optimize a function $f(x, y)$ subject to a constraint $k = g(x, y)$.

1. Form the Lagrangian $F(x, y, \lambda) = f(x, y) + \lambda(k - g(x, y))$.
2. Find the three first order partial derivatives $\frac{\partial F}{\partial x}$, $\frac{\partial F}{\partial y}$ and $\frac{\partial F}{\partial \lambda}$ of F.
3. Solve the simultaneous equations

$$\frac{\partial F}{\partial x} = 0, \qquad \frac{\partial F}{\partial y} = 0, \qquad \frac{\partial F}{\partial \lambda} = 0.$$

   (The last equation is just $k - g(x, y) = 0$ which is the constraint.) This determines the stationary points of F and hence the possible maxima or minima of f subject to the constraint $k = g(x, y)$.

Example 9.9

A firm produces two goods G_1 and G_2. The output of G_1 is denoted by by Q_1 and its price by P_1. Similarly for good G_2. The production functions are

$$P_1 = 20 - Q_1 + 2Q_2$$

and

$$P_2 = 10 + Q_1 - Q_2.$$

The total costs are given as

$$TC = 12Q_1 + Q_1Q_2 + 6Q_2. \tag{9.5}$$

The firm is contracted to produce a total of 20 units of goods of either type. What is the maximum profit possible?

Solution. Essentially, we are to find the values of Q_1 and Q_2 that maximize profit subject to the constraint $Q_1 + Q_2 = 20$. First we need to compute the profit $\pi = TR - TC$. This is the objective function.

Since total revenue $TR = P_1Q_1 + P_2Q_2$, then

$$TR = (20 - Q_1 + 2Q_2)Q_1 + (10 + Q_1 - Q_2)Q_2.$$

After simplifying, we get that

$$TR = 20Q_1 + 10Q_2 + 3Q_1Q_2 - Q_1^2 - Q_2^2. \tag{9.6}$$

The profit is given by $\pi = TR - TC$. Using equations (9.5) and (9.6) and simplifying gives

$$\pi = 8Q_1 + 4Q_2 + 2Q_1Q_2 - Q_1^2 - Q_2^2. \tag{9.7}$$

This is our objective function. The constraint equation is

$$20 - Q_1 - Q_2 = 0.$$

Therefore the Lagrangian is

$$F = 8Q_1 + 4Q_2 + 2Q_1Q_2 - Q_1^2 - Q_2^2 + \lambda(20 - Q_1 - Q_2).$$

For stationary points we need:

$$0 = F_{Q_1} = 8 + 2Q_2 - 2Q_1 - \lambda \tag{9.8}$$
$$0 = F_{Q_2} = 4 + 2Q_1 - 2Q_2 - \lambda \tag{9.9}$$
$$0 = F_\lambda = 20 - Q_1 - Q_2. \tag{9.10}$$

From (9.8) and (9.9) we have

$$8 + 2Q_2 - 2Q_1 = 4 + 2Q_1 - 2Q_2 = \lambda.$$

Therefore $4Q_1 - 4Q_2 = 4$, which simplifies to

$$Q_1 - Q_2 = 1. \tag{9.11}$$

From the constraint equation (9.10), we have

$$Q_1 + Q_2 = 20$$

which when added to (9.11) gives

$$2Q_1 = 21.$$

Therefore $Q_1 = 10.5$, and then from (9.11) we have $Q_2 = Q_1 - 1 = 9.5$. With these values of Q_1, Q_2 we compute from (9.7) the maximum profit to be 121.

Example 9.10

Solve Example 9.8 using Lagrange multipliers.

Solution.

1. Here, we are to find the values of K and L that maximize production subject to the constraint $2K + 4L = 200$.

Objective function:	Q
Constraint equation:	$200 = 2K + 4L$
Lagrangian:	$F = Q + \lambda(200 - 2K - 4L)$

Therefore $F = 6KL + 2L^2 + 200\lambda - 2\lambda K - 4\lambda L$ and so

$$F_K = 6L - 2\lambda,$$
$$F_L = 6K + 4L - 4\lambda,$$
$$F_\lambda = 200 - 2K - 4L.$$

The stationary points of F occur where all of these three first order partial derivatives are zero. That is

$$6L - 2\lambda = 0 \quad \text{or} \quad \lambda = 3L; \tag{9.12}$$
$$6K + 4L - 4\lambda = 0 \quad \text{or} \quad 2\lambda = 3K + 2L; \tag{9.13}$$
$$200 - 2K - 4L = 0 \quad \text{or} \quad 2K + 4L = 200. \tag{9.14}$$

Notice that equation (9.14) is just the constraint equation.

Eliminating λ between (9.12) and (9.13) gives $3K + 2L = 2 \times 3L = 6L$, which implies $3K = 4L$. Putting this in (9.14) gives $200 = 2K + 3K = 5K$. Therefore $K = 40$. Since $3K = 4L$, then $L = \frac{3}{4}K = 30$. We also have the corresponding Lagrange multiplier $\lambda = 3L = 90$, using (9.12). We substitute $K = 40$, $L = 30$ into the production function Q in order to obtain the optimal value:

$$Q(40, 30) = 6 \times 40 \times 30 + 2 \times 30^2 = 9{,}000.$$

This is the maximum production Q subject to the constraint $2K + 4L = 200$. (It's obviously not the minimum production since $K = 100, L = 0$ satisfy the constraint and give zero production.)

2. Here we are to minimize input costs subject to the constraint

$$6KL + 2L^2 = 1{,}200.$$

Objective function:	$C = 2K + 4L$
Constraint equation:	$1{,}200 = 6KL + 2L^2$
Lagrangian:	$F = 2K + 4L + \lambda(1{,}200 - 6KL - 2L^2)$

To find the stationary points of F, we equate all its first order partial derivatives to zero.

$$0 = F_K = 2 - 6\lambda L, \tag{9.15}$$
$$0 = F_L = 4 - 6K\lambda - 4L\lambda, \tag{9.16}$$
$$0 = F_\lambda = 1{,}200 - 6KL - 2L^2. \tag{9.17}$$

First we eliminate λ between equations (9.15) and (9.16). From (9.15) and (9.16) we have $\lambda = \frac{1}{3L}$ and $\lambda = \frac{2}{3K+2L}$, respectively. Equating these values of λ and simplifying gives $3K = 4L$, whence $K = \frac{4}{3}L$. Substituting for K in (9.17), the constraint, gives $1{,}200 = 6KL + 2L^2 = 6(\frac{4}{3}L)L + 2L^2 = 10L^2$. Therefore $L^2 = 120$. Then $L = \sqrt{120}$ and so $L = 10.95$, correct to 2 decimal places. (We've ignored the negative solution $L = -\sqrt{120}$. Then, since $K = \frac{4}{3}L$, we have $K = 14.61$ and $C = 2K + 4L = 73.03$. We also have from (9.15) that $\lambda = \frac{1}{3L} = 0.03$. This information about λ is not needed to optimise C but is nevertheless useful to know, as we shall see.

9.3.3 The Lagrange Multiplier λ: An Interpretation

The Lagrange multiplier λ used in constrained optimization appears at first glance to have no use as it is eliminated from the equations determining a stationary point and does not appear in the optimum value of the objective function or in the values of the variables giving the optimum value. However, λ does have an important and useful interpretation.

Consider the problem of optimizing a function $f(x, y)$ subject to a constraint $k = g(x, y)$, where k is the constraint constant. (Here 'constant' means that k is independent of the value of x or y.) The Lagrangian function

$$F = f(x, y) + \lambda(k - g(x, y))$$

is a function of x, y and λ.

For given f and g, the value M of any optimum (maximum or minimum) of f depends only on the constraint constant k. So M can be considered as a function of k. It can be shown that $\frac{dM}{dk} = \lambda$; that is, λ is the rate of change of M relative to k. This means that:

> The Lagrange multiplier λ is approximately the change in an optimum value of the objective function resulting from a one unit increase in the constraint constant.

In Example 9.10, if the total input costs were fixed at 201, the maximum production would increase approximately by $\lambda = 90$ to 9,090. If we were to carry out the computation with the new constraint constant of 201, we would find the actual new maximum production to be 9,090.225, which is close to the approximation.

If the input costs were fixed at 199, the maximum production would *decrease* by approximately $\lambda = 90$ to 8,910.

For non-unit changes in the constraint constant, the computations are pro rata. For instance, if the input costs are fixed at 205, the maximum production would increase by about $5 \times 90 = 450$ to 9,450. (The actual figure is 9,455.625.)

Example 9.11

A company allocates £600,000 to spend on advertising and research. The company estimates that by spending x thousand pounds on advertising and y thousand pounds on research, they will sell a total of approximately $30x^{4/5}y^{1/3}$ units of its product. How much should the company spend on research and advertising in order to maximize sales?

Solution. We work in units of £1,000. Then the objective function is

$$f(x, y) = 30x^{4/5}y^{1/3}$$

and the constraint equation is

$$x + y = 600 \text{ or } 600 - x - y = 0.$$

The problem can be solved by the substitution method, but the Lagrange multiplier method is used here: The Lagrangian is

$$F = 30x^{4/5}y^{1/3} + \lambda(600 - x - y).$$

For stationary points:

$$0 = F_x = 24x^{-1/5}y^{1/3} - \lambda, \tag{9.18}$$

$$0 = F_y = 10x^{4/5}y^{-2/3} - \lambda, \tag{9.19}$$

$$0 = F_\lambda = 600 - x - y. \tag{9.20}$$

Equations (9.18) and (9.19) give

$$\lambda = 24x^{-1/5}y^{1/3} = 10x^{4/5}y^{-2/3}. \tag{9.21}$$

Therefore $x = 2.4y$. Substituting in (9.20) gives $600 - 2.4y - y = 0$ or $y = \frac{600}{3.4} = 176.47$ (to 2 decimal places). Then $x = 423.53$ and using (9.21) gives $\lambda = 40.15$. The maximum sales total is therefore $f(423.53, 176.47) = 21{,}257.83$.

If in this example we change the advertising and research budget, we can use the Lagrange multiplier λ to estimate the resulting maximum sales total. For instance, suppose the budget is increased by 1% to £606,000: an increase of 6 in the constraint constant, since we are working in units of £1,000. Therefore the maximum sales total will increase by about $6\lambda = 6 \times 40.15 = 240.90$. This is an increase of about 1.13% on the previous sales maximum. (If you were to work through the problem again, with the new budget of £606,000, the actual increase would be 241.08.)

If now the budget decreases by 1.5% to £591,000, the maximum sales total would decrease by approximately $9\lambda = 9 \times 40.15 = 361.35$ (because the constraint constant decrease by 9). This is about a 1.7% decrease in the maximum sales. (The actual decrease in maximum sales is 361.02 to 2 decimal places.)

Example 9.12

A consumer's utility function $U(x, y)$ is given by $U = 30x^{2/5}y^{1/3}$, where x is the number of units of good G_X and y the number of units of good G_Y. Each unit of G_X costs €1 and each of G_Y costs €2. If the consumer's total income $Y =$ €1,100, find the maximum utility U_{max} for the consumer.

Solution. The objective function is U and the constraint is $x + 2y = 1100$. The Lagrangian is therefore:

$$F = 30x^{2/5}y^{1/3} + \lambda(1{,}100 - x - 2y).$$

Its stationary points occur where

$$0 = F_x = 30 \times \frac{2}{5}x^{2/5-1}y^{1/3} - \lambda; \qquad \text{i.e., } \lambda = 12x^{-3/5}y^{1/3} \qquad (9.22)$$

$$0 = F_y = 30x^{2/5} \times \frac{1}{3}y^{1/3-1} - 2\lambda; \qquad \text{i.e., } \lambda = 5x^{2/5}y^{-2/3} \qquad (9.23)$$

$$\text{and } 0 = F_\lambda = 1{,}100 - x - 2y. \qquad \text{(Constraint Equation)} \qquad (9.24)$$

From equations (9.22) and (9.23), we eliminate λ to obtain:

$$5x^{2/5}y^{-2/3} = 12x^{-3/5}y^{1/3},$$

which simplifies to $x = 2.4y$. Substituting for x in the constraint equation (9.24) gives $(2.4 + 2)y = 1{,}100$ and therefore

$$y = \frac{1{,}100}{4.4} = 250.$$

Then $x = 2.4y = 600$. Thus $U_{\max} = U(600, 250) = 2{,}441.72$ (correct to 2 decimal places). From (9.22) we obtain $\lambda = 1.63$ in this case.

If, for instance, the total income Y increases to €1,101, the maximum utility increases by approximately $1 \times \lambda = 1.63$. If Y decreases to €1,050, the maximum utility decreases, approximately, by $50 \times 1.63 = 81.50$. (The actual figures, correct to 2 decimal places, are 1.63 and 81.89, respectively.)

9.4 Iso Curves

If we have a function $f(x, y)$ of two variables x and y, then (see Section 8.7.5) we can visualise the function as a family of nested curves, each with equation of the form:

$$f(x, y) = c,$$

where c is a constant. The curves are known generally as **iso curves** or iso lines. (The word iso comes from the Greek for equal.)

The combinations of values for x and y that give the same value c for the function f are all the coordinate pairs x, y of the points on the iso curve $f(x, y) = c$. Each point on this iso curve corresponds to such a combination.

Iso curves may have specific names depending on what the function f represents. If f is a utility function, then, as noted in Section 8.7.5, the iso curves are called indifference curves.

Other examples are when f is a production, profit, or cost function. Then the iso curves are known respectively, as **isoquants**, **isoprofit**, or **isocost curves**.

Using iso curves we can visualize constrained optimization. We illustrate this using the utility function $U = 30x^{2/5}y^{1/3}$ of Example 9.12. The constraint is the budget of €1,100 which requires $x + 2y = 1{,}100$. This equation represents the budget line. Its graph is given in Fig. 9.4.

For any given positive number c, the points (x, y) of the indifference curve $U = 30x^{2/5}y^{1/3} = c$ provide all combinations x, y that give the same utility c (see Fig. 9.5). The bigger c is, the further the curve is away from the origin. For example, the indifference curve $U = 2$ lies between the indifference curves $U = 1$ and $U = 3$ (see Fig. 9.6).

Every point (x, y), with $x, y > 0$, is on exactly one indifference curve. If c is large enough, an indifference curve $U = c$ will not meet the constraint/budget line $x+2y = 1{,}100$. As c decreases, the indifference curve approaches the budget line and for one value $c = c_0$ the indifference curve $U = c_0$ will touch the budget line at one point P (see Fig. 9.7).

The coordinates of P satisfy the constraint $x + 2y = 1{,}100$ since P is on the budget line. The indifference curve $U = c_0$ on P has maximum utility. We have already computed $c_0 = 2{,}441.72$ and found the coordinates of P to be $x = 600$, $y = 250$.

Any indifference curve $U = a$ with $a > c_0$ misses the budget line; so no point on the curve has coordinates satisfying the budget constraint. An indifference curve $U = b$, with $b < c_0$, meets the budget line in two points whose coordinates satisfy the budget constraint (since the points are on the budget line) but the corresponding utility b is less than c_0 (see Fig. 9.8).

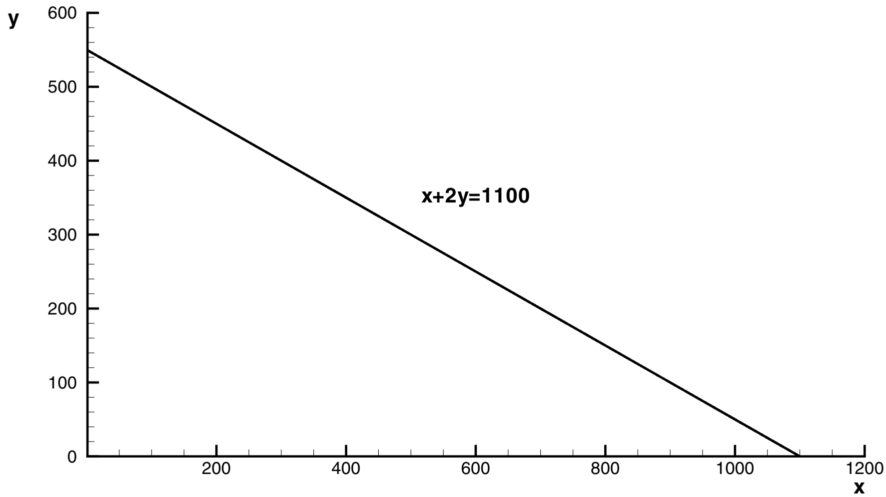

Figure 9.4 The graph of the budget line $x + 2y = 1100$.

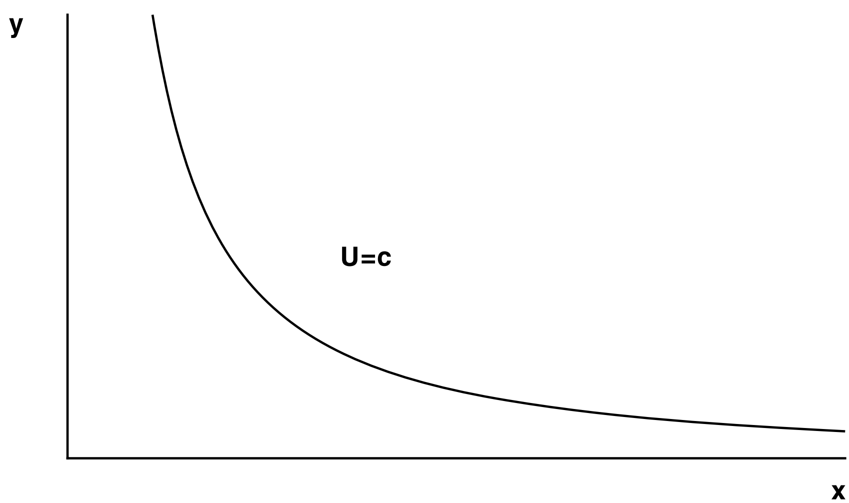

Figure 9.5 An iso curve (or indifference curve) of the utility function $U = 30x^{2/5}y^{1/3}$ for $U = c$ where c is some constant.

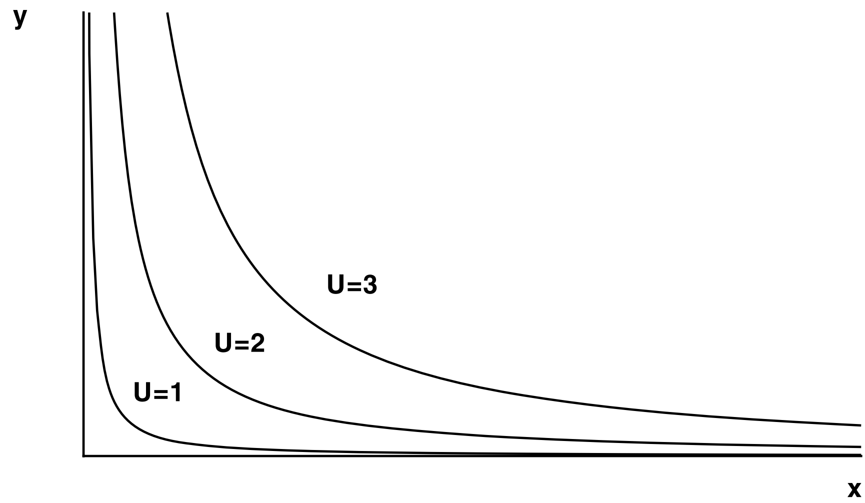

Figure 9.6 Iso curves of the utility function $U = 30x^{2/5}y^{1/3}$ corresponding to $U = 1$, $U = 2$ and $U = 3$.

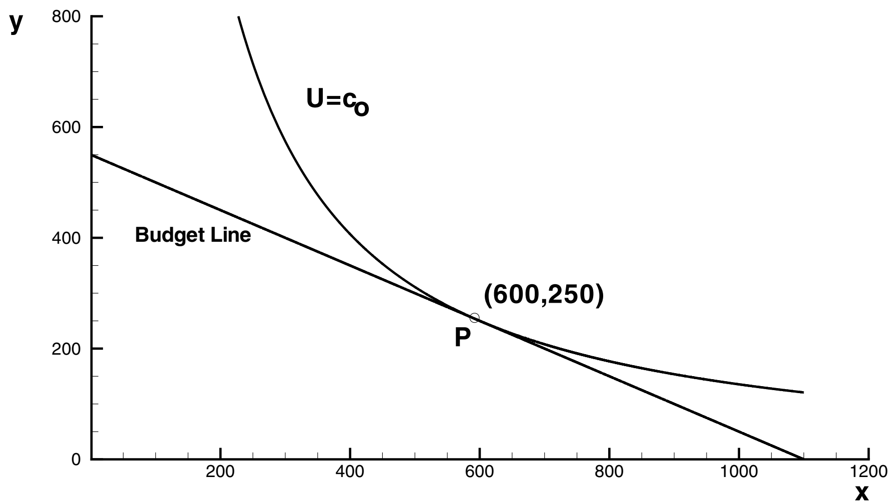

Figure 9.7 The graphs of the budget line $x + 2y = 1{,}100$ and the indifference curve $U = c_0$.

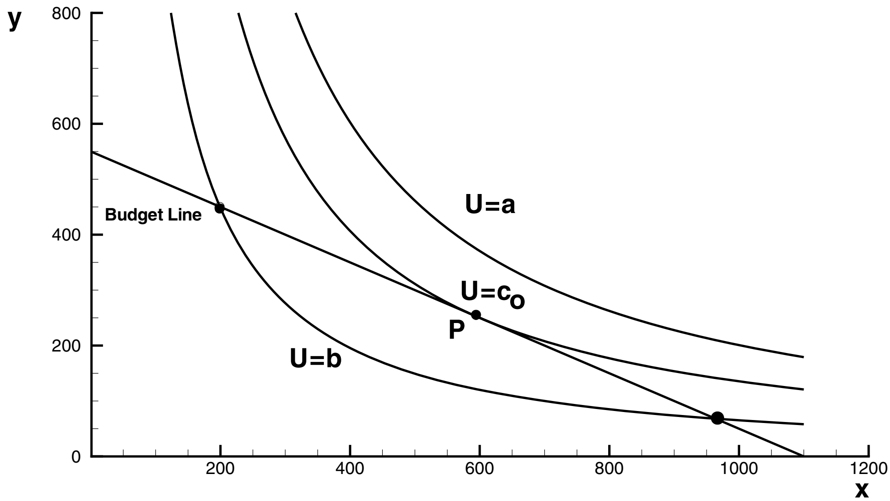

Figure 9.8 The graphs of the budget line $x + 2y = 1{,}100$ and the indifference curves $U = c_0$, $U = a$ and $U = b$ for $b < c_0 < a$.

EXERCISES

9.1. Maximize $4 + xy$ subject to the constraint $2x + y = 1$ using the substitution method.

9.2. A firm's total costs are given by

$$TC = 10L^2 + 10K^2 - 25L - 50K - 5KL + 1{,}500$$

where L is the workforce size (in thousands) and K the capital invested (in thousands of dollars).

Use the substitution method to find the combination of labour and capital that will minimize TC.

9.3. A firm's production function is given by $Q = 12K^{1/2}L^{1/4}$. Unit capital and labour costs are respectively 6 and 4. If the firm must provide 120 units of output, find the minimum total cost of production using the substitution method.

9.4. Optimize xy subject to the constraint $x^2+y^2 = 2$ using the Lagrange multiplier method.

9.5. Use the Lagrange multiplier method to maximize output

$$Q = 4KL + L^2$$

subject to the constraint $TC = K + 2L = 175$.

a) *Estimate* the maximum output if TC is fixed at 174.5.

b) *Estimate* at what value TC should be fixed if the maximum output is to be 17,600.

9.6. A digital media company has a budget of €120,000 to spend on recording and promoting a new DVD.

The company estimates that if it spends x thousand euros on recording and y thousand euros on promotion, it will sell approximately $6yx^{2/3}$ units of DVDs.

a) Use the Lagrange multiplier method to show how the company should allocate its budget to maximize sales?

Evaluate the maximum sales.

b) *Estimate* the maximum sales if the budget is

i. increased by €1,000;

ii. decreased by €500.

10
Matrices and Determinants

10.1 Introduction

Matrix theory is a powerful mathematical tool for dealing with data as a whole rather than the individual items of data. Matrices are especially useful in the theory of equations. They can be used to solve systems of simultaneous linear equations. Determinants are related to matrices and are useful for determining whether or not a unique solution exists. In some cases using determinants, the solution for each unknown can be expressed explicitly in terms of the coefficients of the equations by applying what is known as Cramer's rule. Systems of simultaneous linear equations occur, for example, when optimizing a function using Lagrange multipliers or when trying to find the equilibrium prices of interdependent commodities. As we shall see, matrices can be added and in some cases multiplied together. In economics, business, and finance, many basic theoretical models are linear in that they are described in some way by linear functions. Analyzing these models is made simpler by matrix algebra.

10.2 Matrix Operations

A rectangular array of mn numbers in m rows and n columns is called a **matrix** of size $m \times n$ ('m by n'). The array is enclosed in square or, sometimes, curved brackets. The (i, j)-entry of a matrix M is the entry in the i^{th} row and j^{th}

column of M. This entry can be denoted simply by M_{ij}. For example:

$$M = \begin{bmatrix} 3 & -1 & 0 \\ 0 & 2 & -5 \end{bmatrix} \text{ is a } 2 \times 3 \text{ matrix}$$

and

$$N = \begin{bmatrix} 5 & 1 \\ 2 & -1 \end{bmatrix} \text{ is a } 2 \times 2 \text{ matrix.}$$

For the matrix M, we have $M_{11} = 3$, $M_{13} = 0$, $M_{23} = -5$, and so on.

An $n \times n$ matrix, that is one that has as many rows as columns, is called a **square matrix**. The matrix N is square. A $1 \times n$ matrix is called a **row matrix** or **row vector** of length n. An $n \times 1$ matrix is called a **column matrix** or **column vector** of height n. For example: $\begin{bmatrix} 4 \\ 1 \end{bmatrix}$ is a column matrix of height 2 and $\begin{bmatrix} 0 & 1 & -3 \end{bmatrix}$ is a row matrix of length 3.

The **transpose** of an $m \times n$ matrix M is the $n \times m$ matrix whose i^{th} row is the i^{th} column of M $(i = 1, 2, \ldots n)$. The matrix is denoted by M^t and called 'M transpose'. Another way to define M^t is as the $n \times m$ matrix whose (i, j)-entry is M_{ji}. It is immediately clear that the transpose of M^t is M. That is, $(M^t)^t = M$. For example,

$$\begin{bmatrix} 3 & 5 & -1 \end{bmatrix}^t = \begin{bmatrix} 3 \\ 5 \\ -1 \end{bmatrix}, \quad \begin{bmatrix} 4 & 3 \\ 0 & -2 \end{bmatrix}^t = \begin{bmatrix} 4 & 0 \\ 3 & -2 \end{bmatrix},$$

$$\begin{bmatrix} 1 & 4 & -1 \\ 2 & 0 & -3 \end{bmatrix}^t = \begin{bmatrix} 1 & 2 \\ 4 & 0 \\ -1 & -3 \end{bmatrix}.$$

A matrix M that is its own transpose, so that $M^t = M$, is said to be **symmetric**. Obviously, only square matrices can be symmetric.

$$\begin{bmatrix} 3 & 1 & 0 \\ 1 & -2 & 6 \\ 0 & 6 & 2 \end{bmatrix} \text{ is a symmetric matrix.}$$

We shall introduce some operations that can be performed on matrices. The three basic ones are: scalar multiplication, matrix addition, and matrix multiplication.

10.2.1 Scalar Multiplication

The term 'scalar' in this context simply means a number as opposed to a matrix. The reasons for use of this term are historical. It is still a useful term if we wish to distinguish scalar from matrix multiplication of matrices.

If λ is any number and M any matrix, the **scalar multiple** of M by λ is the matrix, denoted by λM, obtained by multiplying each entry of M by λ, so the (i, j)-entry of λM is λM_{ij}. Obviously, λM and M have the same size. The following are some examples:

$$\begin{bmatrix} 36 & -16 \\ 0 & 24 \end{bmatrix} = 4 \begin{bmatrix} 9 & -4 \\ 0 & 6 \end{bmatrix},$$

$$-2 \begin{bmatrix} 4 & 1 & -2 \\ -3 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -8 & -2 & 4 \\ 6 & 0 & -2 \end{bmatrix}.$$

We write $-M$ rather than $(-1)\,M$. So, for instance

$$-\begin{bmatrix} 1 & -3 \\ -4 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 3 \\ 4 & 0 \end{bmatrix}.$$

Suppose that two firms A and B each produce two goods G_1, G_2. A consumer is supplied with both goods by both firms. The quantities supplied over a particular period can be represented by a matrix

$$Q = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix},$$

which we shall call the supply matrix, where the first row gives the quantity supplied of goods G_1 and G_2, respectively, by firm A and the second row that by firm B. If the consumer increases all the quantities bought from each firm by 20%, the new supply matrix would be $(1.2)Q$.

In theoretical discussions, it is sometimes useful to distinguish scalars from matrices by denoting scalars by lowercase Greek letters and matrices by uppercase latin letters. For instance, the following easy to see matrix rule

$$\lambda\,(\mu M) = (\lambda\mu)\,M$$

says that multiplying a matrix M by a scalar μ and then by a scalar λ is the same as multiplying M by $\lambda\mu$. For example, $3\,(5M) = 15M$.

10.2.2 Matrix Addition

Matrices of the same size can be added. If M and N are matrices of the same size, then their sum $M+N$ is the matrix whose (i,j)-entry is $M_{ij}+N_{ij}$. That is $M+N$ is obtained by adding corresponding entries of M and N. Clearly $M+N$ will have the same size as M and N. For example,

$$\begin{bmatrix} 5 & -4 \\ 3 & 2 \end{bmatrix} + \begin{bmatrix} -2 & 4 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 3 & 0 \\ 4 & 1 \end{bmatrix},$$

$$\begin{bmatrix} 5 \\ -2 \\ 1 \end{bmatrix} + \begin{bmatrix} 3 \\ -4 \\ -2 \end{bmatrix} = \begin{bmatrix} 8 \\ -6 \\ -1 \end{bmatrix}.$$

Suppose

$$Q = \begin{bmatrix} 10 & 35 \\ 15 & 18 \end{bmatrix}, \text{ and } Q' = \begin{bmatrix} 25 & 10 \\ 40 & 5 \end{bmatrix}$$

are the supply matrices (see Section 10.2.1) for two successive periods of a year. Then their sum

$$Q + Q' = \begin{bmatrix} 35 & 45 \\ 55 & 23 \end{bmatrix}$$

is the supply matrix for the combined two-year period.

We can also define a matrix $M-N$ in an obvious way by subtracting corresponding entries:

$$\begin{bmatrix} 5 & -4 \\ 3 & 2 \end{bmatrix} - \begin{bmatrix} -2 & 4 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 7 & -8 \\ 2 & 3 \end{bmatrix}.$$

It is clear that $M-N = M+(-N)$.

Matrix addition is commutative. This is the formal way of saying that the order in which addition is performed is unimportant. That is, $M+N = N+M$ for any two matrices M, N of the same size.

The $m \times n$ **zero matrix** is the $m \times n$ matrix with all zero entries. It is denoted simply by 0, the size $m \times n$ being clear usually from the context. Clearly $M+0 = 0+M = M$ for any matrix M. We have also $M-M = 0 = -M+M$.

10.2.3 Matrix Multiplication

If $A = \begin{bmatrix} a_1 & a_2 & \dots & a_m \end{bmatrix}$ is a $1 \times m$ row matrix and $B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$ is an $m \times 1$ column matrix , then we define the product AB to be the number $a_1 b_1 + a_2 b_2 + \ldots + a_m b_m$.

More generally, if M, N are matrices, then we can define their product MN *if the number of columns in M is the number of rows in N*; say M is $m \times p$ and N is $p \times n$. Then MN is the $m \times n$ matrix whose (i, j)-entry is the number $M_i N_j$, where M_i is the i^{th} row of M and N_j the j^{th} column of N. The following example illustrates this operation for some particular matrices.

Example 10.1

1. If $M = \begin{bmatrix} 3 & 4 & -1 \end{bmatrix}$ and $N = \begin{bmatrix} 2 \\ 1 \\ 7 \end{bmatrix}$, then

$$MN = 3 \times 2 + 4 \times 1 + (-1) \times 7 = 6 + 4 - 7 = 3.$$

If $M = \begin{bmatrix} 2 & -3 & 1 \end{bmatrix}$ and $N = \begin{bmatrix} 4 \\ 1 \\ -5 \end{bmatrix}$, then

$$MN = 2 \times 4 + (-3) \times 1 + 1 \times (-5) = 8 - 3 - 5 = 0.$$

2. If $M = \begin{bmatrix} 2 & -1 & 4 \\ 1 & 0 & 2 \\ 2 & 3 & -8 \end{bmatrix}$ and $N = \begin{bmatrix} 5 & 2 \\ 4 & -3 \\ 1 & 0 \end{bmatrix}$, then

$$MN = \begin{bmatrix} 10 & 7 \\ 7 & 2 \\ 14 & -5 \end{bmatrix},$$

since

$$\begin{aligned} M_1N_1 &= 2 \times 5 + (-1) \times 4 + 4 \times 1 = 10, \\ M_1N_2 &= 2 \times 2 + (-1) \times (-3) + 4 \times 0 = 7, \\ M_2N_1 &= 1 \times 5 + 0 \times 4 + 2 \times 1 = 7, \\ M_2N_2 &= 1 \times 2 + 0 \times (-3) + 2 \times 0 = 2, \\ M_3N_1 &= 2 \times 5 + 3 \times 4 + (-8) \times 1 = 14, \\ M_3N_2 &= 2 \times 2 + 3 \times (-3) + (-8) \times 0 = -5. \end{aligned}$$

3. If $A = \begin{bmatrix} 2 & 4 \\ 1 & 0 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix}$, then

$$\begin{aligned} AB &= \begin{bmatrix} 2 & 4 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 2\times 1 + 4 \times 2 & 2 \times (-1) + 4 \times 1 \\ 1 \times 1 + 0 \times 2 & 1 \times (-1) + 0 \times 1 \end{bmatrix} \\ &= \begin{bmatrix} 10 & 2 \\ 1 & -1 \end{bmatrix} \end{aligned}$$

and

$$\begin{aligned} BA &= \begin{bmatrix} 1 & -1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} 2 & 4 \\ 1 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 1 \times 2 + (-1) \times 1 & 1 \times 4 + (-1) \times 0 \\ 2 \times 2 + 1 \times 1 & 2 \times 4 + 1 \times 0 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 4 \\ 5 & 8 \end{bmatrix}. \end{aligned}$$

4. If $A = \begin{bmatrix} 5 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 3 & 1 & 2 \\ -2 & 0 & 4 \end{bmatrix}$, then

$$\begin{aligned} AB &= \begin{bmatrix} 5 \times 3 + 1 \times (-2) & 5 \times 1 + 1 \times 0 & 5 \times 2 + 1 \times 4 \end{bmatrix} \\ &= \begin{bmatrix} 13 & 5 & 14 \end{bmatrix}. \end{aligned}$$

5. If $A = \begin{bmatrix} 3 & 1 & -4 \\ 0 & 2 & 1 \\ 5 & -2 & -3 \end{bmatrix}$ and $B = \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix}$, then

$$\begin{aligned} AB &= \begin{bmatrix} 3 & 1 & -4 \\ 0 & 2 & 1 \\ 5 & -2 & -3 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \\ 1 \end{bmatrix} \\ &= \begin{bmatrix} 3 \times 2 + 1 \times (-1) + (-4) \times 1 \\ 0 \times 2 + 2 \times (-1) + 1 \times 1 \\ 5 \times 2 + (-2) \times (-1) + (-3) \times 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ -1 \\ 9 \end{bmatrix}. \end{aligned}$$

6. If $A = \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$, then

$$\begin{aligned} AB &= \begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 \times 1 + 1 \times (-1) & 1 \times (-1) + 1 \times 1 \\ (-1) \times 1 + (-1) \times (-1) & (-1) \times (-1) + (-1) \times 1 \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}. \end{aligned}$$

Example 10.2

Two firms A and B each produce three goods G_1, G_2, and G_3. The prices per unit for each good from the two firms are represented by the matrix $P = \begin{bmatrix} P_{A1} & P_{A2} & P_{A3} \\ P_{B1} & P_{B2} & P_{B3} \end{bmatrix}$, where the first row gives the prices per unit for goods G_1, G_2, G_3, respectively, supplied by firm A and the second row gives those supplied by firm B.

A consumer wishes to buy quantities Q_1, Q_2, Q_3, respectively, from one of the firms. The matrix $Q = \begin{bmatrix} Q_1 \\ Q_2 \\ Q_3 \end{bmatrix}$ represents the quantities. Then PQ is the 2×1 column matrix $\begin{bmatrix} P_{A1}Q_1 + P_{A2}Q_2 + P_{A3}Q_3 \\ P_{B1}Q_1 + P_{B2}Q_2 + P_{B3}Q_3 \end{bmatrix}$, where the top entry in the column is the cost of buying the goods from firm A and the bottom entry the cost from firm B.

Notes

1. To get the first row entries of the product MN of two matrices M and N, multiply the first row of M in turn by each column of N. Do this for the second row and so on. This constructs MN row by row. So for instance, in Example 10.1.2,

$$\begin{bmatrix} 2 & -1 & 4 \end{bmatrix} \begin{bmatrix} 5 \\ 4 \\ 1 \end{bmatrix} = 10 - 4 + 4 = 10$$

and

$$\begin{bmatrix} 2 & -1 & 4 \end{bmatrix} \begin{bmatrix} 2 \\ -3 \\ 0 \end{bmatrix} = 4 + (-1)(-3) + 0 = 7$$

are the entries of the first row of MN.

2. If M and N are matrices, it is easy to see that the matrix products MN and NM are both defined when and only when M, N are square matrices of the same size. However, MN and NM are not in general the same. See Example 10.1.3 for instance where $AB \neq BA$. Thus, unlike matrix addition, matrix multiplication is not commutative.

   Note that in Example 10.1.2, we cannot define the product NM.

   In Example 10.1.3, the matrices A, B can be multiplied in two ways to give the products AB and BA. To specify AB, for instance, we can say this is the product of A multiplied by B on the right or B multiplied on the left by A.

The **diagonal** of a matrix M consists of all the entries of the form M_{ii}. For instance, the diagonal entries of the matrix

$$M = \begin{bmatrix} 3 & 4 & 1 & 1 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & -5 & 1 \end{bmatrix}$$

are 3, 1, -5. For any matrix M, it is easy to see that M^t and M will have the same diagonal.

The square $n \times n$ matrix in which every diagonal entry is 1 and every entry off the diagonal is 0 is called the **identity** $n \times n$ matrix. This matrix is denoted by I_n; or simply by I if its size is clear from the context.

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \qquad I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Identity matrices act like 'ones' in the sense that $AI_n = I_m A = A$ if A is any $m \times n$ matrix.

A useful matrix rule is

$$(\lambda M) N = M (\lambda N) = \lambda (MN).$$

Essentially, this means that when multiplying matrices together, any scalar factor, such as λ, can be taken 'outside' the multiplication process. For example, if A and B are as in Example 10.1.3, then

$$\left(\frac{1}{3}A\right)(6B) = \left(\frac{1}{3}\right)(6)\, AB = 2AB = \begin{bmatrix} 20 & 4 \\ 2 & -2 \end{bmatrix}.$$

Two more rules, known as the **distributive laws**, allows us to 'open' brackets:

$$A (B + C) = AB + AC$$

and

$$(A+B)\,C = AC + BC.$$

Conversely, the rules can be regarded as one of factorization. For instance, if A, B, C are matrices then:
(a) $2AB - AC = A\,(2B - C)$;
(b) $BA - A = (B - I)\,A$.
(Here, I is the $m \times m$ identity matrix, where m is the number of rows in A.) This is true because $I_m A = A$. (Why would it be wrong to write $B - 1$ instead of $B - I$?)

If M is a square $m \times m$ matrix, we can define powers of M: $M^2 = MM$, $M^3 = M^2M = MMM$, and so on. In general, for any integer $n \geq 1$, we have $M^n = M^{n-1}M$, where we take $M^0 = I_m$, the identity $m \times m$ matrix. It is easy to see that M^n is just M multiplied by itself n times. This defines the n^{th} power M^n of a square matrix for any integer $n \geq 0$, similar to the way in which x^n is defined for a number x.

Example 10.3

If A is the matrix defined by

$$A = \begin{bmatrix} 1 & 3 \\ 2 & 0 \end{bmatrix}$$

compute A^2 and A^3.

Solution.

$$A^2 = \begin{bmatrix} 1 & 3 \\ 2 & 0 \end{bmatrix}\begin{bmatrix} 1 & 3 \\ 2 & 0 \end{bmatrix} = \begin{bmatrix} 7 & 3 \\ 2 & 6 \end{bmatrix}$$

$$A^3 = AA^2 = \begin{bmatrix} 1 & 3 \\ 2 & 0 \end{bmatrix}\begin{bmatrix} 7 & 3 \\ 2 & 6 \end{bmatrix} = \begin{bmatrix} 13 & 21 \\ 14 & 6 \end{bmatrix}$$

The question that naturally arises is whether this analogy of powers of matrices with powers of numbers extends to negative integer exponents. In particular, can we assign any meaning to M^{-1}? It turns out that we can define M^{-1} in certain cases and we can test whether or not M^{-1} exists.

We shall say that a square matrix M is **invertible** if there is a matrix N such that $MN = NM = I$. If N exists, we write M^{-1} for N. We call M^{-1} the **inverse matrix** of M. (For short, we can say 'M inverse' for the matrix M^{-1}.) It is easy to see that $\left(M^{-1}\right)^{-1} = M$.

It is important to note that not all square matrices are invertible. To determine whether or not a matrix is invertible, we need the concept of the **determinant** of a square matrix.

First we consider a 2×2 matrix

$$M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

The **determinant** of M denoted by $|M|$ or

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix},$$

is the number $ad - bc$. For example:

$$\begin{vmatrix} 2 & -1 \\ 3 & 5 \end{vmatrix} = 10 - (-3) = 13,$$

$$\begin{vmatrix} 6 & 3 \\ 4 & 2 \end{vmatrix} = 12 - 12 = 0.$$

It can be shown that a 2×2 matrix is invertible if, and only if, its determinant is not zero. If M is invertible, its inverse is obtained in the following way.

If $M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ and $|M| = ad - bc \neq 0$, then $M^{-1} = \dfrac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$

The matrix M^{-1} is therefore a scalar multiple of the matrix

$$\begin{bmatrix} d & -b \\ -c & a \end{bmatrix};$$

the scalar being $1/|M|$. To illustrate this, perform the matrix multiplication

$$\begin{aligned} \frac{1}{M} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \begin{bmatrix} a & b \\ c & d \end{bmatrix} &= \frac{1}{ad - bc} \begin{bmatrix} ad - bc & 0 \\ 0 & ad - bc \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \\ &= I. \end{aligned}$$

For example,

$$\begin{vmatrix} 6 & 3 \\ 4 & 2 \end{vmatrix} = 0,$$

so the matrix

$$\begin{bmatrix} 6 & 3 \\ 4 & 2 \end{bmatrix}$$

is not invertible.

Since

$$\begin{vmatrix} 3 & 5 \\ 2 & 4 \end{vmatrix} = 12 - 10 = 2,$$

the matrix

$$M = \begin{bmatrix} 3 & 5 \\ 2 & 4 \end{bmatrix}$$

is invertible and

$$M^{-1} = \frac{1}{2}\begin{bmatrix} 4 & -5 \\ -2 & 3 \end{bmatrix}.$$

Example 10.4

Given that

$$A = \begin{bmatrix} 1 & -2 \\ 3 & 4 \end{bmatrix} \text{ and } B = \begin{bmatrix} 4 & 1 \\ 0 & -3 \end{bmatrix},$$

determine the 2×2 matrices X and Y satisfying

1. $AX = B$;
2. $YA = B$.

Solution. Since $|A| = 4 - (-6) = 10 \neq 0$, then A^{-1} exists and

$$A^{-1} = \frac{1}{10}\begin{bmatrix} 4 & 2 \\ -3 & 1 \end{bmatrix}.$$

1. Multiply both sides of the equation $AX = B$ on the *left* by A^{-1} to get $A^{-1}AX = A^{-1}B$. That is, $IX = A^{-1}B$ and so $X = A^{-1}B$. Therefore,

$$X = \frac{1}{10}\begin{bmatrix} 4 & 2 \\ -3 & 1 \end{bmatrix}\begin{bmatrix} 4 & 1 \\ 0 & -3 \end{bmatrix} = \frac{1}{10}\begin{bmatrix} 16 & -2 \\ -12 & -6 \end{bmatrix}.$$

(Here we used the fact that $A^{-1}A = I$, the identity 2×2 matrix, and that $IX = X$.)

2. Multiply both sides of the equation on the *right* by A^{-1} to get $YAA^{-1} = BA^{-1}$, which gives $Y = BA^{-1}$. Therefore,

$$Y = \frac{1}{10}\begin{bmatrix} 4 & 1 \\ 0 & -3 \end{bmatrix}\begin{bmatrix} 4 & 2 \\ -3 & 1 \end{bmatrix} = \frac{1}{10}\begin{bmatrix} 13 & 9 \\ 9 & -3 \end{bmatrix}.$$

10.3 Solutions of Linear Systems of Equations

Two simultaneous equations in two unknowns can be solved uniquely if the matrix of coefficients is invertible. Specifically, if

$$ax + by = c$$

$$px + qy = r$$

are the equations in the unknowns x, y, the **matrix of coefficients** is

$$M = \begin{bmatrix} a & b \\ p & q \end{bmatrix}.$$

Then

$$M \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} ax + by \\ px + qy \end{bmatrix} = \begin{bmatrix} c \\ r \end{bmatrix}$$

so that the two given simultaneous equations are equivalent to one matrix equation:

$$M \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} c \\ r \end{bmatrix}.$$

If M is invertible, then multiplying both sides on the left by M^{-1} gives the unique solution

$$\begin{bmatrix} x \\ y \end{bmatrix} = M^{-1} \begin{bmatrix} c \\ r \end{bmatrix}.$$

Example 10.5

Solve the linear system of equations

$$\begin{aligned} 7x + 3y &= 41 \\ 3x + 2y &= 19 \end{aligned} \qquad (10.1)$$

Solution. The matrix form of (10.1) is

$$\begin{bmatrix} 7 & 3 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 41 \\ 19 \end{bmatrix}.$$

Since

$$\begin{vmatrix} 7 & 3 \\ 3 & 2 \end{vmatrix} = 14 - 9 = 5 \neq 0,$$

the matrix of coefficients $M = \begin{bmatrix} 7 & 3 \\ 3 & 2 \end{bmatrix}$ is invertible and $M^{-1} = \frac{1}{5}\begin{bmatrix} 2 & -3 \\ -3 & 7 \end{bmatrix}$. The solution is

$$\begin{bmatrix} x \\ y \end{bmatrix} = M^{-1}\begin{bmatrix} 41 \\ 19 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 2 & -3 \\ -3 & 7 \end{bmatrix}\begin{bmatrix} 41 \\ 19 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 25 \\ 10 \end{bmatrix}.$$

So $x = \frac{25}{5} = 5$ and $y = \frac{10}{5} = 2$.

Example 10.6

The demand and supply equations for a good are given by $P + 4Q_D = 70$ and $P - Q_S = 5$. Determine the equilibrium price and quantity.

Solution. To find the equilibrium price P and quantity $Q = Q_D = Q_S$, we solve the equations

$$\begin{aligned} P + 4Q &= 70 \\ P - Q &= 5. \end{aligned}$$

In matrix form, this is equivalent to

$$\begin{bmatrix} 1 & 4 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} P \\ Q \end{bmatrix} = \begin{bmatrix} 70 \\ 5 \end{bmatrix}.$$

There is a unique solution:

$$\begin{aligned} \begin{bmatrix} P \\ Q \end{bmatrix} &= \begin{bmatrix} 1 & 4 \\ 1 & -1 \end{bmatrix}^{-1}\begin{bmatrix} 70 \\ 5 \end{bmatrix} \\ &= -\frac{1}{5}\begin{bmatrix} -1 & -4 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 70 \\ 5 \end{bmatrix} \\ &= -\frac{1}{5}\begin{bmatrix} -90 \\ -65 \end{bmatrix} \\ &= \begin{bmatrix} 18 \\ 13 \end{bmatrix}. \end{aligned}$$

The equilibrium values are therefore $P = 18$, $Q = 13$.

10.4 Cramer's Rule

This simple rule allows us to express the solution of two simultaneous equations in two unknowns explicitly, assuming there is a unique solution. Given the equations

$$ax + by = c$$
$$px + qy = r$$

then Cramer's rule states that if $\begin{vmatrix} a & b \\ p & q \end{vmatrix} \neq 0$ (equivalently, if the matrix of coefficients is invertible), then

$$x = \frac{\begin{vmatrix} c & b \\ r & q \end{vmatrix}}{\begin{vmatrix} a & b \\ p & q \end{vmatrix}} \quad \text{and} \quad y = \frac{\begin{vmatrix} a & c \\ p & r \end{vmatrix}}{\begin{vmatrix} a & b \\ p & q \end{vmatrix}}.$$

Observe that the determinant $\begin{vmatrix} c & b \\ r & q \end{vmatrix}$ in the equation for x is obtained by replacing the column of coefficients of x in the determinant of the matrix of coefficients by the column of constants on the right-hand side of the equation; and similarly for y.

Example 10.7

Solve the simultaneous equations

$$6x + 7y = 10$$
$$4x + 5y = 8$$

using Cramer's rule.

Solution. Using Cramer's rule, we have

$$x = \frac{\begin{vmatrix} 10 & 7 \\ 8 & 5 \end{vmatrix}}{\begin{vmatrix} 6 & 7 \\ 4 & 5 \end{vmatrix}} = \frac{50 - 56}{30 - 28} = -3$$

and

$$y = \frac{\begin{vmatrix} 6 & 10 \\ 4 & 8 \end{vmatrix}}{\begin{vmatrix} 6 & 7 \\ 4 & 5 \end{vmatrix}} = \frac{48 - 40}{2} = 4.$$

(In geometric terms, x and y are the coordinates of the point of intersection of the two lines with equations $6x + 7y = 10$ and $4x + 5y = 8$.)

10.5 More Determinants

In Section 10.2, we introduced for 2×2 matrices the idea of a determinant. Now we consider 3×3 matrices. The determinant of a 3×3 matrix

$$M = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

is the number

$$a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - b \begin{vmatrix} d & f \\ g & i \end{vmatrix} + c \begin{vmatrix} d & e \\ g & h \end{vmatrix}$$

which we denote by $|M|$ or $\begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix}$.

Example 10.8

Evaluate the determinant

$$\begin{vmatrix} 2 & 3 & 5 \\ 3 & 1 & 2 \\ 1 & 4 & 3 \end{vmatrix}.$$

Solution. Using the definition of a determinant of a 3×3 matrix

$$\begin{aligned} \begin{vmatrix} 2 & 3 & 5 \\ 3 & 1 & 2 \\ 1 & 4 & 3 \end{vmatrix} &= 2\begin{vmatrix} 1 & 2 \\ 4 & 3 \end{vmatrix} - 3\begin{vmatrix} 3 & 2 \\ 1 & 3 \end{vmatrix} + 5\begin{vmatrix} 3 & 1 \\ 1 & 4 \end{vmatrix} \\ &= 2(3-8) - 3(9-2) + 5(12-1) \\ &= -10 - 21 + 55 \\ &= 24. \end{aligned}$$

As in the 2×2 case, it is true that a 3×3 matrix M is invertible if, and only if, its determinant $|M| \neq 0$; but it is not as easy to describe M^{-1} in the 3×3 case. A method for constructing M^{-1} (when it exists) is known as the **adjoint method**. To describe this, we need the concept of a **cofactor**.

The (i, j)-cofactor of a 3×3 matrix M is the determinant of the 2×2 matrix obtained by deleting the i^{th} row and j^{th} column of M, multiplied by $(-1)^{i+j}$.

Note that $(-1)^{i+j}$ is $+1$ or -1 according to whether $i+j$ is even or odd, respectively. The pattern

$$\begin{matrix} + & - & + \\ - & + & - \\ + & - & + \end{matrix}$$

gives the (i, j) positions corresponding to $+1$ and -1. For example, the $(3, 1)$-cofactor of the matrix

$$M = \begin{bmatrix} 3 & 2 & -4 \\ 1 & 3 & 2 \\ 4 & 1 & -1 \end{bmatrix}$$

is

$$+\begin{vmatrix} 2 & -4 \\ 3 & 2 \end{vmatrix} = 4 + 12 = 16.$$

Similarly, the $(2, 3)$-cofactor is

$$-\begin{vmatrix} 3 & 2 \\ 4 & 1 \end{vmatrix} = -(3-8) = 5.$$

Note The value of the (i, j)-cofactor does **not** depend on the value of the $(i, j)-$entry.

The **cofactor matrix** of a 3×3 matrix M is the 3×3 matrix whose (i, j)-entry is the (i, j)-cofactor of M. For example, the cofactor matrix of the matrix M in the previous example is

$$\begin{bmatrix} \begin{vmatrix} 3 & 2 \\ 1 & -1 \end{vmatrix} & -\begin{vmatrix} 1 & 2 \\ 4 & -1 \end{vmatrix} & \begin{vmatrix} 1 & 3 \\ 4 & 1 \end{vmatrix} \\ -\begin{vmatrix} 2 & -4 \\ 1 & -1 \end{vmatrix} & \begin{vmatrix} 3 & -4 \\ 4 & -1 \end{vmatrix} & -\begin{vmatrix} 3 & 2 \\ 4 & 1 \end{vmatrix} \\ \begin{vmatrix} 2 & -4 \\ 3 & 2 \end{vmatrix} & -\begin{vmatrix} 3 & -4 \\ 1 & 2 \end{vmatrix} & \begin{vmatrix} 3 & 2 \\ 1 & 3 \end{vmatrix} \end{bmatrix} = \begin{bmatrix} -5 & 9 & -11 \\ -2 & 13 & 5 \\ 16 & -10 & 7 \end{bmatrix}.$$

The transpose of the cofactor matrix of a 3×3 matrix M is known as the **adjoint matrix**, denoted by adj$M$.

The following statement describes how M^{-1} may be computed by the method known as the adjoint method.

> If M is a 3×3 matrix and $|M| \neq 0$, then $M^{-1} = \dfrac{1}{|M|}\text{adj}M.$

Thus M^{-1} is a scalar multiple of adj$M$; the scalar being the number $\frac{1}{|M|}$.

Example 10.9

Find the inverse of the matrix

$$M = \begin{bmatrix} 3 & 2 & -4 \\ 1 & 3 & 2 \\ 4 & 1 & -1 \end{bmatrix}.$$

Solution. The determinant of M is

$$\begin{aligned} \begin{vmatrix} 3 & 2 & -4 \\ 1 & 3 & 2 \\ 4 & 1 & -1 \end{vmatrix} &= 3\begin{vmatrix} 3 & 2 \\ 1 & -1 \end{vmatrix} - 2\begin{vmatrix} 1 & 2 \\ 4 & -1 \end{vmatrix} + (-4)\begin{vmatrix} 1 & 3 \\ 4 & 1 \end{vmatrix} \\ &= 3(-5) - 2(-9) - 4(-11) \\ &= -15 + 18 + 44 \\ &= 47. \end{aligned}$$

Since $|M| = 47 \neq 0$, then M^{-1} exists and $M^{-1} = \frac{1}{|M|}\text{adj}M$, where adj$M$ is the transpose of the cofactor matrix of M (see above). Therefore

$$M^{-1} = \frac{1}{47}\begin{bmatrix} -5 & -2 & 16 \\ 9 & 13 & -10 \\ -11 & 5 & 7 \end{bmatrix}.$$

If we look again at the definition of the determinant of a 3×3 matrix M, we see that $|M|$, the determinant of M, is obtained by multiplying each entry in the first row of M by the corresponding cofactor and adding.

An interesting fact is that there is nothing special about the first row of M. Multiplying each term in any row (or column) by the corresponding cofactor and adding gives the same number; namely $|M|$. (This is a useful check that the cofactor matrix has been computed correctly.)

For example, for the matrix

$$M = \begin{bmatrix} 3 & 2 & -4 \\ 1 & 3 & 2 \\ 4 & 1 & -1 \end{bmatrix},$$

the cofactor matrix is

$$\begin{bmatrix} -5 & 9 & -11 \\ -2 & 13 & 5 \\ 16 & -10 & 7 \end{bmatrix}.$$

The determinant of M is $|M| = 3(-5)+2(9)+(-4)(-11) = -15+18+44 = 47$. This is the **expansion** of the determinant of M by its first row. Expanding by, say, the third column gives $(-4)(-11) + 2(5) + (-1)(7) = 44 + 10 - 7 = 47$ again.

Incidentally, if we expand a row using the cofactor of another row, we will always get 0. For example, expanding along the first row of M using the cofactors of the third row gives $3(16) + 2(-10) + (-4)(7) = 48 - 20 - 28 = 0$. These expansion properties of determinants form the basis of the adjoint method for finding inverses and of Cramer's rule.

Example 10.10

Determine the inverse of the matrix M, where

$$M = \begin{bmatrix} 1 & -3 & -2 \\ 4 & 1 & 2 \\ 0 & 6 & 5 \end{bmatrix}.$$

Solution. The cofactor matrix of M is

$$C = \begin{bmatrix} -7 & -20 & 24 \\ 3 & 5 & -6 \\ -4 & -10 & 13 \end{bmatrix}.$$

The determinant is

$$|M| = 1(-7) - (-3)(20) + (-2)(24) = -7 + 60 - 48 = 5.$$

Therefore

$$M^{-1} = \frac{1}{|M|}\text{adj}M = \frac{1}{5}C^t = \frac{1}{5}\begin{bmatrix} -7 & 3 & -4 \\ -20 & 5 & -10 \\ 24 & -6 & 13 \end{bmatrix}. \tag{10.2}$$

Solving a system of two simultaneous equations in two unknowns using inverse matrices or Cramer's rule extends naturally to the case of three equations in three unknowns. If the matrix

$$M = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix}$$

of coefficients of a system of three equations

$$\begin{aligned} a_1x + a_2y + a_3z &= p \\ b_1x + b_2y + b_3z &= q \\ c_1x + c_2y + c_3z &= r \end{aligned}$$

has non-zero determinant and so is invertible, then the simultaneous equations have a unique solution given by:

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = M^{-1} \begin{bmatrix} p \\ q \\ r \end{bmatrix}.$$

Alternatively, we can use Cramer's rule, extended from Section 10.4 in an obvious way to three equations in three unknowns. This gives

$$x = \frac{|M_x|}{|M|}, \quad y = \frac{|M_y|}{|M|}, \quad z = \frac{|M_z|}{|M|},$$

where M_x is the matrix obtained by replacing the column $\begin{bmatrix} a_1 \\ b_1 \\ c_1 \end{bmatrix}$ of coefficients of x in M by the column of constants $\begin{bmatrix} p \\ q \\ r \end{bmatrix}$. Similarly for M_y and M_z.

Example 10.11

Solve the system of three simultaneous equations

$$\begin{aligned} x - 3y - 2z &= 5, \\ 4x + y + 2z &= 116, \\ 6y + 5z &= 47. \end{aligned}$$

Solution. In matrix form, the system can be written as one equation:

$$M \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 5 \\ 116 \\ 47 \end{bmatrix},$$

where

$$M = \begin{bmatrix} 1 & -3 & -2 \\ 4 & 1 & 2 \\ 0 & 6 & 5 \end{bmatrix}$$

is the matrix of coefficients.

From Example 10.10, we know that M is invertible and M^{-1} is given in (10.2). Therefore

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = M^{-1}\begin{bmatrix} 5 \\ 116 \\ 47 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} -7 & 3 & -4 \\ -20 & 5 & -10 \\ 24 & -6 & 13 \end{bmatrix}\begin{bmatrix} 5 \\ 116 \\ 47 \end{bmatrix}$$

$$= \frac{1}{5}\begin{bmatrix} -35+348-188 \\ -100+580-470 \\ 120-696+611 \end{bmatrix} = \frac{1}{5}\begin{bmatrix} 125 \\ 10 \\ 35 \end{bmatrix} = \begin{bmatrix} 25 \\ 2 \\ 7 \end{bmatrix}.$$

Therefore $x = 25$, $y = 2$, and $z = 7$.

Alternatively, we can use Cramer's rule. Here,

$$M_x = \begin{bmatrix} 5 & -3 & -2 \\ 116 & 1 & 2 \\ 47 & 6 & 5 \end{bmatrix},$$

so

$$|M_x| = 5(5-12) - (-3)(580-94) - 2(696-47) = 125.$$

Therefore

$$x = \frac{|M_x|}{|M|} = \frac{125}{5} = 25,$$

and then y, z are obtained in a similar way, noting that

$$M_y = \begin{bmatrix} 1 & 5 & -2 \\ 4 & 116 & 2 \\ 0 & 47 & 5 \end{bmatrix}$$

and

$$M_z = \begin{bmatrix} 1 & -3 & 5 \\ 4 & 1 & 116 \\ 0 & 6 & 47 \end{bmatrix}.$$

Example 10.12

A consumer's utility function is $U(x, y) = xy + x + 2y$, where x is the number of units of good G_x and y the number of units of good G_y. The price per unit of G_x is 2 (units of money), and the price per unit of G_y is 5. What is the maximum utility if the consumer's budget is 91?

Solution. The objective function is U and the constraint is $2x + 5y = 91$. Form the Lagrangian function

$$F = xy + x + 2y + \lambda(91 - 2x - 5y).$$

The stationary points of F occur where

$$\begin{aligned} 0 &= F_x = y + 1 - 2\lambda, \\ 0 &= F_y = x + 2 - 5\lambda \\ \text{and} \quad 0 &= F_\lambda = 91 - 2x - 5y. \end{aligned}$$

We therefore have the following three equations in the three unknowns x, y, and λ:

$$\begin{aligned} -y + 2\lambda &= 1, \\ x \qquad - 5\lambda &= -2, \\ 2x + 5y \qquad &= 91. \end{aligned}$$

We can easily solve these as before by eliminating λ between the first two equations and then using the constraint equation to solve for x and y.

Another method is to use matrix inversion. The matrix of coefficients is

$$M = \begin{bmatrix} 0 & -1 & 2 \\ 1 & 0 & -5 \\ 2 & 5 & 0 \end{bmatrix},$$

so that

$$M \begin{bmatrix} x \\ y \\ \lambda \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ 91 \end{bmatrix}.$$

Using the adjoint method, we can compute

$$M^{-1} = \frac{1}{20} \begin{bmatrix} 25 & 10 & 5 \\ -10 & -4 & 2 \\ 5 & -2 & 1 \end{bmatrix}.$$

Then

$$\begin{bmatrix} x \\ y \\ \lambda \end{bmatrix} = M^{-1} \begin{bmatrix} 1 \\ -2 \\ 91 \end{bmatrix} = \frac{1}{20} \begin{bmatrix} 25 - 20 + 455 \\ -10 + 8 + 182 \\ 5 + 4 + 91 \end{bmatrix} = \frac{1}{20} \begin{bmatrix} 460 \\ 180 \\ 100 \end{bmatrix}.$$

Therefore, $x = 23$, $y = 9$, and $\lambda = 5$. So the maximum utility is $U(23, 9) = 248$.

Cramer's rule could also have been used to solve the previous problem. We use it for the next problem.

Example 10.13

Find the equilibrium prices of three interdependent commodities whose prices P_1, P_2, P_3 satisfy:

$$\begin{aligned} P_1 + P_2 + 3P_3 &= 37, \\ 3P_1 + 2P_2 + 4P_3 &= 79, \\ 2P_1 + 3P_2 + 5P_3 &= 76. \end{aligned}$$

Solution. The determinant of the matrix of coefficients is

$$\begin{vmatrix} 1 & 1 & 3 \\ 3 & 2 & 4 \\ 2 & 3 & 5 \end{vmatrix} = (10 - 12) - (15 - 8) + 3(9 - 4) = 6.$$

Therefore by Cramer's rule

$$P_1 = \begin{vmatrix} 37 & 1 & 3 \\ 79 & 2 & 4 \\ 76 & 3 & 5 \end{vmatrix} \div 6 = 90 \div 6 = 15,$$

$$P_2 = \begin{vmatrix} 1 & 37 & 3 \\ 3 & 79 & 4 \\ 2 & 76 & 5 \end{vmatrix} \div 6 = 42 \div 6 = 7,$$

$$P_3 = \begin{vmatrix} 1 & 1 & 37 \\ 3 & 2 & 79 \\ 2 & 3 & 76 \end{vmatrix} \div 6 = 30 \div 6 = 5.$$

10.6 Special Cases

The solutions of systems of simultaneous equations considered in this chapter have been for the case when there are as many unknowns as equations and also the matrix of coefficients is invertible (or, equivalently, has non-zero determinant). In this case, the solution is unique. What happens if the matrix of coefficients is not invertible; that is, if its determinant is zero? In this case, the system is either inconsistent and has no simultaneous solution or else it has infinitely many solutions.

To illustrate this point with a very simple example, consider the two systems of equations

(a) $$\begin{aligned} x + 3y &= 4 \\ 2x + 6y &= 8 \end{aligned}$$

(b) $$\begin{aligned} x + 3y &= 5 \\ 2x + 6y &= 9 \end{aligned}$$

In both cases, the matrix of coefficients is

$$\begin{bmatrix} 1 & 3 \\ 2 & 6 \end{bmatrix}$$

and its determinant is

$$\begin{vmatrix} 1 & 3 \\ 2 & 6 \end{vmatrix} = 0.$$

In case (a), the second equation is twice the first equation. (The second equation is therefore redundant.) Therefore, any solution of the equation

$$x + 3y = 4$$

will satisfy both equations simultaneously. So there are infinitely many solutions because for any choice of value α for y, the equation $x + 3y = 4$ is satisfied by $x = 4 - 3\alpha$, $y = \alpha$.

In case (b), multiplying the first equation by 2 gives $2x + 6y = 10$, which is inconsistent with the second equation. So the two given equations cannot have any simultaneous solution.

For the case of three equations in three unknowns, a similar conclusion holds. If the matrix of coefficients has zero determinant, then the equations have either no simultaneous solutions or infinitely many. A method known as Gauss-Jordan elimination can be applied to a system of equations to reduce it to a simple form that eliminates redundant equations. The reduction is akin to the usual process of simplifying equations; subtracting multiples of one equation from another to eliminate variables between them. We shall say no more about this in this book.

EXERCISES

10.1. Find the inverse, where it exists, of each of the matrices

$$\begin{bmatrix} 5 & 2 \\ 6 & 3 \end{bmatrix}, \quad \begin{bmatrix} 3 & -1 \\ -15 & 5 \end{bmatrix}, \quad \begin{bmatrix} 3 & -2 \\ -4 & 3 \end{bmatrix},$$

$$\begin{bmatrix} -2 & 2 \\ 4 & -3 \end{bmatrix}, \quad \begin{bmatrix} 3 & 0 \\ 0 & 5 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

10.2. The equilibrium prices P_1, P_2 for two goods satisfy the equations

$$\begin{aligned} 4P_1 - 3P_2 &= 11 \\ -6P_1 + 7P_2 &= 8. \end{aligned}$$

Express these equations in matrix form. Hence, by inverting the matrix, solve for P_1, P_2. Also solve these equations by Cramer's rule.

10.3. Let $B = \begin{bmatrix} 3 & 2 \\ 0 & -1 \end{bmatrix}$, $C = \begin{bmatrix} 2 & 2 \\ -1 & -2 \end{bmatrix}$, $D = \begin{bmatrix} 1 & 1 \\ 0 & 2 \end{bmatrix}$.

If $B = ED$, $B = DF$, and $AB = AC + D$, determine E, F, and A.

10.4. If $XB = X + C$, where B and C are as in the previous exercise, find the 2×2 matrix X.

10.5. Determine which of the matrices

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 2 & 2 & 3 \\ 3 & 1 & 5 \\ 1 & -7 & 3 \end{bmatrix}$$

are invertible.

(You do not have to find inverses.)

10.6. Find (i) $|A|$; (ii) the cofactor matrix of A; (iii) adj A, where A is the matrix

$$\begin{bmatrix} 2 & 5 & 3 \\ 4 & 5 & 2 \\ 7 & 7 & 1 \end{bmatrix}.$$

Hence find A^{-1}.

10.7. Determine the equilibrium prices P_1, P_2, P_3 of three interdependent commodities that satisfy

$$\begin{array}{ccccccc} 2P_1 & + & 5P_2 & + & 3P_3 & = & 136 \\ 4P_1 & + & 5P_2 & + & 2P_3 & = & 132 \\ 7P_1 & + & 7P_2 & + & P_3 & = & 160 \end{array}$$

using matrices or Cramer's rule.

11 Integration

11.1 Introduction

Differentiating a function $f(x)$ gives its derivative $f'(x)$, which is also a function of x. Geometrically, we can view $f'(x)$ as giving the slope of the tangent at any point on the graph of $y = f(x)$ or, equivalently, the rate of change of $f(x)$ with respect to x at that point.

Integrating a function $f(x)$ also gives a function $F(x)$ of x whose derivative $F'(x) = f(x)$. For this reason, integration can be used to recover an economic function from its corresponding marginal. For instance, it can be used to find the total revenue function TR given the marginal revenue MR.

A useful geometric interpretation of $F(x)$ is as a measure of the area under the graph of $y = f(x)$. This can be used to compute either the consumer's or producer's surplus. Another application is to compute the extra cost to a company for increasing production, given the company's marginal cost function.

Integration can be regarded as the inverse operation to differentiation in that it operates on a function $f(x)$ to produce a function $F(x)$ whose derivative $F'(x) = f(x)$. The function $F(x)$ is the **integral** of $f(x)$ with respect to the variable x. Symbolically, we write

$$F(x) = \int f(x)dx$$

to mean

$$F'(x) = f(x).$$

Integrating a function $f(x)$ means finding its integral; that is, finding a function whose derivative is $f(x)$.

For example, to integrate the function $f(x) = x$, with respect to x, means to find a function whose derivative is x. We know that

$$\frac{d}{dx}(x^2) = 2x;$$

so

$$\frac{d}{dx}\left(\frac{1}{2}x^2\right) = \frac{1}{2}\frac{d}{dx}(x^2) = x.$$

Therefore,

$$x^2 = \int 2x dx$$

and

$$\frac{1}{2}x^2 = \int x dx.$$

However, note that if k is any constant, then also

$$\frac{d}{dx}(x^2 + k) = 2x,$$

since the derivative of k is 0. Therefore, we should really write

$$\int 2x dx = x^2 + k,$$

where k is an arbitrary constant, known as a **constant of integration**.

More generally, if

$$\int f(x) dx = F(x),$$

then $F(x)$ is unique only to within the addition of an arbitrary constant. That is,

$$\int f(x) dx = F(x) + k,$$

where k is any constant. As we shall see later, in certain circumstances we can determine k.

To sum up, if u, v are functions of x, then

$$u = \int v dx \qquad \text{means} \qquad \frac{du}{dx} = v.$$

Example 11.1

Determine u given that $\dfrac{du}{dx} = 2x$ and that $u = 5$ when $x = 1$.

Solution. Since $\frac{du}{dx} = 2x$, then $u = \int 2x dx = x^2 + k$, where k is some constant. When $x = 1$, then $5 = u = 1^2 + k = 1 + k$. Therefore, $k = 4$ and so $u = x^2 + 4$.

Our aim now is to see how to integrate polynomials in general. We have noted already that $\frac{1}{2}x^2 = \int x dx$. More generally, we know from (6.5) that

$$\frac{d}{dx}(x^n) = nx^{n-1}.$$

Therefore if $n \neq 0$, then

$$\frac{d}{dx}\left(\frac{1}{n}x^n\right) = \frac{1}{n}\frac{d}{dx}(x^n) = x^{n-1}.$$

It follows that $\frac{1}{n}x^n = \int x^{n-1} dx$. If we write m for $n - 1$, so that $n = m + 1$, then we can rewrite this as

$$\boxed{\frac{1}{m+1}x^{m+1} = \int x^m dx \qquad (m \neq -1)}$$

Thus we can now integrate any power of x, except x^{-1}. However, from (6.12) we know that

$$\frac{d}{dx}(\ln x) = \frac{1}{x} = x^{-1}.$$

Therefore

$$\boxed{\ln x = \int x^{-1} dx}$$

Example 11.2

Integrate each of the following functions with respect to x

1. x^5
2. x^{-2}
3. $x^{\frac{1}{2}}$
4. 1.

Solution.

1. $\displaystyle\int x^5 dx = \frac{1}{5+1}x^{5+1} = \frac{1}{6}x^6.$
2. $\displaystyle\int x^{-2} dx = \frac{1}{-2+1}x^{-2+1} = -x^{-1}.$

3. $\int x^{\frac{1}{2}}dx = \frac{1}{\frac{1}{2}+1}x^{\frac{1}{2}+1} = \frac{2}{3}x^{\frac{3}{2}}.$
4. $\int 1dx = \int x^0 dx = \frac{1}{0+1}x^{0+1} = x.$

Note that it is customary to write $\int dx$ rather than $\int 1dx$.

We can check each of these integrations by differentiating the right-hand side to see if we get what is under the integral sign. For instance

$$\frac{d}{dx}\left(\frac{2}{3}x^{\frac{3}{2}}\right) = \frac{2}{3}\frac{d}{dx}(x^{\frac{3}{2}}) = \frac{2}{3} \times \frac{3}{2}x^{\frac{3}{2}-1} = x^{\frac{1}{2}}$$

in the third example and

$$\frac{dx}{dx} = 1$$

in the last.

11.2 Rules of Integration

1. If $f(x)$ is a function of x and α is any constant, then

$$\int \alpha f(x)dx = \alpha \int f(x)dx.$$

   (Thus a constant factor may be taken outside the integral sign.)

2. If u, v are functions of x, then

$$\int (u+v)dx = \int udx + \int vdx.$$

That is, to integrate the sum of two functions, integrate each function and add the two integrals. Similarly for differences:

$$\int (u-v)dx = \int udx - \int vdx.$$

These two rules combined give the rule that if $u, v, \ldots$ are functions of x and if $\alpha, \beta, \ldots$ are constants, then

$$\int (\alpha u + \beta v + \ldots)dx = \alpha \int udx + \beta \int vdx + \ldots$$

Example 11.3

Integrate each of the following functions with respect to x

1. $4x^2$
2. $4x^2 - 3x + 5$
3. $(3x - 1)x$
4. $\dfrac{2x - 3}{x}$
5. $(8 - \sqrt{x})$.

Solution.

1. $\int 4x^2 dx = 4 \int x^2 dx = 4 \times \frac{1}{2+1} x^{2+1} = \frac{4}{3} x^3$. Strictly, the answer is $\frac{4}{3}x^3 + k$ where k is any constant. However in this and in subsequent examples, the arbitrary constant is tacitly understood and only noted if required.
2. $\int (4x^2 - 3x + 5) dx = 4 \int x^2 dx - 3 \int x dx + 5 \int dx = 4 \times \frac{1}{3} x^3 - 3 \times \frac{1}{2} x^2 + 5x = \frac{4}{3} x^3 - \frac{3}{2} x^2 + 5x$. (Recall that $\int dx = \int 1 dx = x$.)
3. $\int (3x - 1) x dx = \int (3x^2 - x) dx = 3 \int x^2 dx - \int x dx = 3 \times \frac{1}{3} x^3 - \frac{1}{2} x^2 = x^3 - \frac{1}{2} x^2$.
4. $\displaystyle \int \frac{2x - 3}{x} dx = \int \left(2 - \frac{3}{x}\right) dx = 2 \int dx - 3 \int x^{-1} dx = 2x - 3 \ln x.$
5. $\int (8 - \sqrt{x}) dx = \int (8 - x^{\frac{1}{2}}) dx = 8 \int dx - \int x^{\frac{1}{2}} dx = 8x - \frac{1}{\frac{1}{2} + 1} x^{\frac{1}{2} + 1} = 8x - \frac{2}{3} x^{\frac{3}{2}}$.

In the next three examples, integration is used to determine a standard function in economics when the corresponding marginal function is given.

Example 11.4

A firm's marginal cost function is $MC = Q^2 + 3Q + 8$. Find the total cost function TC if the fixed costs are 250 units of money.

Solution. By definition,

$$\frac{d}{dQ}(TC) = MC = Q^2 + 3Q + 8.$$

Therefore,

$$TC = \int (Q^2 + 3Q + 8) dQ.$$

(Here the variable is now Q rather than x.) Therefore,

$$TC = \int Q^2 dQ + 3 \int Q dQ + 8 \int dQ = \frac{1}{3} Q^3 + 3 \left(\frac{1}{2} Q^2\right) + 8Q + k,$$

where k is some constant. When the output Q is 0, the only costs are the fixed costs. So $TC = 250$ when $Q = 0$. This means that $250 = k$. Therefore,

$$TC = \frac{1}{3}Q^3 + \frac{3}{2}Q^2 + 8Q + 250.$$

Example 11.5

Given that the marginal propensity to consume

$$MPC = 0.15 + \frac{0.2}{\sqrt{Y}},$$

where Y denotes income, find the consumption function C and savings function S if consumption is 135 units when $Y = 100$ money units.

Solution. Since, by definition,

$$MPC = \frac{dC}{dY},$$

then

$$\frac{dC}{dY} = 0.15 + \frac{0.2}{\sqrt{Y}} = 0.15 + 0.2Y^{-\frac{1}{2}}.$$

Therefore,

$$\begin{aligned} C &= \int (0.15 + 0.2Y^{-\frac{1}{2}})dY \\ &= 0.15 \int dY + 0.2 \int Y^{-\frac{1}{2}} dY \\ &= 0.15Y + 0.2 \times \frac{1}{(-\frac{1}{2}) + 1} Y^{(-\frac{1}{2})+1} + k, \end{aligned}$$

where k is some constant. Therefore,

$$C = 0.15Y + 0.4Y^{\frac{1}{2}} + k.$$

Since $C = 135$ when $Y = 100$, then

$$135 = 0.15 \times 100 + 0.4 \times 100^{\frac{1}{2}} + k = 19 + k.$$

So $k = 116$ and therefore

$$C = 0.15Y + 0.4\sqrt{Y} + 116.$$

Since $Y = C + S$, then $S = Y - C$. Therefore,

$$S = 0.85Y - 0.4\sqrt{Y} - 116.$$

Example 11.6

If the marginal revenue function $MR = 15 - 6Q$, determine the total revenue function TR and the demand function.

Solution. Since, by definition,

$$\frac{d}{dQ}(TR) = MR = 15 - 6Q,$$

then

$$TR = \int (15 - 6Q)dQ = 15 \int dQ - 6 \int QdQ = 15Q - 6\left(\frac{1}{2}Q^2\right) + k,$$

where k is some constant. Therefore,

$$TR = 15Q - 3Q^2 + k.$$

Obviously $TR = 0$ when demand $Q = 0$. So $k = 0$ and therefore

$$TR = 15Q - 3Q^2.$$

Since $TR = PQ$, where P is the unit price of the good, then

$$P = \frac{TR}{Q} = \frac{15Q - 3Q^2}{Q} = 15 - 3Q.$$

The demand function is therefore $P = 15 - 3Q$.

Before ending this section, we mention one more standard integral. We saw in (6.11) that

$$\frac{d}{dx}(e^{ax}) = ae^{ax}$$

if a is a constant. It follows that if $a \neq 0$, then

$$\frac{d}{dx}\left(\frac{1}{a}e^{ax}\right) = \frac{1}{a}\frac{d}{dx}(e^{ax}) = e^{ax}.$$

Therefore,

$$\boxed{\frac{1}{a}e^{ax} = \int e^{ax}dx.}$$

Example 11.7

Integrate the function $5e^{-2x}$ with respect to x.

Solution.

$$\int 5e^{-2x}dx = 5\int e^{-2x}dx = 5\left(\frac{1}{-2}\right)e^{-2x} = -\frac{5}{2}e^{-2x}.$$

Example 11.8

A model for the population N (in millions) of a certain country over 10 years, from the beginning of the year 2000 until the end of 2010, assumes the population will decrease exponentially and that the rate of decrease is given by

$$\frac{dN}{dt} = -15e^{-0.5t}$$

where t is the number of years since the beginning of the year 2000.

Express N as a function of t, given that at the start of 2000, the population is 100 million.

What will the population be at the end of 2010? In what year will the population fall to 75 million? If we assume this model is valid indefinitely, what population is predicted in the long run?

Solution. From the given expression for $\frac{dN}{dt}$ we deduce

$$N = \int(-15e^{-0.5t})dt = -15\int e^{-0.5t}dt.$$

Therefore,

$$N = -15\left(\frac{1}{-0.5}\right)e^{-0.5t} + k = 30e^{-0.5t} + k,$$

where k is a constant. Since we are given that $N = 100$ when $t = 0$, then $100 = 30e^0 + k = 30 + k$. So $k = 70$. Therefore

$$N = 30e^{-0.5t} + 70.$$

At the end of 2010, $t = 10$. This means $N = 30e^{-5}+70 = 70.2$ (to 1 decimal place).

The population reaches 75 million when

$$75 = 30e^{-0.5t} + 70.$$

That is

$$e^{-0.5t} = \frac{5}{30},$$

or

$$e^{0.5t} = \frac{30}{5} = 6$$

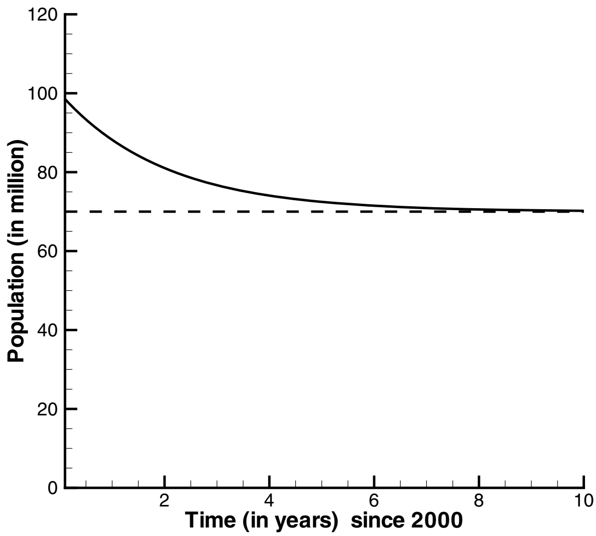

Figure 11.1 The graph of $N(t) = 30e^{-0.5t} + 70$. The dashed line corresponds to $N = 70$.

(since $e^{-x} = \frac{1}{e^x}$). Therefore $0.5t = \ln 6$ and so

$$t = \frac{\ln 6}{0.5} = 3.6 \text{ (to 1 decimal place).}$$

The population is therefore 75 million in 2004.

In the long run, the term $e^{-0.5t}$ approaches 0 and so N will approach the value 70 million (see Fig. 11.1).

11.3 Definite Integrals

The integral of a function of a variable x, as discussed previously, is itself a function of x. More precisely, it is called an **indefinite integral**. A **definite integral** has, as the name might suggest, a definite numerical value.

To be more specific, let $F(x) = \int f(x)dx$. Then the function $F(x)$ is the indefinite integral of $f(x)$ with respect to x. The definite integral of $f(x)$ between $x = a$ and $x = b$ is denoted by $\int_a^b f(x)dx$ and defined to be the number

$F(b) - F(a)$. The numbers a and b are the **limits of the integration**. For short, we can write $[F(x)]_a^b$ for $F(b) - F(a)$.

Notice that although $F(x)$ can have an arbitrary constant added to it, the constant will disappear when evaluating $F(b) - F(a)$. To sum up

The definite integral

$$\int_a^b f(x)dx = [F(x)]_a^b = F(b) - F(a),$$

where $F(x) = \int F(x)dx$ is the indefinite integral.

Thus to evaluate a definite integral, one needs first to find the indefinite integral, evaluate it at the two limits of integration, and then take the difference of these values.

Example 11.9

Evaluate the following

1. $\int_0^2 x^2 dx$,
2. $\int_0^1 e^{4x} dx$,
3. $\int_1^3 \frac{1}{x} dx$,
4. $\int_1^2 (6x^2 - 3x + 5)dx$.

Solution.

1. $\int_0^2 x^2 dx = \left[\frac{1}{3}x^3\right]_0^2$ (since $\frac{1}{3}x^3 = \int x^2 dx$).

   Therefore $\int_0^2 x^2 dx = \left[\frac{1}{3}x^3\right]_0^2 = \frac{1}{3}(8 - 0) = \frac{8}{3}$.
2. $\int_0^1 e^{4x} dx = \left[\frac{1}{4}e^{4x}\right]_0^1 = \frac{1}{4}(e^4 - e^0) = \frac{1}{4}(e^4 - 1) = 13.400$ (correct to 3 decimal places).
3. $\int_1^3 \frac{1}{x} dx$ (sometimes written as $\int_1^3 \frac{dx}{x}$) $= [\ln x]_1^3 = \ln 3 - \ln 1 = \ln 3 = 1.099$ (correct to 3 decimal places).
4. $\int_1^2 (6x^2 - 3x + 5)dx = \left[2x^3 - \frac{3}{2}x^2 + 5x\right]_1^2 = (16 - 6 + 10) - (2 - \frac{3}{2} + 5) = 14.5$.

11.4 Definite Integration: Area and Summation

The integral sign $\int$ originated in the 17th century as an elongated S, suggestive of the aspect of integration as a summation process, leading to a method for calculating areas under graphs.

Given a function $y = f(x)$, let A, B be two points on the x-axis with coordinates $(a, 0)$, $(b, 0)$ respectively, where $a \le b$ (see Fig. 11.2). We refer to the area enclosed by the graph of $y = f(x)$, the x-axis and the vertical lines $x = a$ and $x = b$, simply as the area under the graph between A and B (or between $x = a$ and $x = b$).

(For the following discussions, we assume that the function $f(x)$ is continuous between $x = a$ and $x = b$, which essentially means that the graph of $y = f(x)$ for this range of x is continuous, that is has no breaks.)

It can be shown that the area under the graph of $y = f(x)$ between $x = a$ and $x = b$ is the value of the definite integral $\int_a^b f(x)dx$. Here is a rough outline why this is so. Let P with coordinates $(x, 0)$ be a general point on the x-axis between A and B. Let $S(x)$ be the area under the graph of $y = f(x)$ between A and P.

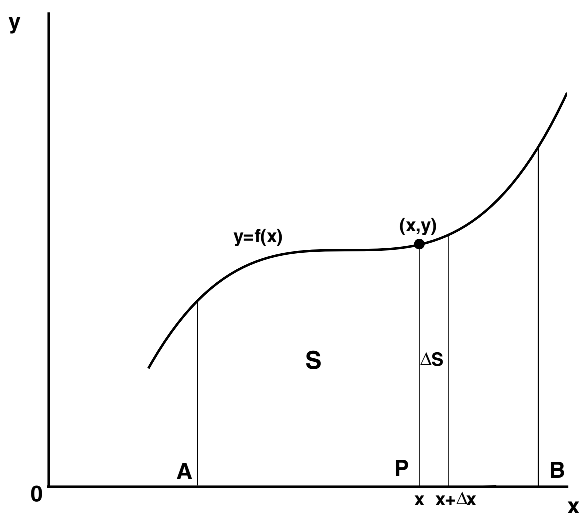

Figure 11.2 An illustration of the definite integral as the area under a graph.

A very small vertical strip of width Δx of this area can be regarded as rectangular of width Δx and height $y\,(= f(x))$. If ΔS denotes the area of this strip, then $\Delta S = y\Delta x$ so that $\frac{\Delta S}{\Delta x} = y$. As Δx gets smaller and smaller (approaches 0), the ratio $\frac{\Delta S}{\Delta x}$ approaches, by definition, the derivative $\frac{dS}{dx}$. Therefore $\frac{dS}{dx} = y = f(x)$. So $S(x) = F(x) + k$, where $F(x)$ is the (indefinite) integral $\int f(x)dx$ and k is some constant.

Obviously $S(a) = 0$, since $S(a)$ is just the area under the graph between A and A (the case $P = A$). Therefore $0 = S(a) = F(a) + k$ and so $k = -F(a)$. Therefore $S(x) = F(x) + k = F(x) - F(a)$.

In particular the area under the graph between $x = a$ and $x = b$ (the case $P = B$) is $S(b) = F(b) - F(a)$, which, by definition, is $\int_a^b f(x)dx$, since $F(x)$ is the indefinite integral $\int f(x)dx$. To sum up:

The area under the graph of $y = f(x)$ between $x = a$ and $x = b$ is the definite integral

$$\int_a^b f(x)dx.$$

Here are some rules for definite integrals. They follow easily from the definition.

1. $\displaystyle\int_a^b f(x)dx = -\int_b^a f(x)dx;$
2. $\displaystyle\int_a^a f(x)dx = 0;$
3. $\displaystyle\int_a^b f(x)dx + \int_b^c f(x)dx = \int_a^c f(x)dx.$

To see why these rules hold, let the indefinite integral $\int f(x)dx = F(x)$. Then, by definition, $\int_a^b f(x)dx = F(b) - F(a) = -(F(a) - F(b)) = -\int_b^a f(x)dx$. This proves Rule 1. Since $F(a) - F(a) = 0$, Rule 2 follows immediately. Finally, Rule 3 follows easily from the fact that $F(b) - F(a) + F(c) - F(b) = F(c) - F(a)$.

A brief word of warning. The mathematics involved treats areas below the x-axis as negative (since y is negative there). Therefore, more precisely, $\int_a^b f(x)dx$ is the difference of the total area above the x-axis and that under the x-axis, between $x = a$ and $x = b$.

Example 11.10

Find the area under the graph of $y = x^2 + 5$:

1. between $x = 0$ and $x = 1$;

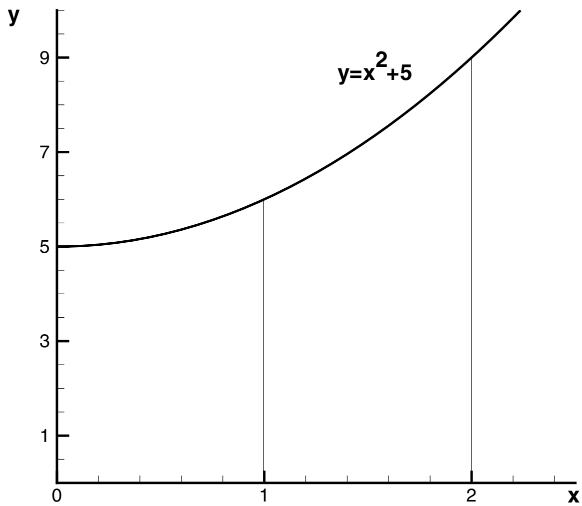

Figure 11.3 The graph of $y = x^2 + 5$.

2. between $x = 1$ and $x = 2$;
3. between $x = 0$ and $x = 2$.

(See Fig. 11.3.)

Solution.

1. $\int_0^1 (x^2+5)dx = \left[\frac{1}{3}x^3 + 5x\right]_0^1 = (\frac{1}{3}+5) - 0 = 5\frac{1}{3}$;
2. $\int_1^2 (x^2+5)dx = \left[\frac{1}{3}x^3 + 5x\right]_1^2 = (\frac{8}{3}+10) - (\frac{1}{3}+5) = 7\frac{1}{3}$;
3. $\int_0^2 (x^2+5)dx = \left[\frac{1}{3}x^3 + 5x\right]_0^2 = (\frac{8}{3}+10) - 0 = 12\frac{2}{3}$;

This illustrates Rule 3 for definite integrals since the value of the integral in Example 11.10.3 is the sum of the values of those in Example 11.10.1 and 11.10.2.

Example 11.11

Find the area in the first quadrant enclosed by the graph of $y = 4x^2$ and the y-axis and the line $y = 1$ (the shaded area in Fig 11.4).

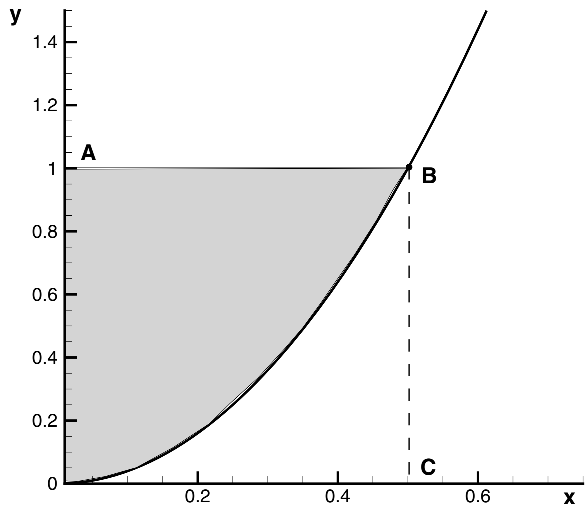

Figure 11.4 The graph of $y = 4x^2$.

Solution. Note that when $y = 1$ then $x^2 = \frac{1}{4}$ and so $x = \frac{1}{2}$ in the first quadrant. We want the area of the rectangle $ABCO$ (see Fig. 11.4) less the area under the graph between $x = 0$ and $x = \frac{1}{2}$. That is

$$\begin{aligned} 1 \times \frac{1}{2} - \int_0^{\frac{1}{2}} 4x^2 dx &= \frac{1}{2} - 4\int_0^{\frac{1}{2}} x^2 dx = \frac{1}{2} - 4\left[\frac{1}{3}x^3\right]_0^{\frac{1}{2}} \\ &= \frac{1}{2} - 4\left(\frac{1}{24} - 0\right) = \frac{1}{2} - \frac{1}{6} = \frac{1}{3}. \end{aligned}$$

Example 11.12

Find the area under the graph of $y = e^{2x}$ between $x = 0$ and $x = 1$ (the shaded area in Fig. 11.5).

Solution.

$$\begin{aligned} \int_0^1 e^{2x} dx &= \left[\frac{1}{2}e^{2x}\right]_0^1 = \frac{1}{2}(e^2 - e^0) \\ &= \frac{1}{2}(e^2 - 1) = 3.19 \qquad \text{(correct to 2 decimal places)} \end{aligned}$$

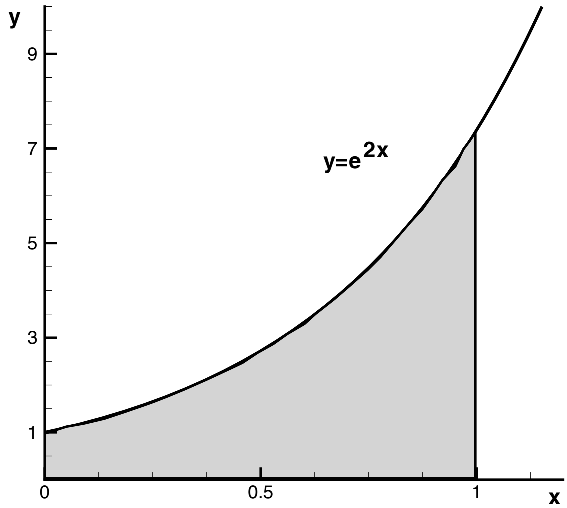

Figure 11.5 The graph of $y = e^{2x}$.

Example 11.13

Find the area under the graph of $y = \frac{1}{x}$ between $x = 1$ and $x = 2$.

Solution. Since $\int \frac{1}{x} dx = \ln x$, then $\int_1^2 \frac{1}{x} dx = [\ln x]_1^2 = \ln 2 - \ln 1 = \ln 2 - 0 = \ln 2 = 0.69$ (to 2 decimal places).

Example 11.14

Find the area that is enclosed completely between the graphs of $y = x^2$ and $y = 8x - x^2$ (see Fig 11.6).

Solution. The graphs meet at (x, y) where $x^2 = y = 8x - x^2$. That is, $2x^2 - 8x = 0$, or $2x(x - 4) = 0$. The only solutions to this equation are $x = 0$ and $x = 4$. So the graphs meet at the origin and at the point (4, 16). The

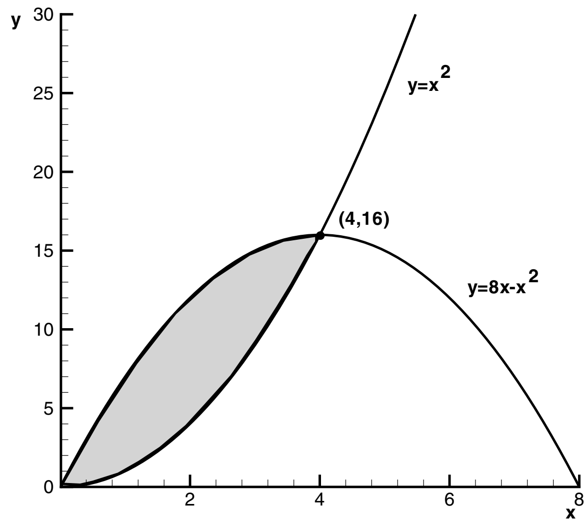

Figure 11.6 The graphs of the functions $y = x^2$ and $y = 8x - x^2$.

required area is

$$\begin{aligned}\int_0^4 (8x - x^2)dx - \int_0^4 x^2 dx &= \int_0^4 (8x - 2x^2)dx = 2\int_0^4 (4x - x^2)dx \\ &= 2\left[2x^2 - \frac{1}{3}x^3\right]_0^4 = 2\left(32 - \frac{64}{3} - 0\right) \\ &= \frac{64}{3} = 21\frac{1}{3}.\end{aligned}$$

The summation aspect of definite integration can be seen from the sketched proof that $\int_a^b f(x)dx$ is the area under the graph of $y = f(x)$ between $x = a$ and $x = b$.

In Fig 11.2, each small strip has area $y\Delta x$. Integrating gives the sum of all these small areas in the limiting case as strip widths Δx tend to 0. Put another way, if for each value of x between a and b we evaluate $f(x)$ and multiply it by a small increment of x and sum, the total is the definite integral $\int_a^b f(x)dx$ in the limit as the increments approach 0.

(If we abuse notation and regard dx as an infinitesimal increment of x, then $\int_a^b f(x)dx$ 'sums' all the products $f(x)dx$ as x ranges between a and b. We have touched here on the subject of differentials that is discussed in Appendix A.

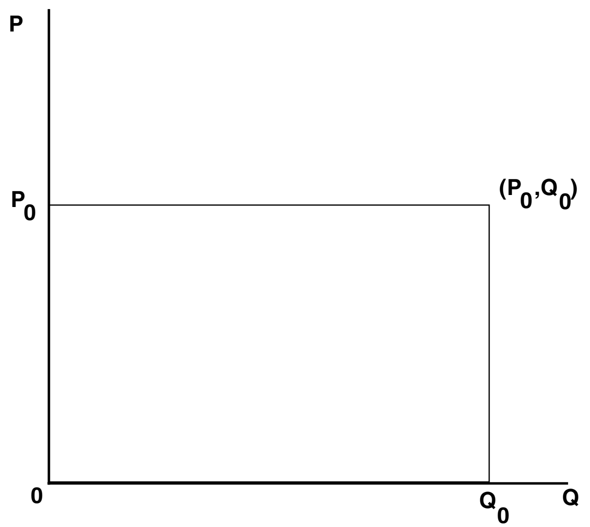

Figure 11.7 The graph of the supply curve $P = P_0$, where P_0 is constant. The area of the rectangle provides the revenue.

To illustrate this way of thinking, consider a good where the price P per unit charged by the supplier is a function of Q, the quantity supplied. If P is constant, say $P = P_0$, the revenue from supplying Q_0 units of the good is just P_0Q_0. This is the area of the rectangle in Fig. 11.7 or equivalently the area under the graph of $P = P_0$ between $Q = 0$ and $Q = Q_0$.

Consider the more general situation when the price P varies with Q, the supply curve. If a quantity Q_0 is supplied at the prevailing price P_0, the revenue is P_0Q_0. However, in practice, the quantity Q_0 may be supplied in smaller quantities totalling Q_0. As each small quantity ΔQ is supplied, the price P per unit changes. If ΔQ is very small we can assume the price per unit does not change as Q increases to $Q + \Delta Q$, so that the revenue for supplying the quantity ΔQ is $P \times \Delta Q$ (the shaded area in Fig. 11.8).

The revenue for supplying a total quantity Q_0 in small quantities ΔQ is the sum of all these small areas which, as ΔQ gets smaller and smaller (i.e., tends to 0), approaches in value the area under the supply curve between $Q = 0$ and $Q = Q_0$. This area is

$$\int_0^{Q_0} PdQ.$$

This number can be regarded as, theoretically, the total revenue obtained

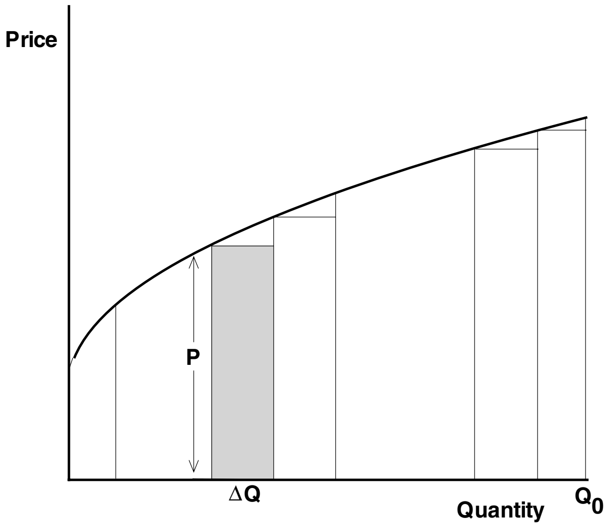

Figure 11.8 Graph illustrating the supplier's revenue as the area under the supply curve.

for supplying Q_0 units of the good in a continuous supply flow. More generally,

$$\int_A^B PdQ,$$

is the revenue produced as the quantity supplied, in this way, increases from A to B.

11.5 Producer's Surplus

This is a measure of the producer's satisfaction or of willingness to supply a good. We assume the price P is a function of quantity Q.

If the prevailing price for a quantity Q_0 of the good is P_0, then at that price the producer's revenue is P_0Q_0, the area of the rectangle OP_0AQ_0 in Fig. 11.9. However, the producer was willing to supply the Q_0 units of the good at lower prices in smaller quantities. The area under the supply curve $\int_0^{Q_0} PdQ$ between $Q = 0$ and $Q = Q_0$ represents the producer's revenue for supplying Q_0 units in a continuous supply flow.

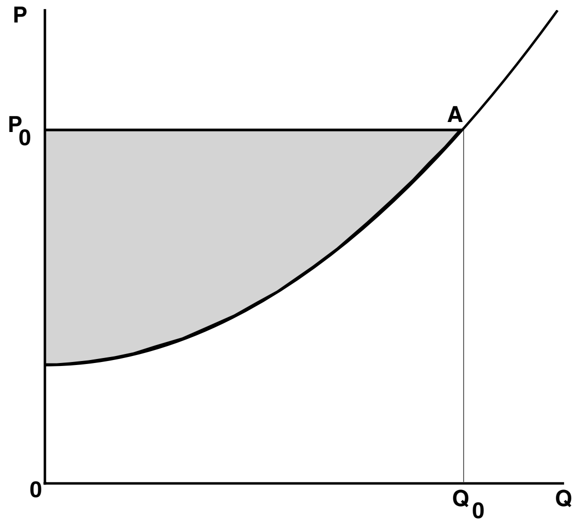

Figure 11.9 Graph showing the producer's surplus.

The difference between these two revenues, represented by the shaded area in Fig. 11.9, is

$$P_0Q_0 - \int_0^{Q_0} P dQ$$

and is known as the **producer's surplus**. It measures the benefit to the producer of supplying all Q_0 units at the prevailing price P_0.

11.6 Consumer's Surplus

This is a measure of consumer utility: benefit, satisfaction, or willingness to buy a particular good at the prevailing price. Again we assume that the price P per unit of the good is a function of the demand Q.

The following account is analogous to that for the producer's surplus. If the prevailing price for a quantity Q_0 is P_0, the consumer would pay P_0Q_0 for the goods, which is the area of the rectangle OP_0BQ_0 in Fig. 11.10.

However, if the consumer were to buy the same quantity Q_0 in a continuous purchase flow, the cost would be $\int_0^{Q_0} P dQ$, the area under the demand graph between $Q = 0$ and $Q = Q_0$.

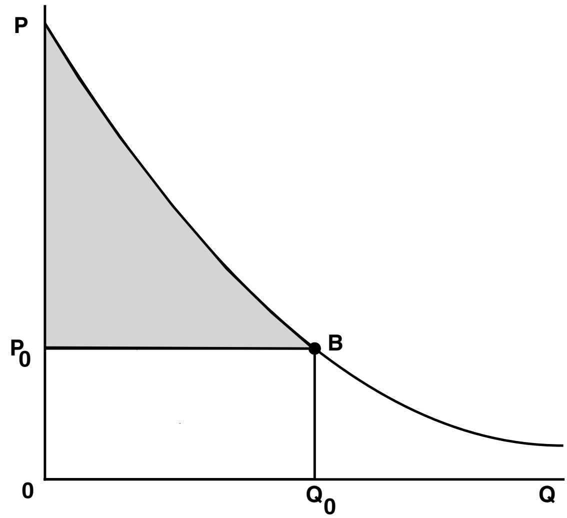

Figure 11.10 Graph showing the consumer's surplus.

Therefore by buying at the prevailing price, the consumer benefits by the difference of these two costs

$$\int_0^{Q_0} PdQ - P_0Q_0$$

which is known as the **consumer's surplus**. It is represented by the shaded area in Fig 11.10.

Example 11.15

Find the consumer's surplus if the demand function is $P = 17 - 5Q$, when the demand Q is 2.

Solution. Here $Q_0 = 2$, so $P_0 = 17 - 5Q_0 = 17 - 10 = 7$. Therefore $P_0Q_0 = 14$ and

$$\int_0^2 PdQ = \int_0^2 (17 - 5Q)dQ = \left[17Q - \frac{5}{2}Q^2\right]_0^2 = (34 - 10) - 0 = 24.$$

Therefore the consumer's surplus is $\int_0^2 PdQ - P_0Q_0 = 24 - 14 = 10$.

Example 11.16

If the demand equation is

$$P = \frac{8}{\sqrt{Q}},$$

find the consumer's surplus when $P = 4$.

Solution. We have $P_0 = 4$, and Q_0 is found from the equation $4 = P_0 = \frac{8}{\sqrt{Q_0}}$. Therefore $\sqrt{Q_0} = \frac{8}{4} = 2$, and so $Q_0 = 4$, $P_0Q_0 = 4 \times 4 = 16$, and

$$\begin{aligned}
\int_0^4 PdQ &= \int_0^4 \frac{8}{\sqrt{Q}}dQ \\
&= 8\int_0^4 Q^{-\frac{1}{2}}dQ \\
&= 8\left[\frac{1}{-\frac{1}{2}+1}Q^{-\frac{1}{2}+1}\right]_0^4 \\
&= 8\left[2Q^{\frac{1}{2}}\right]_0^4 \\
&= 8[2 \times 2 - 0] \\
&= 32.
\end{aligned}$$

Therefore, the consumer's surplus is

$$\int_0^4 PdQ - P_0Q_0 = 32 - 16 = 16.$$

Example 11.17

Find the producer's surplus at $Q = 5$ if the supply function is $P = 7 + 4Q$.

Solution. Since $Q_0 = 5$, then $P_0 = 7 + 4 \times 5 = 27$, so $P_0Q_0 = 135$ and

$$\int_0^5 PdQ = \int_0^5 (7 + 4Q)dQ = [7Q + 2Q^2]_0^5 = 35 + 50 - 0 = 85.$$

The producer's surplus is

$$P_0Q_0 - \int_0^{Q_0} PdQ = 135 - 85 = 50.$$

Figure 11.11 illustrates this example. The shaded area represents the producer's surplus.

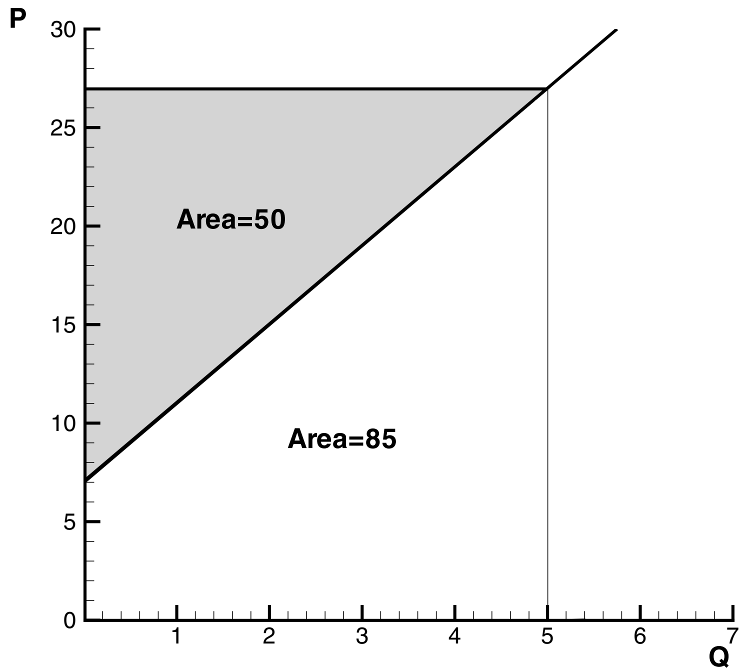

Figure 11.11 Producer's surplus for Example 11.17.

Example 11.18

Given the demand function $P = 70 - 4Q_d$ and the supply function $P = 5 + Q_s$, evaluate the consumer's surplus and the producer's surplus, assuming equilibrium.

Solution. In equilibrium, $Q_d = Q_s$. Let Q be this common value. Then the equilibrium price is given by

$$70 - 4Q = P = 5 + Q.$$

Therefore $70 - 5 = Q + 4Q$ and so $65 = 5Q$ and $Q = 13$. The equilibrium price is then $P = 5 + Q = 18$.

To calculate the consumer's surplus, we use the consumer's demand function in equilibrium: $P = 70 - 4Q$. The consumer's surplus is therefore:

$$\begin{aligned}\int_0^{13} PdQ - 13 \times 18 &= \int_0^{13} (70 - 4Q)dQ - 234 \\ &= \left[70Q - 2Q^2\right]_0^{13} - 234 \\ &= (910 - 338 - 0) - 234 = 338.\end{aligned}$$

For the producer's surplus, use the producer's supply function in equilibrium: $P = 5 + Q$. The producer's surplus is therefore:

$$\begin{aligned}13 \times 18 - \int_0^{13} PdQ &= 234 - \int_0^{13} (5 + Q)dQ \\ &= 234 - \left[5Q + \frac{1}{2}Q^2\right]_0^{13} \\ &= 234 - (65 + 84.5 - 0) = 84.5\end{aligned}$$

Example 11.19

Given the demand function $P = 25 - Q - 0.3Q^2$, by how much does the consumer's surplus change if Q increases from 5 to 6 units?

Solution. If $Q = 5$, then $P = 12.5$ and the consumer's surplus is

$$\begin{aligned}\int_0^{5} PdQ - 5 \times 12.5 &= \int_0^{5} (25 - Q - 0.3Q^2)dQ - 62.5 \\ &= \left[25Q - \frac{1}{2}Q^2 - 0.1Q^3\right]_0^{5} - 62.5 \\ &= 125 - 12.5 - 12.5 - 62.5 \\ &= 37.5\end{aligned}$$

When $Q = 6$, then $P = 8.2$ and the consumer's surplus is

$$\begin{aligned}\int_0^{6} PdQ - 6 \times 8.2 &= \int_0^{6} (25 - Q - 0.3Q^2)dQ - 49.2 \\ &= \left[25Q - \frac{1}{2}Q^2 - 0.1Q^3\right]_0^{6} - 49.2 \\ &= 150 - 18 - 21.6 - 49.2 \\ &= 61.2\end{aligned}$$

The consumer's surplus therefore increases by 23.7 units.

Integration may be used to reconstruct a function from its corresponding marginal function. Put another way, given $f'(x)$, the rate of change relative to x of a function $f(x)$ of x, then integration can be used to determine $f(x)$.

For instance, given a marginal cost function $MC = \frac{d}{dQ}(TC)$, the total cost function TC is given by $TC = \int MCdQ$. This indefinite integral is in general a function of Q.

The extra costs as a result of raising production from $Q = A$ to $Q = B$ units is the difference of the values of this indefinite integral evaluated at $Q = A$ and $Q = B$. From the definition of a definite integral, this difference is $\int_A^B MCdQ$.

This is a special case of the more general equation

$$\int_a^b f'(x)dx = f(b) - f(a) \tag{11.1}$$

which is obvious because the integral of the derivative $f'(x)$ of $f(x)$ is $f(x)$ (since differentiation and integration are inverse operations). Then (11.1) follows by definition of the definite integral.

Example 11.20

A company's marginal cost function is given by $MC = 100 - 2Q + 0.6Q^2$. Calculate the extra cost in increasing production from:

1. 5 to 10 units,
2. 10 to 15 units.

Solution.

1.

$$\begin{aligned}
\int_5^{10} MCdQ &= \int_5^{10} (100 - 2Q + 0.6Q^2)dQ \\
&= \left[100Q - Q^2 + 0.6 \times \frac{1}{3}Q^3\right]_5^{10} \\
&= 1{,}000 - 100 + 200 - (500 - 25 + 25) \\
&= 600
\end{aligned}$$

2.

$$\begin{aligned}\int_{10}^{15} MCdQ &= \int_{5}^{10} (100 - 2Q + 0.6Q^2)dQ \\ &= \left[100Q - Q^2 + 0.2Q^3\right]_{10}^{15} \\ &= 1{,}500 - 225 + 675 - (1{,}000 - 100 + 200) \\ &= 850\end{aligned}$$

Notes

1. The cost of raising production from 5 to 15 units is $\int_5^{15} MCdQ = \int_5^{10} MCdQ + \int_{10}^{15} MCdQ = 600 + 850 = 1{,}450$. Here we used property 3 of definite integrals in §11.4.
2. We have implicitly found the total cost function TC. For as mentioned before, TC is the indefinite integral $\int MCdQ$. Therefore

$$TC = 100Q - Q^2 + 0.2Q^3 + k$$

where k is some constant.

Clearly when $Q = 0$, total costs $TC =$ fixed costs FC. Therefore $k = FC$. So we can determine TC exactly if the fixed costs are given.

The total revenue TR of a company may be regarded as a function of time t. If we know the revenue the company receives each day, the total revenue over t days is simply the sum of the revenues for each of these days. In this case, we implicitly assume that t is a **discrete** variable.

If the company's revenue is in continuous flow and we are given the marginal revenue MR as a function of t, then to calculate the total revenue we use the continuous analogue of discrete summation: integration.

Example 11.21

Find TR when output $Q = 4$ for each of the following MR functions:

1. $MR = 15 - 0.6Q$,
2. $MR = 40Q^{-0.5}$.

In each case, compute the increase in TR as Q is raised from 4 to 9 units.

Solution.

1. When $Q = 4$,

$$TR = \int_0^4 MR\, dQ = \int_0^4 (15 - 0.6Q)dQ$$
$$= \left[15Q - 0.3Q^2\right]_0^4 = 60 - 4.8 = 55.2.$$

The change in TR as Q increases from 4 to 9 is

$$\int_4^9 MR\, dQ = \left[15Q - 0.3Q^2\right]_4^9$$
$$= (135 - 24.3) - 55.2 = 55.5$$

2. When $Q = 4$,

$$TR = \int_0^4 40Q^{-0.5} dQ = 40 \int_0^4 Q^{-0.5} dQ$$
$$= 40 \left[\frac{1}{-0.5+1} Q^{-0.5+1}\right]_0^4$$
$$= 80 \left[Q^{0.5}\right]_0^4$$
$$= 80(4^{0.5} - 0) = 80 \times 2 = 160.$$

The change in TR as Q increases from 4 to 9 is

$$\int_4^9 MRdQ = 80 \left[Q^{0.5}\right]_4^9 = 80(9^{0.5} - 4^{0.5})$$
$$= 80(3 - 2) = 80.$$

EXERCISES

11.1. Integrate each of the following functions with respect to x:

a) x^7,

b) x^{-4},

c) $3x^2 + 2x + 1$,

d) $16x^4 - \dfrac{2}{x^2}$,

e) e^{-4x},

f) $e^{2x} + 4x - 7$,

g) $\dfrac{(x^2 + 4x + 2)}{2x}$,

h) $x(x + 2)^2$,

i) $x^3 + x - 3$.

11.2. Evaluate:

a) $\displaystyle\int_0^1 x^4 dx$,

b) $\displaystyle\int_2^4 (3x + 2)dx$,

c) $\displaystyle\int_0^2 e^{-3x} dx$,

d) $\displaystyle\int_0^1 (3x^4 + 2x^3 - x + 3)dx$,

e) $\displaystyle\int_1^2 \left(3x + \frac{2}{x}\right) dx$.

11.3. Find the total revenue TR and demand functions corresponding to each of the marginal revenue functions:

a) $MR = 15 - 4Q$,

b) $MR = \dfrac{9}{\sqrt{Q}}$.

11.4. Determine the total cost function TC if the marginal cost function $MC = 25 + 8Q$ and fixed costs are 12 units.

11.5. If the marginal propensity to consume

$$MPC = 0.75 + \frac{0.1}{\sqrt{Y}}$$

and consumption is 15 when income is 16, determine the consumption function and hence the corresponding savings function.

11.6. Evaluate the consumer's surplus CS at $Q = 1$ for the demand function $P = 5 - Q - 2Q^2$.

11.7. Evaluate the producer's surplus PS at $Q = 3$ for the supply function $P = 40 + 3Q^2$. Find the change in PS if Q increases to 4.

11.8. Net investment $I(t)$ is defined to be the rate of change of capital stock $K(t)$ relative to time t.

If $I(t) = 240t^{\frac{3}{5}}$ and the initial stock of capital is 100, determine the function $K(t)$.

12
Linear Difference Equations

12.1 Introduction

Problems encountered so far have mostly been **static** in that the quantities and equations involved are for a particular period of time. For instance, the current price of a good depends on the current demand of consumers.

However, it may be that this year's price for a good, such as a car, depends on last year's demand; or a manufacturing quota for a particular month depends on the demand in that month in the previous year (or years). This is the concept of a lagged response. These examples are not static but **dynamic** situations in which economic models are viewed as a sequence of discrete periods. The value of an economic quantity in one period may depend on data from the previous period, or previous n periods for some integer $n > 0$.

Difference equations are used to analyse dynamic models. An n^{th} order difference equation, for instance, might express the price of a good as a function of the demands in the previous n years.

12.2 Difference Equations

Consider a **sequence** X_0, X_1, $\ldots$ of quantities, which we denote simply by $\{X_t\}$. By a sequence we mean a list in a specific order. In the economic situations that will concern us, t denotes time measured in discrete units 0, 1, 2, $\ldots$.

In this case, the period before time $t+1$ starting at time t is referred to as period t.

Thus, period 0 is the initial period. The sequence $\{X_t\}$ can be regarded as the values of a 'step' function X of a continuous time variable i, where $X(i) = X_t$ for all i in period t. Thus, X has constant value X_t in period t (see, for example, Fig. 12.1).

One way of generating a sequence $\{X_t\}$ is by using an n^{th} order **difference equation**, which relates the general term of the sequence to the previous n terms. We shall only be concerned with the **linear difference equations** (LDEs). They are of the form:

$$X_t + a_{t-1}X_{t-1} + a_{t-2}X_{t-2} + \ldots + a_{t-n}X_{t-n} = b,$$

where the coefficients a_i and b are constants (independent of t). If $b = 0$, the difference equation is said to be **homogeneous**; otherwise it is **inhomogeneous**.

By 'solving' a difference equation, we mean expressing the function X mentioned earlier explicitly as a function of t. An n^{th} order difference equation determines the sequence uniquely if the first n terms of the sequence are specified.

For example, consider the first order difference equation

$$X_t - 3X_{t-1} = -1,$$

(or $X_t = 3X_{t-1} - 1$) where $X_0 = 2$. Using the difference equation, we can generate successively the terms of the corresponding sequence $\{X_t\}$:

$$X_1 = 3X_0 - 1 = 3 \times 2 - 1 = 5,$$
$$X_2 = 3X_1 - 1 = 3 \times 5 - 1 = 14,$$

and so on. To compute the term X_{100} in this way would mean explicitly evaluating $X_1, X_2, \ldots, X_{99}$ successively. This is laborious, but if we solve the linear difference equation, the computation is simple. Accept for the moment (for reasons that will be explained later) that the solution of this difference equation is

$$X_t = \frac{1}{2}(1 + 3^{t+1}).$$

Substituting $t = 0$ gives $X_0 = \frac{1}{2}(1 + 3^1) = 2$ (as it should be); $t = 1$ gives $X_1 = \frac{1}{2}(1 + 3^2) = 5$, and $t = 2$ gives $X_2 = \frac{1}{2}(1 + 3^3) = 14$, agreeing with our previous computations. The term X_{100} is simply $\frac{1}{2}(1 + 3^{101})$, which is a huge number and best left expressed in this form.

Difference equations are sometimes known as time series (if t denotes time) or recurrence relations. The sequence $\{X_t\}$ is also known as the **time path** of the function X being measured, giving the successive values in time of X.

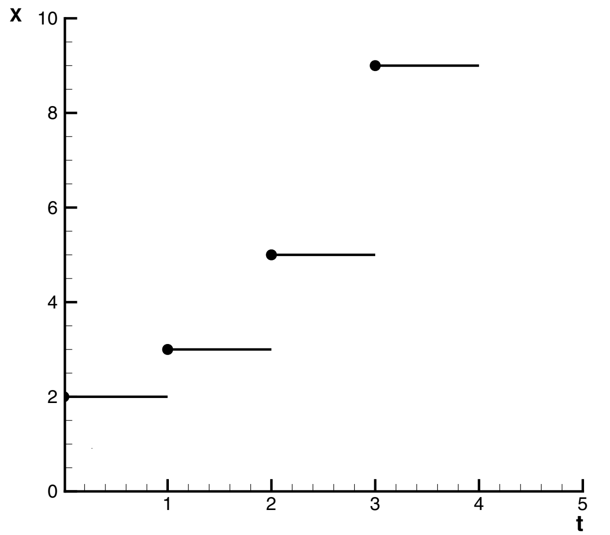

Figure 12.1 Graph representing the solution of the difference equation $X_t = 2X_{t-1} - 1$ with $X_0 = 2$.

A sequence $\{X_t\}$ may be visualised by plotting the graph of the corresponding step function X. For instance, the sequence $\{X_t\}$ given by the linear difference equation

$$X_t = 2X_{t-1} - 1,$$

where $X_0 = 2$, has first four terms: 2, 3, 5, 9. This can be represented graphically as in Fig. 12.1. (On a line segment, the heavy dot represents the end point that belongs to the graph.)

The values of X change only at discrete values of t. So for instance,

$$\begin{aligned} X_t &= X_0 = 2 \quad &\text{for} \quad &0 \leq t < 1, \\ X_t &= X_1 = 3 \quad &\text{for} \quad &1 \leq t < 2, \\ X_t &= X_2 = 5 \quad &\text{for} \quad &2 \leq t < 3, \\ X_t &= X_3 = 9 \quad &\text{for} \quad &3 \leq t < 4, \end{aligned}$$

and so on.

12.3 First Order Linear Difference Equations

First order linear difference equations can be dealt with as a special case of second order linear difference equations. However, it is instructive to see them analysed from basics.

Consider a general first order linear difference equation written in the form

$$X_t = aX_{t-1} + b. \tag{12.1}$$

This means that each term is a times the previous term and then b is added to this product. Therefore, $X_{t-1} = aX_{t-2} + b$, $X_{t-2} = aX_{t-3} + b$, and so on. It follows that

$$\begin{aligned} X_t &= a(aX_{t-2} + b) + b \\ &= a^2 X_{t-2} + (a+1)b \\ &= a^2(aX_{t-3} + b) + (a+1)b \\ &= a^3 X_{t-3} + (a^2 + a + 1)b. \end{aligned}$$

Eventually, we have

$$X_t = a^t X_{t-t} + (a^{t-1} + a^{t-2} + \ldots + a + 1)b.$$

The term

$$a^{t-1} + a^{t-2} + \ldots + a + 1$$

is the sum of a geometric series and is equal to

$$\frac{(1-a^t)}{(1-a)} \text{ if } a \neq 1;$$

otherwise the sum is equal to

$$1 + 1 + \ldots + 1 + 1(t \text{ terms }) = t.$$

Therefore,

$$X_t = \begin{cases} a^t X_0 + \left(\dfrac{1-a^t}{1-a}\right) b & \text{if } a \neq 1, \\ X_0 + bt & \text{if } a = 1. \end{cases}$$

If $a \neq 1$, we can rewrite the solution (collecting together the terms involving a^t) in the form:

$$X_t = Aa^t + \frac{b}{1-a}, \tag{12.2}$$

where A is a constant that can be determined from X_0. (The example given earlier in Section 12.2 had $a = 3$ and $b = -1$.)

Example 12.1

A bank customer borrows \$15,000. Interest is 9.6% per annum on the outstanding balance. The customer can afford to repay at most \$400 each month.

1. How long will it take to repay the loan?
2. How much does the customer owe after 1 year?

Solution.

1. Let X_t be the amount owed after t months. Then $X_0 = 15{,}000$. At the end of t months, interest of $\frac{9.6}{12}\% = 0.8\%$ of the current balance of X_{t-1} is added and a repayment made of \$400, therefore

$$\begin{aligned} X_t &= \left(1 + \frac{0.096}{12}\right) X_{t-1} - 400 \\ &= 1.008X_{t-1} - 400. \end{aligned}$$

From (12.2) we have (with $a = 1.008$ and $b = -400$):

$$X_t = A(1.008)^t - \frac{400}{1 - 1.008} = A(1.008)^t + 50{,}000,$$

where A is a constant.

Since $X_0 = 15{,}000$, then putting $t = 0$ in this equation gives (as $1.008^0 = 1$)

$$15{,}000 = A + 50{,}000.$$

Therefore $A = -35{,}000$ and so

$$X_t = -35{,}000(1.008)^t + 50{,}000.$$

This is explicitly the balance owing after t months. The balance is 0 at time t if

$$35{,}000(1.008)^t = 50{,}000,$$

that is

$$(1.008)^t = \frac{50{,}000}{35{,}000} = \frac{10}{7}.$$

Taking natural logarithms of both sides:

$$t \ln(1.008) = \ln\left(\frac{10}{7}\right).$$

Therefore

$$t = \frac{\ln\left(\frac{10}{7}\right)}{\ln(1.008)},$$

which is approximately 44.76 months. Therefore, the loan will be paid off at the end of the 45th month.

2. $X_{12} = -35{,}000(1.008)^{12} + 50{,}000 \approx \$11{,}488.15$.

Example 12.2

A bank savings account pays 5% per annum interest. Initially, a saver deposits £1,000. After 10 years, what will be the value of this customer's savings account if

1. no further deposits are made;
2. £100 is deposited at the end of each year?

Solution. Let X_t be the value of the savings account after t years.

1. In this case
$$X_t = (1 + 0.05)X_{t-1} = 1.05X_{t-1}.$$
Then from (12.2), with $a = 1.05$ and $b = 0$, we have
$$X_t = A(1.05)^t,$$
where A is a constant.
Since $X_0 = 1{,}000$, then $1{,}000 = X_0 = A(1.05)^0 = A$. Therefore,
$$X_t = 1{,}000(1.05)^t.$$
We want
$$X_{10} = 1{,}000(1.05)^{10} \approx £1{,}628.89.$$

2. The difference equation is now
$$X_t = (1.05)X_{t-1} + 100.$$
Then from (12.2), with $a = 1.05$ and $b = 100$, we have
$$X_t = A(1.05)^t + \frac{100}{1 - 1.05} = A(1.05)^t - 2{,}000.$$
Since $1{,}000 = X_0$, then
$$1{,}000 = X_0 = A(1.05)^0 - 2{,}000 = A - 2{,}000.$$
Consequently, $A = 3{,}000$ and
$$X_t = 3{,}000(1.05)^t - 2{,}000.$$
The value after 10 years is
$$X_{10} = 3{,}000(1.05)^{10} - 2{,}000 \approx £2{,}886.68.$$

12.4 Stability

Suppose an economic model has associated with it the first order linear difference equation

$$X_t = aX_{t-1} + b \qquad (a \neq 1).$$

We showed that this has solution:

$$X_t = Aa^t + \frac{b}{1-a},$$

where A is a constant (independent of time t).

1. If $-1 < a < 1$, then a^t tends to 0 as t tends to infinity (i.e., t increases indefinitely). Then X_t *converges* to the value $\frac{b}{1-a}$, called the **equilibrium value**.

   The convergence is **oscillatory** if a is negative and is **uniform** if a is positive. The model or difference equation in this case is said to be **stable**. See Figs. 12.2 and 12.3, respectively.

2. If $a < -1$ or $a > 1$, then X_t *diverges* in that, numerically, X_t increases without bound. If a is negative, the divergence is **oscillatory**, whereas if a is positive, the divergence is **uniform**. See Figs. 12.4 and 12.5, respectively. The model or difference equation is **unstable** in this case.

Example 12.3

Solve the following linear difference equations:

1. $X_t = -0.5X_{t-1} + 0.25$, where $X_0 = 0.5$,
2. $X_t = \frac{1}{3}X_{t-1} + 1$, where $X_0 = \frac{1}{2}$,
3. $X_t = -2X_{t-1} + 3$, where $X_0 = 3$,
4. $X_t = 3X_{t-1} + 5$, where $X_0 = 3.5$.

In each case, comment on stability and display the solution graphically for $0 \leq t < 5$.

Solution.

1.

$$X_t = A(-0.5)^t + \frac{0.25}{1-(-0.5)} = A(-0.5)^t + \frac{0.25}{1.5}$$
$$= A(-0.5)^t + \frac{1}{6}.$$

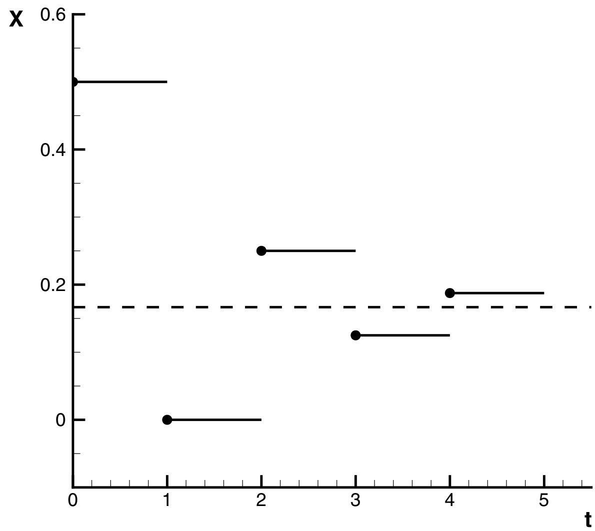

Figure 12.2 Graph of the solution of Example 12.3.1 showing oscillatory convergence to the equilibrium value.

Since $0.5 = X_0 = A + \frac{1}{6}$, then $A = \frac{1}{3}$. Therefore

$$X_t = \frac{1}{3}(-0.5)^t + \frac{1}{6}.$$

This is a stable linear difference equation, with oscillatory convergence to the equilibrium value $\frac{1}{6}$ as shown in Fig. 12.2.

2.

$$X_t = A\left(\frac{1}{3}\right)^t + \frac{1}{1-\frac{1}{3}} = A\left(\frac{1}{3}\right)^t + \frac{3}{2}.$$

Since $\frac{1}{2} = X_0 = A + \frac{3}{2}$, then $A = -1$. Therefore,

$$X_t = \frac{3}{2} - \left(\frac{1}{3}\right)^t.$$

This is a stable linear difference equation, uniformly converging to the equilibrium value $\frac{3}{2}$ as shown in Fig. 12.3.

3.

$$X_t = A(-2)^t + \frac{3}{1-(-2)} = A(-2)^t + 1.$$

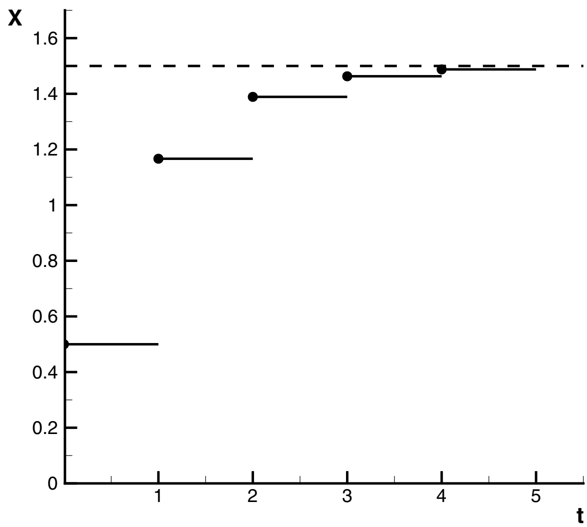

Figure 12.3 Graph of the solution of Example 12.3.2 showing uniform convergence to the equilibrium value.

Since $3 = X_0 = A + 1$, then $A = 2$ and

$$X_t = 2(-2)^t + 1.$$

This is an unstable linear difference equation. The divergence is oscillatory (see Fig. 12.4).

4.
$$X_t = A(3)^t + \frac{5}{1-3} = A(3)^t - 2.5.$$

Since $3.5 = X_0 = A - 2.5$, then $A = 6$. Therefore,

$$X_t = 6(3)^t - 2.5.$$

This is an unstable linear difference equation. The divergence is uniform (see Fig. 12.5).

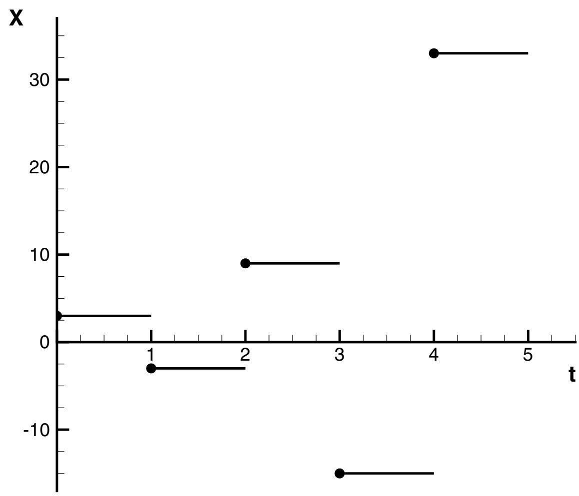

Figure 12.4 Graph of the solution of Example 12.3.4 showing oscillatory divergence.

12.5 The Cobweb Model

The Cobweb model is an economic model for analysing periodic fluctuations in price, supply, and demand that oscillate towards equilibrium. It is assumed that the quantities involved change only at discrete time intervals and that there is a time lag in the response of suppliers to price changes.

For instance, the supply this year of a particular agricultural product depends on the price obtained from the previous year's harvest. The demand for the produce will depend of course on this year's price. Another example is that of package holidays. The holiday company's supply of holidays for this season will depend on the prices obtained for last season's.

In general, we assume that the supply function at time t for a single good is

$$Q_{S,t} = aP_{t-1} + b.$$

Here $Q_{S,t}$ is the supply at time t and P_{t-1} the price at time $t-1$ (the previous period). The demand equation is

$$Q_{D,t} = cP_t + d.$$

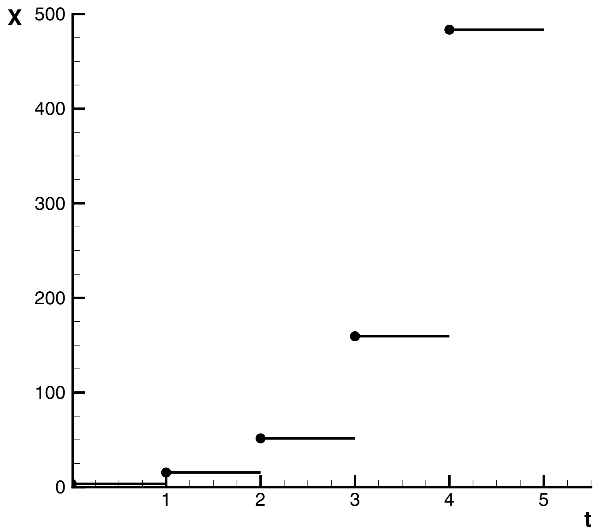

Figure 12.5 Graph of the solution of Example 12.3.3 showing uniform divergence.

Here a, b, c, d are constants, with $a > 0$ and $c < 0$. Initially $t = 0$, and then t increases one unit at a time.

Assuming equilibrium in period t, we have $Q_{D,t} = Q_{S,t}$. That is,

$$aP_{t-1} + b = cP_t + d$$

or

$$P_t = \frac{a}{c}P_{t-1} + \frac{b-d}{c}.$$

This is a first order linear difference equation, with $\frac{a}{c} < 0$, since a is positive and c is negative. The sequence $\{P_t\}$ generated gives the equilibrium prices for each period. Since $a/c \neq 0$, as a/c is negative, we can solve the difference equation using (12.2) to obtain an expression for P_t in the form

$$\begin{aligned} P_t &= A\left(\frac{a}{c}\right)^t + \left(\frac{b-d}{c}\right) / \left(1 - \frac{a}{c}\right) \\ &= A\left(\frac{a}{c}\right)^t + \frac{b-d}{c-a} \end{aligned}$$

where A is a constant.

If $-1 < \frac{a}{c} < 0$, we have stability and oscillatory convergence to an equilibrium value

$$P_e = \frac{b-d}{c-a}. \qquad (12.3)$$

Example 12.4

A car manufacturer's supply and demand functions at time t for a particular car model are

$$Q_{S,t} = 3P_{t-1} - 12, \qquad Q_{D,t} = -4P_t + 28.$$

Show that over time, the car's price will converge and give the equilibrium value.

Solution. For equilibrium in period t,

$$3P_{t-1} - 12 = -4P_t + 28,$$

which simplifies to

$$P_t = \left(-\frac{3}{4}\right) P_{t-1} + 10. \qquad (12.4)$$

Since $0 > -\frac{3}{4} > -1$, the prices P_t converge to the equilibrium value P_e, which we can compute using the formula obtained earlier (equation (12.3)) or in §12.3. Using the latter with $a = -\frac{3}{4}$ and $b = 10$, we have

$$P_e = \frac{b}{1-a} = \frac{10}{1-\left(-\frac{3}{4}\right)} = \frac{40}{7} = 5\frac{5}{7}.$$

Another way to compute P_e, when we know there is convergence, is to note that in the limit as t tends to infinity, $P_t = P_{t-1} = P_e$. Therefore,

$$P_e = \left(-\frac{3}{4}\right) P_e + 10.$$

Then $\left(1 + \frac{3}{4}\right) P_e = 10$, from which it follows that $P_e = 5\frac{5}{7}$, as before.

Observe that the equilibrium price in this problem is independent of the value of P_0. The general solution of linear difference equation (12.4) is

$$P_t = A\left(-\frac{3}{4}\right)^t + \frac{10}{1-\left(\frac{3}{4}\right)} = A\left(-\frac{3}{4}\right)^t + \frac{40}{7}$$

where A is a constant.

From this equation it is again clear that the equilibrium price is $\frac{40}{7}$ and that this does not depend on the value of A (which depends on the value of P_0).

12.6 Second Order Linear Difference Equations

A time path sequence is uniquely determined, given a second order linear difference equation and the values of the first two terms.

Example 12.5

Find the first five terms of the sequence $\{X_t\}$ given by

$$X_t - 5X_{t-1} + 3X_{t-2} = 1,$$

given that $X_0 = 1$, $X_1 = 3$.

Solution. Rearranging the linear difference equation gives

$$X_t = 5X_{t-1} - 3X_{t-2} + 1.$$

Therefore,

$$\begin{aligned}
X_2 &= 5X_1 - 3X_0 + 1 = 5 \times 3 - 3 \times 1 + 1 = 13, \\
X_3 &= 5X_2 - 3X_1 + 1 = 5 \times 13 - 3 \times 3 + 1 = 57, \\
X_4 &= 5X_3 - 3X_2 + 1 = 5 \times 57 - 3 \times 13 + 1 = 247.
\end{aligned}$$

Therefore, the first five terms in order are 1, 3, 13, 57, 247.

The general second order linear difference equation is of the form

$$X_t + aX_{t-1} + bX_{t-2} = c \tag{12.5}$$

where we assume a, b, c are constants (independent of t).

First order linear difference equations may be considered as the special case $b = 0$. This explains the remark made in Section 12.3 when discussing first order linear difference equations: that they are special cases of second order linear difference equations. However, we gave a self-contained account of the general solution for the first order case. The general solution for the second order linear difference equations is a little more involved.

The **associated** homogeneous linear difference equation of (12.5) is

$$X_t + aX_{t-1} + bX_{t-2} = 0. \tag{12.6}$$

If $\{U_t\}$, $\{V_t\}$ are sequences satisfying the linear difference equation (12.5), then

$$U_t + aU_{t-1} + bU_{t-2} = c$$

and

$$V_t + aV_{t-1} + bV_{t-2} = c.$$

Subtracting we have:

$$U_t - V_t + a(U_{t-1} - V_{t-1}) + b(U_{t-2} - V_{t-2}) = 0$$

or

$$W_t + aW_{t-1} + bW_{t-2} = 0$$

where $W_t = U_t - V_t$. Thus $\{W_t\}$ satisfies the homogeneous linear difference equation (12.6).

It follows that any two solutions of (12.5) differ by a solution of the associated linear difference equation (12.6). Thus if we manage to find a **particular solution** of (12.5) by guesswork or from theory, then the **general solution** of an inhomogeneous linear difference equation is obtained by adding the particular solution to the general solution of the associated homogeneous linear difference equation (known as the **complementary solution**). In brief:

$$\textbf{General Solution} = \textbf{Particular Solution} + \textbf{Complementary Solution}. \tag{12.7}$$

This is true generally for linear difference equations and, in particular, for first order ones. We saw in Section 12.3 that the general solution of a first order linear difference equation

$$X_t = aX_{t-1} + b$$

is of the form

$$X_t = Aa^t + \frac{b}{1-a}(\text{if } a \neq 1).$$

It can be verified that $X_t = Aa^t$ is the complementary solution (the general solution of the associated linear difference equation $X_t = aX_{t-1}$) and that the sequence $\{X_t\}$, where $X_t = \frac{b}{1-a}$ (for all t), is a particular solution of $X_t = aX_{t-1} + b$ since

$$\frac{b}{1-a} = a\left(\frac{b}{1-a}\right) + b.$$

12.6.1 Complementary Solutions

In analogy with the complementary solution for the homogeneous first order case, suppose we try a similar solution of (12.6), of the form $X_t = Au^t$, where A is a constant and u a number to be determined. Then

$$Au^t + aAu^{t-1} + bAu^{t-2} = 0.$$

Dividing throughout by Au^{t-2} gives

$$u^2 + au + b = 0.$$

So u is a root of the quadratic equation

$$x^2 + ax + b = 0, \tag{12.8}$$

known as the **characteristic equation** of the linear difference equation. Let its roots be u, v (sometimes called the **characteristic roots**).

It can be shown that the general solution of the homogeneous linear difference equation (12.6) is of the form

$$X_t = \begin{cases} Au^t + Bv^t & \text{if } u \neq v, \\ (A + tB)u^t & \text{if } u = v, \end{cases}$$

where A and B are constants.

Example 12.6

Solve the homogeneous linear difference equation

$$X_t - 7X_{t-1} + 10X_{t-2} = 0$$

where $X_0 = 2$, $X_1 = 13$. Determine X_{10}.

Solution. The characteristic equation is

$$x^2 - 7x + 10 = 0$$

whose roots are 2 and 5. The general solution is therefore

$$X_t = A2^t + B5^t$$

where A and B are constants.

Since $X_0 = 2$, then

$$2 = X_0 = A2^0 + B5^0 = A + B.$$

Similarly, since $X_1 = 13$, then

$$13 = X_1 = A2^1 + B5^1 = 2A + 5B.$$

Thus, we have the simultaneous equations

$$A + B = 2, \tag{12.9}$$

$$2A + 5B = 13. \tag{12.10}$$

Multiplying equation (12.9) by two and subtracting the result from equation (12.10) gives

$$3B = 13 - 4 = 9.$$

Therefore $B = 3$ and so, from (12.9), $A = -1$. The general solution is therefore

$$X_t = -2^t + 3(5^t).$$

Then

$$X_{10} = -2^{10} + 3(5^{10}) = 29{,}295{,}851.$$

Example 12.7

Solve the linear difference equation

$$X_t - 12X_{t-1} + 36X_{t-2} = 0$$

where $X_0 = 1$ and $X_1 = 8$. Determine X_9.

Solution. The characteristic equation is $x^2 - 12x + 36 = 0$ or $(x - 6)^2 = 0$. This equation has two equal roots 6, 6. So the general solution of the linear difference equation is of the form

$$X_t = (A + tB)6^t$$

where A and B are constants.

Since $X_0 = 1$, then

$$1 = X_0 = (A + 0)6^0 = A$$

and since $X_1 = 8$, then

$$8 = X_1 = (A + 1 \times B)6^1 \text{ or } 8 = 6(A + B).$$

Since $A = 1$, then $8 = 6 + 6B$ and therefore $B = \frac{1}{3}$.

The general solution of the linear difference equation is

$$X_t = \left(1 + \frac{t}{3}\right) 6^t.$$

Therefore

$$X_9 = \left(1 + \frac{9}{3}\right) 6^9 = 4 \times 6^9 = 40{,}310{,}784.$$

12.6.2 Particular Solutions

A particular solution for the general second order linear difference equation

$$X_t + aX_{t-1} + bX_{t-2} = c$$

where a, b, c are constants is as follows:

$$X_t = \begin{cases} \dfrac{c}{1+a+b} & \text{if } 1+a+b \neq 0; \\ \dfrac{ct}{2+a} & \text{if } 1+a+b = 0 \text{ and } a \neq -2; \\ \dfrac{1}{2}ct^2 & \text{if } a = -2 \text{ and } b = 1. \end{cases} \tag{12.11}$$

The conditions on a and b in the above are respectively that:

1. 1 is not a characteristic root (i.e., not a root of $x^2 + ax + b = 0$);
2. One characteristic root is 1 and the other is not;
3. Both characteristic roots are 1.

That these are particular solutions is easily verified, but we will illustrate this by example only.

It is not really necessary to remember the formulae for particular solutions. For a second order linear difference equation with constant coefficients, only one of the forms $X_t = K$, $X_t = Kt$, $X_t = Kt^2$ will work as a solution with K a constant. So one need only test three possibilities.

For instance, $X_t = K$ (K constant) cannot be a solution of

$$X_t - 3X_{t-1} + 2X_{t-2} = 21,$$

since this would require $K - 3K + 2K = 21$ or $0 = 21$, which is absurd. However, trying $X_t = Kt$ (so $X_{t-1} = K(t-1)$, $X_{t-2} = K(t-2)$) we have:

$$Kt - 3K(t-1) + 2K(t-2) = 21$$

which simplifies to $-K = 21$. Therefore, $K = -21$ and so $X_t = -21t$ is a solution of the difference equation.

Example 12.8

Find particular solutions of the following difference equations:

1. $X_t + 7X_{t-1} + 12X_{t-2} = 4$,
2. $X_t - 5X_{t-1} + 4X_{t-2} = 9$,
3. $X_t - 2X_{t-1} + X_{t-2} = 4$.

Solution.

1. In this difference equation $1 + a + b = 1 + 7 + 12 = 20 \neq 0$. Therefore,

$$X_t = \frac{c}{1 + a + b} = \frac{4}{20} = \frac{1}{5}$$

is a particular solution. That is, the repeating sequence $\frac{1}{5}$, $\frac{1}{5}$, $\frac{1}{5}, \ldots$ satisfies this difference equation. To check this, note that $\frac{1}{5} + 7 \times \frac{1}{5} + 12 \times \frac{1}{5} = 4$.

2. Here $1 + a + b = 1 - 5 + 4 = 0$ and $a = -5 \neq -2$. Then

$$X_t = \frac{ct}{a + 2} = \frac{9t}{-5 + 2} = -3t$$

is a particular solution. Thus $X_0 = 0$, $X_1 = -3$, $X_2 = -6, X_3 = -9$ and so on.

3. In this case, $a = -2$, $b = 1$, $c = 4$ so a particular solution is

$$X_t = \frac{1}{2}4t^2 = 2t^2.$$

The general second order linear difference equation is solved by adding a particular solution to the complementary solution. We illustrate this in the next example.

Example 12.9

Solve the following difference equations:

1. $X_t + 7X_{t-1} + 12X_{t-2} = 4; X_0 = 1.2,\ X_1 = 2.2,$
2. $X_t - 5X_{t-1} + 4X_{t-2} = 9; X_0 = 0,\ X_1 = 5,$
3. $X_t - 2X_{t-1} + X_{t-2} = 4.\ X_0 = 1,\ X_1 = 2.$

Solution. Particular solutions were found for these equations in the previous example; so we only need the complementary solutions.

1. The characteristic equation is

$$x^2 + 7x + 12 = 0$$

whose roots are -3, -4. The complementary solution is therefore

$$X_t = A(-3)^t + B(-4)^t$$

with A, B constants.

Therefore, the general solution (of the given inhomogeneous difference equation) is obtained by adding the particular solution $X_t = \frac{1}{5} = 0.2$ found in the previous example:

$$X_t = 0.2 + A(-3)^t + B(-4)^t.$$

Since $X_0 = 1.2$, then $1.2 = X_0 = 0.2 + A + B$, which gives

$$A + B = 1.$$

Since $X_1 = 2.2$, then $2.2 = X_1 = 0.2 + A(-3) + B(-4)$, which gives

$$3A + 4B = -2.$$

Solving these two simultaneous equations for A and B gives

$$A = 6, \qquad B = -5.$$

The solution is therefore

$$X_t = 0.2 + 6(-3)^t - 5(-4)^t.$$

2. The characteristic equation is

$$x^2 - 5x + 4 = 0$$

with roots 1, 4. The complementary solution is

$$X_t = A1^t + B4^t = A + B4^t$$

with A, B constants. A particular solution we found, in Example 12.8.2, was $X_t = -3t$. Therefore the general solution is

$$X_t = A + B(4)^t - 3t.$$

Since $X_0 = 0$, then $0 = X_0 = A + B4^0 - 0 = A + B$. Therefore, $A = -B$. Next $X_1 = 5$, gives $5 = X_1 = A + B4^1 - 3 \times 1$. That is $A + 4B = 8$. Since $A = -B$, then $3B = 8$ and so $B = \frac{8}{3}$ and $A = -\frac{8}{3}$. The general solution is therefore

$$X_t = -\frac{8}{3} + \frac{8}{3}(4^t) - 3t.$$

3. The characteristic equation is

$$x^2 - 2x + 1 = 0$$

which has two equal roots: 1, 1.

The complementary solution is therefore $X_t = A+tB$. A particular solution we found was $X_t = 2t^2$. Therefore the general solution is

$$X_t = A + tB + 2t^2.$$

Now $X_0 = 1$ gives $1 = A$; while $X_1 = 2$ gives

$$2 = A + 1 \times B + 2(1)^2 = A + B + 2.$$

Therefore $A + B = 0$. Since $A = 1$, then $B = -1$. The solution is therefore

$$X_t = 1 - t + 2t^2.$$

Example 12.10

A simplified Samuelson model for a national economy is provided by the following difference equation, where X_t is the total national income in year t:

$$X_t - c(1 + w)X_{t-1} + cwX_{t-2} = k.$$

Here c, w, k are positive constants and $c < 1$.

Find the general solution for the case $c = 0.9$, $w = 0.5$, $k = 1$, where $X_0 = 1$ and $X_1 = 1.3$.

Calculate the national income in years 10 and 20.

Solution. The difference equation for the given parameters is

$$X_t - 0.9(1.5)X_{t-1} + 0.9(0.5)X_{t-2} = 1,$$

that is

$$X_t - 1.35X_{t-1} + 0.45X_{t-2} = 1.$$

The characteristic equation is

$$x^2 - 1.35x + 0.45 = 0$$

which has roots 0.75, 0.6. The complementary solution is therefore

$$X_t = A(0.75^t) + B(0.6^t)$$

where A and B are constants.

A particular solution is (see 12.11)

$$X_t = \frac{1}{1 - 1.35 + 0.45} = \frac{1}{0.1} = 10.$$

The general solution is therefore of the form

$$X_t = A(0.75^t) + B(0.6^t) + 10.$$

Since $X_0 = 1$, then $1 = X_0 = A + B + 10$, which gives

$$A + B = -9. \tag{12.12}$$

Similarly, since $X_1 = 1.3$, then $1.3 = X_1 = A(0.75) + B(0.6) + 10$. Therefore

$$0.75A + 0.6B = -8.7. \tag{12.13}$$

Multiplying equation (12.12) by 0.6 and subtracting the result from equation (12.13) gives

$$0.15A = -8.7 + 0.6 \times 9 = -3.3.$$

It follows that $A = -22$ and, from (12.12), $B = -9 - A = -9 + 22 = 13$. The general solution is therefore

$$X_t = -22(0.75^t) + 13(0.6^t) + 10. \tag{12.14}$$

The national income in year 10 is

$$X_{10} = -22(0.75^{10}) + 13(0.6^{10}) + 10 = 8.840 \qquad \text{(to 3 decimal places).}$$

(This is almost 9 times the first year's national income $X_0 = 1$.)

In year 20, the national income is

$$X_{20} = -22(0.75^{20}) + 13(0.6^{20}) + 10 = 9.931 \qquad \text{(to 3 decimal places).}$$

If in the above example, we compute X_t for larger and larger t, the values approach a limiting value of 10. This can be seen from equation (12.14). As t increases, i.e., tends to ∞, the terms involving 0.75^t and 0.6^t tend to 0, so that X_t tends to the value 10; see Fig. 12.6.

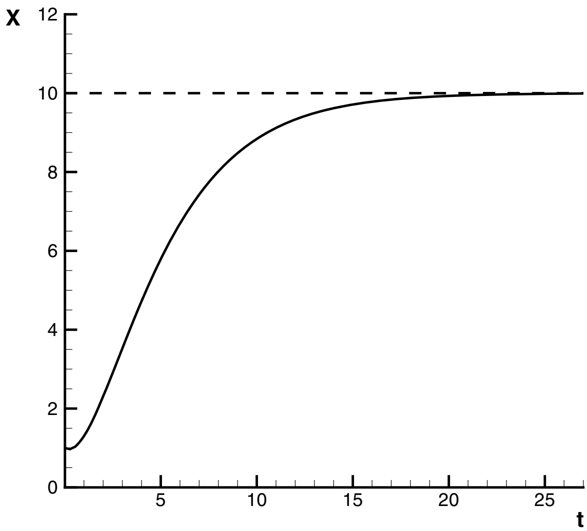

Figure 12.6 Graph of the solution X_t of Example 12.10 in continuous time, illustrating the simplified Samuelson model converging to the value 10.

12.6.3 Stability

A second order linear difference equation

$$X_t + aX_{t-1} + bX_{t-2} = c$$

where a, b, c are constants, is **stable** if both its characteristic roots are strictly between -1 and 1. Otherwise it is **divergent**.

If the characteristic roots are α, β, this condition is the same as requiring $-1 < \alpha < 1$ and $-1 < \beta < 1$. The general solution of the difference equation must then be one of the forms:

$$X_t = \begin{cases} A\alpha^t + B\beta^t + C & \text{if } \alpha \neq \beta, \\ (A + tB)\alpha^t + C & \text{if } \alpha = \beta, \end{cases}$$

where A, B are constants determined by the initial conditions (the values X_0 and X_1) and where

$$C = \frac{c}{1 + a + b}.$$

The particular solution is $X_t = C$ and the complementary solution is either $X_t = A\alpha^t + B\beta^t$ or $X_t = (A + tB)\alpha^t$.

In either case; if $-1 < \alpha < 1$ and $-1 < \beta < 1$, the complementary solution will tend to 0 as t tends to ∞. Then the general solution will tend to the particular solution.

The three difference equations in the second example of the previous section are divergent. The difference equation in the simplified Samuelson model example is stable and the solution converges on the particular solution $X_t = 10$, as is illustrated in Fig. 12.6.

When a characteristic root is 1 or -1, convergence is still possible in certain cases. One is when the sequence is constant: $X_t = k$ for all t, where k is constant. The other case occurs when one characteristic root is 1 and the other one α satisfies $-1 < \alpha < 1$.

In this case the homogeneous difference equation is

$$X_t - (1+\alpha)X_{t-1} + \alpha X_{t-2} = 0$$

and the general solution is of the form

$$X_t = A(1^t) + B(\alpha^t) = A + B\alpha^t$$

where A, B are constants.

Since $-1 < \alpha < 1$, then α^t tends to the value 0 as t tends to ∞ and therefore X_t converges to the solution $X_t = A$. The convergence is oscillatory if $-1 < \alpha < 0$.

Example 12.11

Consider the difference equation

$$2X_t - X_{t-1} - X_{t-2} = 0$$

(So the value of X_t in period t is the average of its values in the previous two periods, $t \geq 2$.)

Find the general solution and comment on stability.

Solution. The characteristic equation is $2x^2 - x - 1 = 0$, whose roots are $1, -\frac{1}{2}$. The general solution is therefore

$$X_t = A + B\left(-\frac{1}{2}\right)^t$$

where A, B are constants.

As t tends to ∞, $\left(-\frac{1}{2}\right)^t$ will tend to the value 0 and therefore X_t converges to the solution $X_t = A$. The value of A can easily be expressed in terms of X_0 and X_1. Since $X_0 = A + B$ and $X_1 = A - \frac{1}{2}B$, then $A = \frac{1}{3}(X_0 + 2X_1)$.

EXERCISES

12.1. Solve the following difference equations, commenting on stability.

a) $X_t = -3X_{t-1} + 5; \quad X_0 = 2$

b) $X_t = -\frac{1}{3}X_{t-1} + 5; \quad X_0 = -\frac{1}{4}$

c) $X_t = 2X_{t-1} - 8; \quad X_0 = 9$

d) $X_t = \frac{3}{2}X_{t-1}$

e) $X_t = \frac{2}{3}X_{t-1} - 10$

f) $X_t = 0.95X_{t-1}$.

Evaluate X_{10} in each case.

12.2. Calculate the equilibrium price of a single good in an isolated market where the supply, $Q_{S,t}$, demand $Q_{D,t}$ and price P_t in period t are given by

$$Q_{S,t} = 4P_{t-1} - 5,$$
$$Q_{D,t} = -5P_t + 10.$$

Show that the prices P_t converge and find the equilibrium price.

12.3. Solve the difference equations:

a) $X_t = X_{t-1} + 3; \quad X_0 = 0$

b) $X_t = 3 - X_{t-1}; \quad X_0 = 1$

In each case, sketch the graph of the function X_t for $t = 0, 1, 2, 3, 4$.

12.4. Solve the following difference equations, commenting on stability in each case:

a) $X_t - 6X_{t-1} + 9X_{t-2} = 2; \quad X_0 = 1.5, X_1 = 2$

b) $X_t + 2X_{t-1} - 3X_{t-2} = 7; \quad X_0 = 0, X_1 = 4$

c) $X_t - 2X_{t-1} + X_{t-2} = 6; \quad X_0 = 1, X_1 = 3$

d) $X_t - 2X_{t-1} - 15X_{t-2} = 8; \quad X_0 = 0, X_1 = 1$

e) $10X_t - 3X_{t-1} - 4X_{t-2} = 18; \quad X_0 = 2, X_1 = 0$

f) $4X_t - X_{t-2} = 9; \quad X_0 = 5, X_1 = 1$

g) $3X_t - 4X_{t-1} + X_{t-2} = 0; \quad X_0 = 4, X_1 = 0$

12.5. Solve the difference equations:

a) $2X_t - X_{t-1} - X_{t-2} = 0; \quad X_0 = 0,\ X_1 = 1$

b) $2X_t - X_{t-1} - X_{t-2} = 1; \quad X_0 = 0,\ X_1 = 1$

12.6. A simple model for total national income X_t in year t satisfies the difference equation

$$X_t - 1.26X_{t-1} + 0.36X_{t-2} = 1.$$

Show that X_t converges and give the equilibrium value.

12.7. A population model for a population X_t in year t is given by the difference equation

$$9X_t - 9X_{t-1} + 2X_{t-2} = 100.$$

Show that the population converges and give its equilibrium value. Evaluate X_{10}, given that $X_0 = 48$ and $X_1 = 49$.

12.8. A second order linear difference equation $X_t + aX_{t-1} + bX_{t-2} = c$ with $b = 0$ may be considered first order. Use this to deduce the solution of a general first order linear difference equation as given in Section 12.3, from the theory of solutions of second order equations given in Section 12.6.

13
Differential Equations

13.1 Introduction

There are close similarities between the theories of linear difference equations and linear differential equations. Indeed, differential equations may be regarded as the continuous analogues of difference equations where the variable quantity, such as time, is assumed to flow continuously rather than occurring in discrete intervals.

In market models where supply, demand, and price vary continuously and where each one is affected by the others, it is important to know their rates of change and whether these rates are increasing or decreasing. For instance, if the current price $P(t)$ is a function of time t, the economist may wish to know the first derivative $P'(t) = \frac{dP(t)}{dt}$ and the second derivative $P''(t) = \frac{d^2P(t)}{dt^2}$.

A particular model may be described by an equation relating P and its derivatives, known as a **differential equation**. Given the equation, the problem is to determine P.

Differential equations have already been encountered. Finding the indefinite integral $F(t) = \int f(t)dt$ is equivalent to solving the differential equation

$$\frac{dF}{dt} = f$$

for F when f is given.

If $y(t)$ is a function of a variable t, an $\mathbf{n^{th}}$ **order** differential equation in y is a differential equation in which n is the highest order of derivative of y that occurs in the equation. For instance, a second order differential equation can only involve $\frac{dy}{dt}$ and $\frac{d^2y}{dt^2}$.

The equation is **linear** if no m^{th} powers of y or its derivatives occur for any m other than $m = 0$ or $m = 1$. Thus, no terms such as $y^{-\frac{1}{2}}$, y^2 or $\left(\frac{dy}{dx}\right)^3$ occur.

Solving a differential equation for y means expressing y as a function of t, either implicitly or explicitly, using the equation.

Finally, we mention some simplified notation that is often used in the theory of differential equations. If y is a function of t, then $\frac{dy}{dt}$ and $\frac{d^2y}{dt^2}$ can be denoted simply by y' and y'', respectively. An alternative notation, due to Newton, is $\dot{y}$ for $\frac{dy}{dt}$ and $\ddot{y}$ for $\frac{d^2y}{dt^2}$.

13.2 First Order Linear Differential Equations

If $y(t)$ is a function of a variable t, an equation of the form

$$\frac{dy}{dt} = ay + b \tag{13.1}$$

where a, b are constants, not both 0, is known as a **first order linear differential equation**. The equation is **homogeneous** if $b = 0$; otherwise it is **inhomogeneous**.

The **associated** linear homogeneous differential equation of (13.1) is

$$\frac{dy}{dt} = ay. \tag{13.2}$$

If $y_1(t)$, $y_2(t)$ are solutions of (13.1), then $y_1' = ay_1 + b$ and $y_2' = ay_2 + b$, which implies $y_1' - y_2' = a(y_1 - y_2)$. Let $y = y_1 - y_2$. Then $y' = y_1' - y_2'$ and therefore $y' = ay$, which means y is a solution of (13.2).

It follows that any two solutions of equation (13.1) differ only in a solution of (13.2), the associated homogeneous equation. Therefore, the general solution of $y' = ay + b$ is obtained by adding a **particular solution** of that equation to the general solution of the homogeneous equation $y' = ay$. The general solution of (13.2) is known as the **complementary solution** of (13.1). The analogy with the theory of difference equations is clear (see Section 12.6).

To find a particular solution of (13.1), try one that does not change with t (i.e., one that is time invariant if t denotes time). Try $y = k$, where k is a constant to be determined. In this case $\frac{dy}{dt} = 0$, since k is constant, and so for (13.5) to hold, we require $0 = ak + b$ or $k = -\frac{b}{a}$ (if $a \neq 0$). Therefore,

$$y = -\frac{b}{a}$$

is a particular solution if $a \neq 0$. Next, we solve the homogeneous equation

$$\frac{dy}{dt} = ay.$$

Since

$$\frac{dt}{dy} = \frac{1}{\frac{dy}{dt}},$$

(see (6.10) then

$$\frac{dt}{dy} = \frac{1}{ay}$$

and so

$$t = \int \frac{1}{ay} dy = \frac{1}{a} \int \frac{1}{y} dy = \frac{1}{a} \ln y + k,$$

where k is a constant. Therefore, putting $c = ka$, this simplifies to

$$\ln y = at - c$$
$$y = e^{at-c} = e^{-c} e^{at}$$

(using the product rule for indices (1.11)). Therefore,

$$y = Ae^{at}$$

where $A (= e^{-c})$ is a constant.

Therefore, the general solution of

$$\frac{dy}{dt} = ay + b,$$

if $a \neq 0$, is

$$y = Ae^{at} - \frac{b}{a}$$

where A is a constant.

In the case $a = 0$, equation (13.1) reduces to

$$\frac{dy}{dt} = b$$

whose solution is $y = \int b dt = b \int dt = bt + K$ (remember b is a constant), where K is a constant of integration. To sum up:

The general solution of the equation $\frac{dy}{dt} = ay + b$ is:

$$y = \begin{cases} Ae^{at} - \frac{b}{a} & \text{if} \quad a \neq 0, \\ bt + K & \text{if} \quad a = 0, \end{cases}$$

where A and K are constants.

Example 13.1

Solve the differential equation

$$\frac{dy}{dt} = (3.4)y + 17$$

where $y = 3$ when $t = 0$.

Solution. Using the general solution in the box above,

$$y = Ae^{3.4t} - \frac{17}{3.4} = Ae^{3.4t} - 5.$$

When $t = 0$, $3 = y = Ae^0 - 5 = A - 5$ so that $A = 8$ and so the solution is

$$y = 8e^{3.4t} - 5.$$

Example 13.2

A model for the population $y(t)$, in millions, of some country at time t states that the rate of change of the population is given by

$$\frac{dy}{dt} = -0.05y + 4.5.$$

The population at time $t = 0$ is 100 million.

1. Evaluate $y(10)$, correct to 2 decimal places.
2. Find the value of t for which $y(t) = 91$, correct to 1 decimal place.

Solution. We are given that

$$\frac{dy}{dt} = -0.05y + 4.5.$$

The solution is

$$y = Ae^{-0.05t} - \frac{4.5}{(-0.05)} = Ae^{-0.05t} + 90.$$

1. Since $100 = y(0) = Ae^0 + 90 = A + 90$, then $A = 10$, so

   $$y(t) = 10e^{-0.05t} + 90.$$

   Therefore,

   $$y(10) = 10e^{-0.5} + 90 = 96.07 \text{ (correct to 2 decimal places)}.$$

2. If $91 = 10e^{-0.05t} + 90$, it follows that $1 = 10e^{-0.05t}$, which gives

$$e^{-0.05t} = 0.1.$$

Take the natural logarithm of each side:

$$\ln 0.1 = \ln e^{-0.05t} = -0.05t.$$

Therefore,

$$t = -\frac{1}{0.05} \ln 0.1 = 46.1 \text{ (correct to 1 decimal place).}$$

A sketch of this solution is shown in Fig. 13.1. The dashed line corresponds to the line $y = 90$.

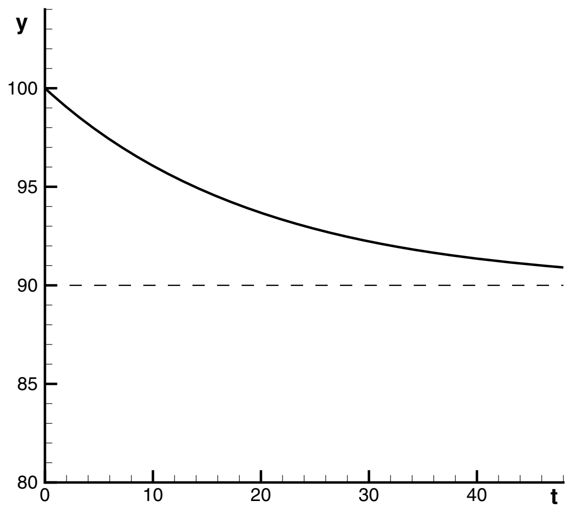

Figure 13.1 The graph of the solution of Example 13.2 on population decline. The population converges to 90 million.

13.2.1 Stability

If α is any positive number, $e^{-\alpha t}$ will tend to 0 and $e^{\alpha t}$ will tend to ∞ (increase without bound) as t tends to ∞. It follows that the solution

$$y = Ae^{at} - \frac{b}{a}$$

of the differential equation

$$\frac{dy}{dt} = ay + b \quad (a \neq 0),$$

will converge on the **equilibrium value** $-\frac{b}{a}$ as t tends to ∞ when $a < 0$. In this case, we say the solution of the equation is **stable**. Solutions that diverge are said to be **unstable**; for instance, when $a > 0$.

In the case $a = 0$, the differential equation is $\frac{dy}{dt} = b$, where $b \neq 0$. The general solution $y = bt + K$, where K is a constant, is evidently divergent since bt will tend to $\pm\infty$ with t. To sum up:

The solution of the differential equation $\dfrac{dy}{dt} = ay + b$ is

1. unstable if $a \geq 0$;
2. stable if $a < 0$. In this case the solution y converges on the particular solution $-b/a$ as equilibrium value.

The solution in Example 13.1 is unstable. The solution in Example 13.2 is stable and the equilibrium value is 90. This means the population converges towards this value as t increases (see Fig 13.1). It takes almost 46 years to reach 91 million (see the second part of Example 13.2). However, by decreasing one of the parameters, namely changing 4.5 to 3.85, it takes less than 10 years to fall from 100 to 91 million.

Note the equilibrium value $-\frac{b}{a}$ is never realized by $y = Ae^{at} - \frac{b}{a}$ if $A \neq 0$. This is because e^{at} never takes the value 0, since $e^x > 0$ for any number x.

13.3 Nonlinear First Order Differential Equations

Nonlinear differential equations are more difficult to analyse. For these equations, there are specialised techniques depending on the type of equation.

One such equation that occurs in economics is the Bernoulli equation

$$\frac{dy}{dt} = ay + by^n \tag{13.3}$$

where a, b, n are constants and $n > 1$.

This can be solved by *linearizing* it in the following way. Let $z = y^{1-n}$. Then by the chain rule (6.8)

$$\frac{dz}{dt} = \frac{dz}{dy} \times \frac{dy}{dt} = \frac{d(y^{1-n})}{dy} \times \frac{dy}{dt} = (1-n)y^{-n}\frac{dy}{dt}.$$

Therefore,

$$y^{-n}\frac{dy}{dt} = \frac{1}{(1-n)}\frac{dz}{dt}.$$

Multiplying equation (13.3) throughout by y^{-n} gives

$$y^{-n}\frac{dy}{dt} = ayy^{-n} + by^n y^{-n}.$$

That is

$$\frac{1}{(1-n)}\frac{dz}{dt} = ay^{1-n} + b = az + b.$$

Therefore,

$$\frac{dz}{dt} = (1-n)az + (1-n)b,$$

which is a linear first order differential equation. This can be solved by the method discussed earlier in Section 13.2.

Example 13.3

Solve

$$\frac{dy}{dt} = y - 2y^2,$$

given that $y(0) = \frac{1}{5}$.

Solution. This is the Bernoulli equation (13.3) with $a = 1$, $b = -2$ and $n = 2$. Let $z = y^{1-n} = y^{1-2} = y^{-1}$. Then following through the above technique, the given differential equation transforms to

$$\frac{dz}{dt} = -z + 2.$$

The general solution is

$$z = Ae^{-t} - \frac{2}{(-1)} = Ae^{-t} + 2.$$

Since $z = y^{-1}$, then

$$y^{-1} = Ae^{-t} + 2.$$

We are given that $y = \frac{1}{5}$ when $t = 0$. Therefore,

$$\left(\frac{1}{5}\right)^{-1} = Ae^0 + 2 = A + 2.$$

It follows that $A = 3$, $y^{-1} = 3e^{-t} + 2$, and therefore

$$y = \frac{1}{3e^{-t} + 2}.$$

13.3.1 Separation of Variables

If a first degree differential equation can be expressed in the form

$$f(y)\frac{dy}{dt} = g(t),$$

where f is a function only of y and g a function only of t, we can sometimes solve the equation by the method of **separation of variables**.

The technique is more easily understood if we treat dy and dt as individual quantities (called **differentials** – see Appendix A) whose ratio is the derivative $\frac{dy}{dt}$. For the purposes of our current discussion, it is enough just to accept dy and dt can be regarded as separate entities. The usefulness of this idea will become apparent from the following examples.

Example 13.4

Solve the differential equation

$$y^2\frac{dy}{dt} = 8t + 1,$$

given that $y(0) = 6$.

Solution. Write the equation as

$$y^2 dy = (8t + 1)dt.$$

Integrate both sides:

$$\int y^2 dy = \int (8t + 1)dt.$$

Then
$$\frac{1}{3}y^3 = 4t^2 + t + K,$$
where K is a constant.

When $t = 0$, $y = 6$ and so $\frac{1}{3}(6^3) = 0 + 0 + K$, which gives $K = 72$. It follows that
$$\frac{1}{3}y^3 = 4t^2 + t + 72$$
which can also be written as
$$y^3 = 3(4t^2 + t + 72).$$

Example 13.5

Solve the equation
$$\frac{dy}{dt} = \frac{1}{2}y^3t^2,$$
given that $y(0) = 1$.

Solution. Rearrange the equation to obtain
$$2y^{-3}dy = t^2dt.$$
Integrating both sides:
$$2\int y^{-3}dy = \int t^2dt,$$
gives
$$\frac{2}{-3+1}y^{-3+1} = \frac{1}{3}t^3 + K,$$
where K is a constant. Then
$$-y^{-2} = \frac{1}{3}t^3 + K.$$
Since $y = 1$ when $t = 0$, then $K = -1$. Therefore
$$-y^{-2} = \frac{1}{3}t^3 - 1.$$
Multiply throughout by -3 to get
$$3y^{-2} = -t^3 + 3$$
or
$$\frac{3}{y^2} = 3 - t^3.$$
We can also write this as
$$y^2 = \frac{3}{3 - t^3}.$$

Example 13.6

Solve the equation

$$t\frac{dy}{dt} = y^2,$$

where $y(1) = -\frac{1}{2}$.

Solution. Write the equation in the form

$$y^{-2}dy = t^{-1}dt$$

and integrate both sides to obtain

$$\int y^{-2}dy = \int t^{-1}dt.$$

That is,

$$-y^{-1} = \ln t + K \tag{13.4}$$

where K is a constant. When $t = 1$, $y = -\frac{1}{2}$; so we have

$$-\left(-\frac{1}{2}\right)^{-1} = \ln 1 + K = 0 + K.$$

Therefore $2 = K$, and substituting this in (13.4) gives

$$-y^{-1} = \ln t + 2$$

which can be rearranged as

$$y = -\frac{1}{2 + \ln t}.$$

13.4 Second Order Linear Differential Equations

The general second order differential equation is of the form

$$\frac{d^2y}{dt^2} + a\frac{dy}{dt} + by = c, \tag{13.5}$$

which can also be written as

$$y'' + ay' + by = c.$$

Here a, b, c are constants,

$$y' = \frac{dy}{dt} \text{ and } y'' = \frac{d^2y}{dt^2}.$$

The equation is **homogeneous** if $c = 0$; otherwise it is **inhomogeneous**. The associated homogeneous differential equation to (13.5) is

$$\frac{d^2y}{dt^2} + a\frac{dy}{dt} + by = 0. \tag{13.6}$$

As in the first order case, it can easily be shown that any two solutions of equation (13.5) differ by a solution of (13.6). Thus, the general solution of (13.5) is any particular solution of (13.5) plus the general solution of (13.6).

As before, the general solution of the associated homogeneous differential equation is known as the complementary solution of (13.5).

13.4.1 The Homogeneous Case

Consider the homogeneous linear differential equation (13.6). The general second order homogeneous linear difference equation had solutions of the form $X_t = A\alpha^t$, where A, α are constants. So we might try solutions of this form for the differential equations case.

However, as the function e^x is easier to differentiate than the general exponential α^x (recall that $\frac{de^x}{dx} = e^x$), we shall try solutions of the form $y = Ae^{\alpha t}$, with A, α constants. This is not a major change because α^x can be expressed as a power of e, noting that $e^{t \ln \alpha} = \alpha^t$ (since $t \ln \alpha = \ln(\alpha^t) = \log_e(\alpha^t)$).

If $y = Ae^{\alpha t}$ $(A \neq 0)$, then $y' = A\frac{d}{dt}(e^{\alpha t}) = A\alpha e^{\alpha t}$ and $y'' = A\alpha\frac{d}{dt}(e^{\alpha t}) = A\alpha^2 e^{\alpha t}$. Therefore, $y = Ae^{\alpha t}$ is a solution of the homogeneous equation if and only if

$$A\alpha^2 e^{\alpha t} + aA\alpha e^{\alpha t} + bAe^{\alpha t} = 0.$$

Dividing throughout by $Ae^{\alpha t}$ gives

$$\alpha^2 + a\alpha + b = 0.$$

This is the condition for α to be a root of the quadratic equation

$$x^2 + ax + b = 0$$

which we will call the **characteristic equation** of the differential equation. Its roots are the **characteristic roots**. The similarity with difference equations is clear (see Section 12.6).

If we allow $A = 0$, then $y = 0$, which is still a solution of the homogeneous differential equation. It follows that $y = Ae^{\alpha t}$ is a solution for any constant

A, where α is any one of the two characteristic roots. If α, β are the two characteristic roots, there are two combinations of this basic type solution that give the general solution of $y'' + ay' + b = 0$ depending on where α, β are equal or not. They are as follows:

$$y = \begin{cases} Ae^{\alpha t} + Be^{\beta t} & \text{if } \alpha \neq \beta, \\ (A + tB)e^{\alpha t} & \text{if } \alpha = \beta, \end{cases}$$

where A, B are constants. The values of A, B can be determined from **boundary conditions**; for instance the values of $y(0)$ and $y'(0)$ are given, or the values of $y(0)$ and $y(1)$.

Example 13.7

Solve the differential equations

1. $\dfrac{d^2y}{dt^2} + 5\dfrac{dy}{dt} + 6y = 0; \qquad y(0) = 0 \text{ and } y'(0) = 4,$
2. $\dfrac{d^2y}{dt^2} - \dfrac{dy}{dt} - 6y = 0; \qquad y(0) = 1 \text{ and } y'(0) = 5,$
3. $\dfrac{d^2y}{dt^2} - 6\dfrac{dy}{dt} + 9y = 0; \qquad y(0) = 1 \text{ and } y'(0) = 1.$

Solution.

1. The characteristic equation is

$$x^2 + 5x + 6 = (x + 2)(x + 3) = 0.$$

The characteristic roots are therefore -2, -3. The solution is therefore

$$y = Ae^{-2t} + Be^{-3t}$$

and so $y' = -2Ae^{-2t} - 3Be^{-3t}$. Since

$$0 = y(0) = Ae^0 + Be^0 = A + B,$$

then $A = -B$. Since also

$$4 = y'(0) = -2A - 3B$$

then

$$4 = -2A - 3(-A) = -2A + 3A = A.$$

Therefore $A = 4 = -B$, and the solution is

$$y = 4e^{-2t} - 4e^{-3t}.$$

2. The characteristic equation is

$$x^2 - x - 6 = (x+2)(x-3) = 0.$$

The characteristic roots are therefore -2, 3. The solution is therefore

$$y = Ae^{-2t} + Be^{3t}.$$

Then

$$y' = -2Ae^{-2t} + 3Be^{3t}.$$

Since $y(0) = 1$, then

$$A + B = 1.$$

Since $y'(0) = 5$, then

$$-2A + 3B = 5.$$

Solving the simultaneous equations gives $A = -0.4$ and $B = 1.4$. The solution is therefore

$$y = -0.4e^{-2t} + 1.4e^{3t}.$$

3. The characteristic equation is

$$x^2 - 6x + 9 = (x-3)^2 = 0.$$

Therefore, there are two equal characteristic roots 3, 3. The solution is therefore

$$y = (A + tB)e^{3t}.$$

We have

$$1 = y(0) = (A+0)e^0 = A.$$

Since

$$y' = Be^{3t} + (A+tB)3e^{3t},$$

using the rule for differentiation of a product of functions (6.6), then

$$\begin{aligned} y'(0) &= Be^0 + (A+0)3e^0 \\ &= B + 3A \\ &= B + 3 \quad (\text{since } A = 1). \end{aligned}$$

Therefore, since $y'(0) = 1$, then $B = -2$. It follows that the solution is

$$y = (1-2t)e^{3t}.$$

13.4.2 The General Case

We have shown how to solve homogeneous linear differential equations and therefore we can find complementary solutions in the inhomogeneous case. All we need now is to find **particular solutions** of equation (13.5):

$$y'' + ay' + by = c.$$

There are three cases of particular solutions:

$$y = \begin{cases} \dfrac{c}{b} & \text{if} \quad b \neq 0, \\ \dfrac{c}{a}t & \text{if} \quad b = 0, a \neq 0, \\ \dfrac{1}{2}ct^2 & \text{if} \quad a = b = 0. \end{cases}$$

It is a simple exercise to show that these are indeed particular solutions of equation (13.5). Note that the solution $y = \frac{c}{b}$ is that obtained by assuming y is constant (i.e., time invariant if t represents time).

The cases for a particular solution correspond, in order, to the cases when: 0 is not a characteristic root; exactly one characteristic root is 0; both characteristic roots are 0. Compare this with the corresponding case for difference equations in Chapter 12.

Example 13.8

If $y = y(t)$ is a function of t, solve the following differential equations for y:

1. $y'' - y' - 6y = 6$; $\quad y(0) = 0$ and $y'(0) = 5$,
2. $y'' + 5y' + 6y = -12$; $\quad y(0) = 2$ and $y'(0) = 3$,
3. $y'' - 6y' + 9y = 18$; $\quad y(0) = 0$ and $y'(0) = 1$,
4. $y'' - 4y' = 8$; $\quad y(0) = 0$ and $y(1) = 3$.

Solution.

1. From Example 13.7.2, we know that the complementary solution is of the form
$$y = Ae^{-2t} + Be^{3t}.$$
(We do not apply boundary conditions until we have the complete general solution.)

A particular solution is $y = \frac{6}{-6} = -1$, so the general solution is

$$y = Ae^{-2t} + Be^{3t} - 1.$$

Then

$$y' = -2Ae^{-2t} + 3Be^{3t}.$$

Since

$$0 = y(0) = A + B - 1,$$

then $A + B = 1$. We also have

$$5 = y'(0) = -2A + 3B.$$

Solving the simultaneous equations $A + B = 1$ and $-2A + 3B = 5$ gives $A = -0.4$ and $B = 1.4$. The solution is therefore

$$y = -0.4e^{-2t} + 1.4e^{3t} - 1.$$

2. From Example 13.7.1, the complementary solution is

$$y = Ae^{-2t} + Be^{-3t}.$$

A particular solution is $y = \frac{-12}{6} = -2$. Therefore, the general solution is

$$y = Ae^{-2t} + Be^{-3t} - 2.$$

Then

$$y' = -2Ae^{-2t} - 3Be^{-3t}.$$

Since $y(0) = 2$, then $2 = A+B-2$ and since $y'(0) = 3$, then $3 = -2A-3B$. Solving the simultaneous equations $A + B = 4$ and $2A + 3B = -3$ gives $A = 15$ and $B = -11$. The solution is therefore

$$y = 15e^{-2t} - 11e^{-3t} - 2.$$

3. From Example 13.7.3, the complementary solution is

$$y = (A + tB)e^{3t}.$$

A particular solution is $y = \frac{18}{9} = 2$. The general solution is therefore

$$y = (A + tB)e^{3t} + 2.$$

Since $0 = y(0) = (A + 0)e^0 + 2 = A + 2$, then $A = -2$. Since

$$y' = Be^{3t} + (A + tB)3e^{3t},$$

using the rule for differentiation of a product of functions (6.6), then

$$y'(0) = B + 3A.$$

Therefore, since $y'(0) = 1$ and $A = -2$, then $B = 7$. The general solution is therefore

$$y = (7t - 2)e^{3t} + 2.$$

4. The characteristic equation is

$$x^2 - 4x = x(x-4) = 0$$

and the characteristic roots are therefore 4, 0. The complementary solution is

$$y = Ae^{4t} + Be^{0t} = Ae^{4t} + B.$$

A particular solution is $y = \frac{8}{-4}t = -2t$. The general solution is therefore

$$y = Ae^{4t} + B - 2t.$$

Since $0 = y(0) = A + B$ and $3 = y(1) = Ae^4 + B - 2$, then $B = -A$ and $5 = Ae^4 + B$. Then $5 = Ae^4 - A = A(e^4 - 1)$. Therefore, $A = \frac{5}{e^4-1} = -B$ and the general solution is

$$y = \frac{5}{e^4 - 1}(e^{4t} - 1) - 2t.$$

13.4.3 Stability

To discuss the stability of the second order linear differential equation

$$\frac{d^2y}{dt^2} + a\frac{dy}{dt} + by = c,$$

we shall assume $b \neq 0$ in order to simplify matters by avoiding degenerate cases. The condition $b \neq 0$ is equivalent to the condition that the characteristic roots α, $\beta \neq 0$.

The general solution of the differential equation is then

$$y = \begin{cases} Ae^{\alpha t} + Be^{\beta t} + \dfrac{c}{b} & \text{if} \quad \alpha \neq \beta, \\ (A + tB)e^{\alpha t} + \dfrac{c}{b} & \text{if} \quad \alpha = \beta, \end{cases}$$

where A, B are constants.

Since $e^{\gamma t}$ tends to 0 or ∞ according as $\gamma < 0$ or $\gamma > 0$, the solution y will diverge if either α or β is positive; while if α, β are both negative, the complementary solution tends to 0 and so y converges on the particular solution $\frac{c}{b}$, the equilibrium value.

In Examples 13.8.1 and 13.8.3, the solution diverges, while in Example 13.8.2 it converges to the equilibrium value -2, the particular solution in that case.

EXERCISES

13.1. Solve the following differential equations for the function $y = y(t)$ of t.

a) $\dfrac{dy}{dt} = 5y + 6; \qquad y(0) = 1,$

b) $\dfrac{dy}{dt} = -3y + 4; \qquad y(0) = \dfrac{1}{3},$

c) $\dfrac{dy}{dt} = 0.8y + 12; \qquad y(0) = 5.$

Comment on stability for each of these equations and sketch the graph of y against t.

13.2. Solve the following differential equations for $y = y(t)$ and sketch the graph of y against t.

a) $\dfrac{dy}{dt} = 4; \qquad y(0) = 7,$

b) $\dfrac{dy}{dt} = 4t; \qquad y(0) = 1.$

13.3. In a population model, the population $y(t)$ (thousands) at time t (years) satisfies

$$y' = -0.05y + 2.$$

The initial population is 100,000. What is the equilibrium value of the population? When does the population fall to within 1,000 of this equilibrium value? Sketch the graph of population against time.

13.4. Solve the following differential equations for $y = y(t)$.

a) $y' = 1.2y - y^2; \qquad y(0) = 2,$

b) $y' = -2y + 5y^{1.2}; \qquad y(0) = 1.$

13.5. Solve the following differential equations for $y = y(t)$.

a) $y^2 \dfrac{dy}{dt} = 4t; \qquad y(0) = 3,$

b) $yt \dfrac{dy}{dt} = 1; \qquad y(1) = 1,$

c) $\dfrac{dy}{dt} = yt; \qquad y(0) = 3,$

d) $\dfrac{dy}{dt} = \dfrac{2t+1}{6y^2}; \qquad y(0) = 1,$

e) $e^t \dfrac{dy}{dt} = y^2; \qquad y(0) = 0.5.$

13.6. Solve the following differential equations for $y = y(t)$. In each case, comment on stability.

a) $\dfrac{d^2y}{dt^2} - 2\dfrac{dy}{dt} - 15y = 0; \qquad y(0) = 5, \quad y'(0) = 1,$

b) $\dfrac{d^2y}{dt^2} + 8\dfrac{dy}{dt} + 15y = 30; \qquad y(0) = 2, \quad y'(0) = 1,$

c) $\dfrac{d^2y}{dt^2} - 8\dfrac{dy}{dt} + 16y = 4; \qquad y(0) = 4, \quad y'(0) = 10.$

13.7. Solve the following differential equations.

a) $\dfrac{d^2y}{dt^2} + 3\dfrac{dy}{dt} = 0; \qquad y(0) = 1, \quad y'(0) = 3,$

b) $\dfrac{d^2y}{dt^2} - 4y = 12; \qquad y(0) = 0, \quad y'(0) = 6,$

c) $\dfrac{d^2y}{dt^2} - 10\dfrac{dy}{dt} = 5; \qquad y(0) = 0, \quad y'(0) = \dfrac{1}{2},$

d) $\dfrac{d^2y}{dt^2} = 10; \qquad y(0) = 1, \quad y'(0) = 2.$

A
Differentials

In defining the derivative $\frac{dy}{dx}$ of a function y of x in Chapter 6, we said that dy and dx should not be regarded as separate quantities. However, with the appropriate interpretation, we can regard dx and dy individually (they are then called **differentials**) and regard $\frac{dy}{dx}$ as their ratio.

The geometric meaning of a differential can be seen in Fig. A.1, which shows part of the graph of a function $y = f(x)$. A general point P on the graph has coordinates $(x,\, y)$, where $y = f(x)$. If x changes a small amount Δx, the corresponding point Q on the curve has coordinates $(x + \Delta x,\, y + \Delta y)$, where $y + \Delta y = f(x + \Delta x)$. Since $y = f(x)$, then $\Delta y = f(x + \Delta x) - f(x)$. In Fig. A.1, Δx is the length PB and Δy the length QB.

The tangent slope $\frac{dy}{dx}$ at P is the rate of change of y relative to x at P; or approximately the change in y resulting from a unit increase in x. So an estimate for Δy is $\frac{dy}{dx} \times \Delta x$. This is the length AB in Fig A.1.

The **differential** dy of any function $y = f(x)$ of x is defined by

$$dy = \frac{dy}{dx} \times \Delta x = f'(x) \times \Delta x. \tag{A.1}$$

In particular, since x is itself a function of x, then taking $y = x$ we have $dx = \frac{dx}{dx} \times \Delta x = 1 \times \Delta x = \Delta x$. Therefore, $dx = \Delta x$. It follows that if x changes by a very small amount dx, then

$$dy = \frac{dy}{dx} \times dx = f'(x)dx$$

is the change in y, calculated using the current rate of change $\frac{dy}{dx}$ of y relative to x.

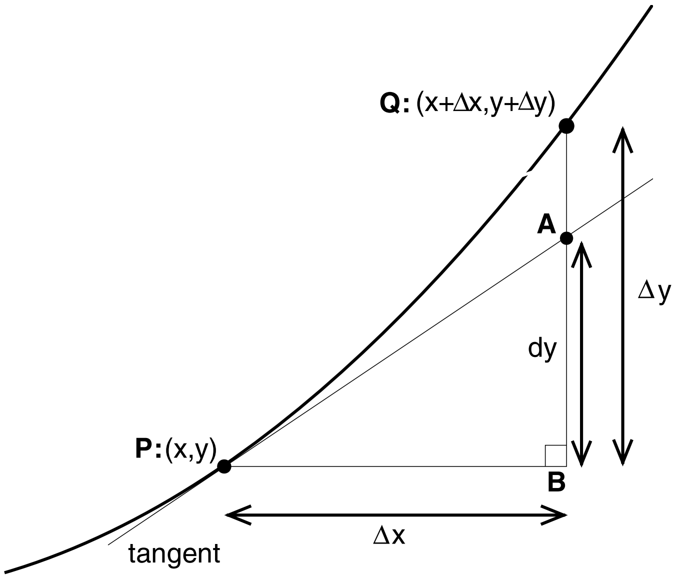

Figure A.1 Geometric interpretation of a differential.

For example, for the function $y = f(x) = x^3$, we have

$$\frac{dy}{dx} = f'(x) = 3x^2.$$

Therefore

$$dy = 3x^2 dx.$$

Thus, if $x = 2$ and x increases to 2.001, the change in x is $\Delta x = dx = 0.001$ and $f'(2) = 3 \times 2^2 = 12$. Therefore

$$dy = f'(2) \times dx = 12 \times 0.001 = 0.012.$$

The actual change in y is

$\Delta y = f(2.001) - f(2) = (2.001)^3 - 2^3 = 0.012006$ (correct to 6 decimal places).

This concept of differentials extends to functions of two or more variables in a natural way. If $z = f(x, y)$, the differentials dx, dy, dz are related by

$$dz = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy. \tag{A.2}$$

This relation can be used to obtain the total derivative formula (see Chapter 8).

If x and y are functions of a variable t and if dx, dy, dz are the differentials corresponding to a change dt in t, then dividing both sides of (A.2) by dt gives

$$\frac{dz}{dt} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}.$$

Index