Mathematics 2

Martin Anthony

Contents

Chapter 1

General introduction

Mathematics 1 and Mathematics 2

If you are studying this subject, you will already have studied (or be concurrently studying) **Mathematics 1** or an equivalent. This subject builds upon Mathematics 1. Everything in Mathematics 1 is essential to Mathematics 2. So, although Mathematics 2 is formally a separate subject, it is best thought of as an extension of Mathematics 1. Given this, it is essential that you have a good understanding of Mathematics 1. In this subject guide, we will briefly review some of the important ideas and techniques from Mathematics 1 that we shall need for Mathematics 2, but you should refer to the Mathematics 1 guide and the textbooks if you feel the need to refresh yourself on some of the more basic Mathematics 1 topics.

In Mathematics 2, we explore further some topics introduced in Mathematics 1 and we study some more applications to the social sciences, particularly economics. In particular, we investigate further the applications of differentiation and integration, and functions of several variables: we shall see some new applications and also some new techniques.

This subject also introduces some important new topics: for example, we shall meet new techniques for solving linear equations, and we introduce differential and difference equations.

Studying mathematics

I make a number of points in the introductory chapter of the Mathematics 1 guide about the nature of studying mathematics, and these are worth repeating here.

The study of mathematics can be very rewarding. It is particularly satisfying to solve a problem and know that it is solved. Unlike many of the other subjects you will study, in the mathematics subjects, there is always a right answer. Although there may be only one right (final) answer, there could be a number of different ways of obtaining that answer, some more complex than others. Thus, a given problem will have only one 'answer', but many 'solutions' (by which we mean routes to finding the answer). Generally, a mathematician likes to find the simplest solution possible to a given problem, but that does not mean that any other solution is wrong. (There may be different, equally simple, solutions.)

With mathematical questions, you first have to work out precisely what it is that the question is asking, and then try to find a method (hopefully a nice, simple one) which will solve the problem. This second step involves some degree of creativity, especially at an advanced level. You must realise that you can hardly be expected to look at every mathematics problem and write down a beautiful and concise solution, leading to the correct answer, straight away. For obvious reasons, teachers, lecturers, and textbooks rarely give that impression: they present the solution right there on the page or the blackboard, with no indication of the time a student might be expected to spend thinking—or of the dead-end paths he or she might understandably follow—before a solution can be found. It is a good idea to have scrap paper to work with so that you can try out various methods of solution. You must not get frustrated if you can't solve a problem immediately. As you proceed through the subject, gathering more experience, you will develop a feel for which techniques are likely to be useful for particular problems. You should not be afraid to try different techniques, some of which may not work, if you cannot immediately recognise which technique to use.

How to use the subject guide

This subject guide is **absolutely not** a substitute for the textbooks. It is only what its name suggests: a guide to the study and reading you should undertake. In each of the subsequent chapters, brief discussions of the syllabus topics are presented, together with pointers to recommended readings from the textbooks. **It is essential that you use textbooks.** Generally, it is a good idea to read the texts as you work through a chapter of the guide. It is most useful to read what the guide says about a particular topic, then do the necessary reading, then come back and re-read what the guide says to make sure you fully understand the topic. Textbooks are also an invaluable source of examples for you to attempt.

You should not necessarily spend the same amount of time on each chapter of the guide: some chapters cover much more material than others. The given division into chapters is the way it is in order to group together topics on particular central themes.

The discussions of some topics in this guide are rather more extensive than others. Often, this is not because those topics are more significant, but because the textbook treatments are not as extensive as they might be.

Within each chapter of the guide you will encounter 'Activities'. You should carry out these activities as you encounter them: they are designed to help you understand the topic under discussion. Solutions to the activities are given near the ends of the chapters, but do make a serious attempt at them before consulting the solutions.

To help your time management, the chapters and topics of the subject are converted below into **approximate** percentages of total time. However, this is purely for indicative purposes. Some of you will know the basics quite well and need to spend less time on the earlier material, while others might have to work hard to comprehend the very basic topics before proceeding onto the more advanced.

At the end of each chapter, you will find a list of 'Learning Outcomes'. This

Chapter	Title	% Time
2	Further Differentiation and Integration	15
3	Functions of Several Variables	10
4	Linear Algebra	25
5	Differential Equations	25
6	Difference Equations	25

indicates what you should be able to do having studied the topics of that chapter. At the end of each chapter, there are sample examination questions, which are based largely on past exam questions. Some of the 'sample examination questions' are really only samples of **parts** of exam questions. Solutions are given at the ends of the chapters.

Recommended books

None of the books in the following reading list covers absolutely everything in the syllabus. You should therefore ensure that you have access to a sufficient number of the books. At the beginning of each chapter, appropriate references are made to the books dealing with the material of that chapter.

The main recommended text is the book by Anthony and Biggs. This covers most of the required material, and uses the same notations as this guide. This is the recommended book for **Mathematics 1**, so most of you will already have a copy.

Main text

⋆ Anthony, M. and Biggs, N., *Mathematics for Economics and Finance.* (Cambridge University Press, Cambridge, UK, 1996.) [ISBN 0 521 55113 7 (hardback) and ISBN 0 521 55913 8 (paperback)].

⋆ Recommended for purchase.

The following other books are recommended.

Binmore, K. and Davies, J. *Calculus.* (Cambridge University Press, Cambridge, UK, 2001) [ISBN 0521775418]

Black, J., and Bradley, J.F. *Essential Mathematics for Economists.* Second Edition (J. Wiley and Sons, Chichester, England, 1980) [ISBN 0-471-27659-6 (cloth), 0-471-27660-X (paperback)].

Bradley, T., and Patton, P. *Essential Mathematics for Economics and Business.* (Wiley, 1998) [ISBN 0 471 97511 7].

Dowling, Edward T. *Introduction to Mathematical Economics.* Second Edition. Schaum's Outline Series. (McGraw-Hill, 1992) [ISBN 0-07-017674-4]. This is a revised edition of the author's *Mathematics for Economists* (McGraw-Hill, 1980.)

Holden, K. and Pearson, A.W. *Introductory Mathematics for Economics and Business.* Second Edition. (The Macmillan Press, London, 1992) [ISBN 0-333-57649-7 (hardback), 0-333-57650-0 (paperback)]. This is a revised and expanded edition of *Introductory Mathematics for Economists* (The Macmillan Press, 1983).

Ostaszewski, A. *Mathematics in Economics: Models and Methods.* (Blackwell, Oxford, UK, 1993) [ISBN 0-631-18055-9 (hardback), 0-631-18056-7 (paperback)].

Simon, C.P. and Blume, L. *Mathematics for Economists.* (W.W. Norton and Company Ltd, New York and London, 1994) [ISBN 0-393-95733-0].

The book by Anthony and Biggs covers most, but not all, of the required material. (You will need to refer elsewhere, as indicated in the guide, for the topics of numerical integration, Taylor series, and diagonalisation and its applications in differential and difference equations.) Each chapter of Anthony and Biggs has a large section of fully worked examples, and a selection of exercises for the reader to attempt. Since this is the key text for Mathematics 1, it also will be useful for revision of that subject.

Binmore and Davies cover all the calculus you will need, and a lot more.

Black and Bradley cover much of the necessary mathematics, explaining the economic applications.

The book by Bradley and Patton covers most of the more basic material, and has plenty of worked examples.

Dowling's book contains lots of worked examples. It is, however, less concerned with explaining the techniques. It would not be suitable as your main text, but it is a good source of additional examples.

The book by Holden and Pearson covers most of the material, and has discussions of economic applications.

The book by Ostaszewski is very suitable for a number of the topics, and provides many examples.

The book by Simon and Blume is a large book covering everything in this subject and also many topics outside the coverage of this subject.

There are many other books which cover the material of this subject, but those listed above are the ones I shall refer to explicitly.

The single most important point to be made about learning mathematics is that to learn it properly, you have to do it. **Do** work through the worked examples in a textbook and **do** attempt the exercises. This is the real way to learn mathematics. In the examination, you are hardly likely to encounter a question you have seen before, so you must have practised enough examples to ensure that you know your techniques well enough to be able to cope with new problems.

Examinations advice

A sample exam paper may be found at the end of this subject guide. You will see that there is a section of compulsory questions and a section from which you choose questions. Any changes to exam format will be announced in examiners' reports. It is worth making a few comments about exam technique. Perhaps the most important, though obvious, point is that you do not have to answer the questions in any particular order; choose the order that suits you best. Some students will want to do easy questions first to boost their confidence, while others will like to get the difficult ones out of the way. It is entirely up to you. Another point, often overlooked by students, is that you should **always** include your working. This means two things. First, do not simply write down the answer in the exam script, but explain your method of obtaining it (that is, what I called the 'solution' earlier). Secondly, include your rough working. Why should you do this? First, if you have just written down the answer without explaining how you obtained it, then you have not convinced the examiner that you know the techniques, and it is the techniques that are important in this subject. (The examiners want you to get the right answers, of course, but it is more important that you prove you know what you are doing: that is what is really being examined.) Secondly, if you have not completely solved a problem, you may still be awarded marks for a partial, incomplete, or slightly wrong, solution: if you have written down a wrong answer and nothing else, no marks can be awarded. (You may have carried out a lengthy calculation somewhere on scrap paper where you made a silly arithmetical error. Had you included this calculation in the exam answer book, you would probably not have been heavily penalised for the arithmetical error.) It is useful, also, to let the examiner know what you are thinking. For example, if you know you have obtained the wrong answer to a problem, but you can't see how to correct it, say so!

As mentioned above, you will find that, wherever appropriate, there are sample exam questions at the end of the chapters. These are based in large part on the questions appearing in past examination papers. As such, they are an indication of the types of question that might appear in future exams. But they are **just** an indication. The examiners want to test that you know and understand a number of mathematical methods and, in setting an exam paper, they are trying to test whether you do indeed know the methods, understand them, and are able to use them, and not merely whether you vaguely remember them. Because of this, you will quite possibly encounter some questions in your exam which seem unfamiliar. Of course, you will only be examined on material in the syllabus. Furthermore, you should **not** assume that your exam will be almost identical to the previous year's: for instance, just because there was a question, or a part of a question, on a certain topic last year, you should not assume there will be one on the same topic this year. For this reason, you cannot guarantee passing if you have concentrated only on a very small fraction of the topics in the subject. This may all sound a bit harsh, but it has to be emphasised.

The use of calculators

You will not be permitted to use calculators of any type in the examination. This is not something that you should panic about: the examiners are interested in assessing that you understand the key methods and techniques, and will set questions which do not require the use of a calculator.

In this guide, I will perform some calculations for which a calculator would be needed, but you will not have to do this in the exam questions. Look carefully at the answers to the sample exam questions to see how to deal with calculations. For example, if the answer to a problem is $\sqrt{2}$, then leave the answer like that: there is no need to express this number as a decimal (for which one would need a calculator).

Chapter 2

Further differentiation and integration, with applications

Essential reading

(For full publication details, see Chapter 1.)

Anthony, M. and Biggs, N. *Mathematics for Economics and Finance.* Chapters 6, 7, 8, 9, 25 and 26.

Further reading

Binmore, K. and Davies, J., *Calculus.* Section 2.13 (on Taylor's theorem) and Section 10.5 (on producer and consumer surplus)

Black, J. and Bradley, J.F. *Essential Mathematics for Economists.* Chapters 1 (for analysis of tax), 2, 5, 6, and 9.

Bradley, T., and Patton, P. *Essential Mathematics for Economics and Business.* Chapters 6 and 8.

Dowling, Edward T. *Introduction to Mathematical Economics.* Second Edition. Chapters 3, 4, 16,17.

Holden, K, and Pearson, A.W., *Introductory Mathematics for Economics and Business.* Chapters 5 and 6.

Ostaszewski, A. *Mathematics in Economics: Models and Methods.* Chapters 10 and 14.

Simon, C.P. and Blume, L., *Mathematics for Economists*: Sections 3.5, 3.6 and 20.1.

Introduction

You will have studied **Mathematics 1** or its equivalent, and be familiar with differentiation and integration. In this chapter of the subject guide we will look at some more applications of both differentiation and integration.

Revision of Mathematics 1 differentiation

From Mathematics 1, you should remember the meaning of the derivative, and know how to calculate derivatives. You should also know about some of the applications of the derivative.

Just to remind you, here are some of the important things we learned in Mathematics 1.

- The derivative of a function $f(x)$ at a point a is the instantaneous rate of change of the function at a. Formally,

$$f'(a) = \lim_{h \to 0} \frac{f(a+h) - f(a)}{h}.$$

- We have the following standard derivatives:

$f(x)$	$f'(x)$
x^k	kx^{k-1}
e^x	e^x
$\ln x$	$1/x$
$\sin x$	$\cos x$
$\cos x$	$-\sin x$

- We have the following rules for calculating derivatives

  The sum rule: If $h(x) = f(x) + g(x)$ then $h'(x) = f'(x) + g'(x)$.

  The product rule: If $h(x) = f(x)g(x)$, $h'(x) = f'(x)g(x) + f(x)g'(x)$.

  The quotient rule: If $h(x) = f(x)/g(x)$ and $g(x) \neq 0$ then

$$h'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}.$$

  The composite function rule: if $f(x) = s(r(x))$, $f'(x) = s'(r(x))r'(x)$.

- The derivative and second derivative can be used in optimisation and in curve-sketching. Particular applications in economics include profit-maximisation.

Using derivatives for approximations

Another way of looking at the definition of f' is to think of it as an **approximation** [1], which tells us how a small change in the input x affects the output $f(x)$. If we denote a small change in x by Δx, then the resulting change in $f(x)$ is

$$\Delta f = f(x + \Delta x) - f(x).$$

[1] *See Anthony and Biggs, Section 6.1*

Since $f'(x)$ is the limit of $\Delta f/\Delta x$ as Δx approaches zero, for small values of Δx we have

$$f'(x) \simeq \frac{\Delta f}{\Delta x}, \quad \text{or} \quad \Delta f \simeq f'(x)\Delta x,$$

where the symbol '$\simeq$' means 'is approximately equal to'. In 'd' notation (in which the derivative is denoted by df/dx rather than f'), we have

$$\Delta f \simeq \frac{df}{dx}\Delta x.$$

We made mention of this idea when we discussed marginal cost in Mathematics 1, and also when we looked at the meaning of the Lagrange multiplier.

A couple of examples will serve to illustrate this simple use of the derivative for approximations.

Example: Let us use the derivative to find the approximate change in the function $f(x) = x^4$ when x changes from 3 to 3.005.

The derivative of f is $f'(x) = 4x^3$ and we therefore have the approximation

$$\Delta f \simeq f'(3)\Delta x = 4(3)^2 \times 0.005 = 0.540.$$

The actual value of the change is

$$\Delta f = f(3.005) - f(3) = (3.005)^4 - 3^4 = 0.54135,$$

so our approximation is correct to two decimal places.

Example: Suppose the demand equation for a good is $p^3 q = 8000$ where q is the number of units demanded (in thousands per week) and p is the price per unit, in dollars. If p is increased from \$20 to \$21, what will be the approximate fall in sales, approximately? If, on the other hand, production were to be increased from 1000 units per week to 1100, what would be the approximate fall in price?

The demand function and its derivative are

$$q^D(p) = 8000p^{-3}, \quad (q^D)'(p) = 8000 \times (-3)p^{-4} = \frac{-24000}{p^4}.$$

Therefore when $p = 20$ and $\Delta p = 1$ we have

$$\Delta q \simeq (-24000/p^4)\Delta p = -(24000/20^4)(1) = -0.15.$$

Remembering that q is measured in thousands of units, it follows that 150 fewer units will be sold per week.

For the second question we have to consider p as a function of q, and so we need the inverse demand function and its derivative:

$$p^D(q) = 20q^{-1/3}, \quad (p^D)'(q) = 20 \times (-1/3)q^{-4/3}.$$

So when $q = 1$ and $\Delta q = 0.1$ we have

$$\Delta p \simeq (-20/3)q^{-4/3} \times 0.1 = -2/3.$$

The conclusion is that the price falls by about 67 cents.

Activity 2.1 *Suppose that the demand for a good is given by the equation*

$$p^2 q = 6000,$$

where q is the quantity (in thousands) and p is the price in dollars. If the price is increased from \$20 to \$21, what is the approximate fall in expected sales?

Taylor's theorem

We have just seen that the derivative may be used to give an approximation to a small change in value of a function. A more precise result along these lines is **Taylor's theorem**. This relates the function value not just to the derivative, but higher-order derivatives (second-order, third-order, and so on).

Under certain circumstances, functions can be approximated well by certain polynomials, referred to as **power series**. [2] The appropriate way to do this is by using 'Taylor series'. To explain this, we first introduce a piece of notation. The derivative of a function f is denoted f' and the second derivative is denoted by f''. If we differentiate the second derivative, we obtain the third derivative $f^{(3)}$, and if we differentiate this we obtain the fourth derivative $f^{(4)}$, and so on. In general, we have the nth derivative $f^{(n)}$. Taylor's theorem states that if f is a function which can be differentiated n times, then, for certain ranges of x, the approximation

$$f(x) \simeq f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f^{(3)}(0)}{3!}x^3 + \frac{f^{(4)}(0)}{4!}x^4 + \cdots + \frac{f^{(n)}(0)}{n!}x^n,$$

is valid, where $n! = n(n-1)(n-2)\dots 2$ (called 'n factorial') is the product of all the positive integers up to and including n. (Generally, the larger n, the better the approximation.) The right-hand side of this approximation is known as the Taylor series for f. Strictly speaking, the stated approximation is Taylor's theorem 'about 0', often called Maclaurin's theorem, and the right-hand side is the Maclaurin series. The right-hand side is known as a power series expansion of the function f.

[2] *See, for example, Holden and Pearson, Section 5.15, Ostaszewski, Sections 14.1–14.2, and Binmore and Davies, ***, for discussion of Taylor's theorem.*

A more general form of Taylor's theorem is often useful. This states that (again, with certain qualifications on the set of x for which the approximation is true), we have

$$f(x) \simeq f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f^{(3)}(a)}{3!}(x-a)^3 + \cdots + \frac{f^{(n)}(a)}{n!}(x-a)^n.$$

This is called the Taylor expansion of f **about** a. Taking $n = 1$ we obtain

$$f(x) \simeq f(a) + f'(a)(x-a).$$

Writing $x = a + \Delta h$, and $\Delta f = f(x) - f(a)$, this becomes the simple rule we met earlier:

$$\Delta f \simeq f'(a)\Delta x.$$

Thus, Taylor's theorem, for $n > 1$, is a generalisation of our simple approximation rule and will, in general, be more accurate. When we refer simply to Taylor's theorem or Taylor series we shall mean, for the sake of simplicity, those about 0 (that is, Maclaurin's theorem and Taylor's theorem): so, unless it is otherwise specified, we take $a = 0$.

Example: The exponential function $f(x) = e^x$ has all its derivatives equal to e^x. Since $e^0 = 1$, we therefore have that, for all n, $f^{(n)}(0)/n! = 1/n!$ and hence

$$e^x \simeq 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots + \frac{x^n}{n!}.$$

In fact, this approximation is valid for all x (though it requires mathematics outside the content of this subject to explain why).

Other important Taylor series are as follows:

$$\ln(1+x) \simeq x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \ldots,$$

valid for $-1 < x \leq 1$,

$$\sin x \simeq x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \ldots,$$

valid for all x, and

$$\cos x \simeq 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ldots,$$

valid for all x.

You should convince yourself that the Taylor series for these functions are as stated, by calculating the derivatives. These standard Taylor series should be remembered.

The power series expansions for more complicated functions can often be determined by using the standard power series given above, as the following example illustrates.

Example: We expand, as a power series up to x^4, the function $f(x) = \cos(\ln(1+x))$. This means that we find the Taylor approximation with $n = 4$. One approach is to calculate the first, second, third, and fourth derivatives of f and use Taylor's theorem directly. However, we'll take a different approach. (It is a good exercise to check that the other approach gives the same answer.) We'll use the facts that

$$\ln(1+x) \simeq x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4}$$

and

$$\cos y \simeq 1 - \frac{y^2}{2!} + \frac{y^4}{4!}.$$

It follows from these, by taking

$$y = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} \simeq \ln(1+x),$$

that

$$\begin{aligned} \cos(\ln(1+x)) &\simeq 1 - \frac{1}{2}y^2 + \frac{1}{24}y^4 \\ &= 1 - \frac{1}{2}\left(x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4}\right)^2 + \frac{1}{24}\left(x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4}\right)^4. \end{aligned}$$

Because we only need the terms involving x^4 or lower powers of x, the only relevant term one gets from the $y^4/24$ term is $x^4/24$. To explain why this is so, note that

$$\left(x-\frac{x^2}{2}+\frac{x^3}{3}-\frac{x^4}{4}\right)^4 \text{ is}$$

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

The terms which arise from this product are obtained by multiplying together four objects, one from each bracket. Since the term with lowest power of x in each bracket is x, it is only by taking the x from each bracket that we obtain a term with degree no more than 4; and since all terms of degree greater than 4 are to be ignored, this is therefore the only term we consider. Consider the y^2 term. This is

$$\begin{aligned}
\frac{1}{2}\left(x-\frac{x^2}{2}+\frac{x^3}{3}-\frac{x^4}{4}\right)^2 &= \frac{1}{2}\left(x-\frac{x^2}{2}+\frac{x^3}{3}-\frac{x^4}{4}\right)\left(x-\frac{x^2}{2}+\frac{x^3}{3}-\frac{x^4}{4}\right)\\
&= \frac{1}{2}\left(xx-2\frac{x^2}{2}x+\left(\frac{x^2}{2}\right)^2+2\frac{x^3}{3}x+\cdots\right)\\
&= \frac{x^2}{2}-\frac{x^3}{2}+\frac{11}{24}x^4+\cdots,
\end{aligned}$$

where '$\cdots$' indicates terms which involve powers of x greater than 4 (and terms which we can therefore ignore). Therefore,

$$\begin{aligned}
\cos(\ln(1+x)) &\simeq 1-\left(\frac{x^2}{2}-\frac{x^3}{2}+\frac{11}{24}x^4\right)+\frac{x^4}{24}\\
&= 1-\frac{x^2}{2}+\frac{x^3}{2}-\frac{5x^4}{12},
\end{aligned}$$

and this is the required expansion.

Activity 2.2 *By taking* $x = 0.1$ *in the expansion just obtained, find an approximate value for* $\cos(\ln(1.1))$.

Elasticities

When the price of a commodity increases from p to $p+\Delta p$, there is a change in the quantity demanded (usually a decrease). Suppose that the quantity demanded changes from q to $q+\Delta q$ (where Δq can be negative). Then, $q = q^D(p)$ and $q+\Delta q = q^D(p+\Delta p)$. It follows that

$$\Delta q = q^D(p+\Delta p) - q^D(p) \simeq (q^D)'(p)\Delta p.$$

The **relative change** in quantity is $\Delta q/q$ and the **relative change** in price is $\Delta p/p$. The ratio of these is

$$\frac{(\Delta q/q)}{(\Delta p/p)} = \frac{p}{q}\frac{\Delta q}{\Delta p} \simeq \frac{p}{q}(q^D)'(p).$$

Since, generally, Δq will be negative, and since $\Delta p > 0$,

$$-\frac{p}{q}\frac{\Delta q}{\Delta p}$$

will be positive. We define the **point elasticity of demand** to be

$$\varepsilon = -\frac{p}{q}\frac{dq}{dp},$$

where q denotes the demand quantity $q^D(p)$.

For a typical demand function, the elasticity of demand is positive. It should be noted that many texts omit the negative sign in the definition of elasticity of demand. The demand is said to be **elastic** if the elasticity is greater than 1, and **inelastic** if the elasticity is less than 1.

Example: Suppose that the demand curve has equation $q = 2 - p$. Then the point elasticity of demand is

$$-\frac{p}{q}\frac{dq}{dp} = -\frac{p}{q}(-1) = \frac{p}{q}.$$

This can be writen solely in terms of q as $(2-q)/q$ or in terms of p as $p/(2-p)$.

Example: Suppose that the demand function is given by $q = 5/p^2$. Then $dq/dp = -10/p^3$ and so the elasticity of demand is

$$\varepsilon = -\frac{p}{q}\frac{dq}{dp} = -\frac{p}{q}\left(\frac{-10}{p^3}\right) = -\frac{p}{(5/p^2)}\left(\frac{-10}{p^3}\right) = 2.$$

This is a constant, not depending on p or q.

Suppose a firm is producing a particular good, and let us think about how its total revenue changes if the selling price increases. If the price p rises, the quantity sold q falls; but the revenue $R = qp$ is the product of these two things, and it may rise or fall. We shall assume that a price rise applies uniformly to the entire supply of the good under consideration, which would certainly be the case if the firm is a monopoly. Using the product rule to differentiate total revenue $TR = qp$ with respect to p, remembering that q is a function of p, we get

$$TR' = (qp)' = q'p + q.$$

Thus the condition that total revenue increases, $TR' > 0$, is equivalent to $q'p + q > 0$, which is equivalent to

$$-\frac{q'p}{q} < 1,$$

or $\varepsilon < 1$. Thus the elasticity determines whether revenue increases or decreases as price increases:

• If $\varepsilon < 1$ (that is, demand is inelastic), then a small increase in price results in an increase in revenue.

• If $\varepsilon > 1$ (that is, demand is elastic), then a small increase in price results in an decrease in revenue.

The **point elasticity of supply** is defined in a way similar to that used to define point elasticity of demand, but using the supply function q^S rather than q^D, and omitting the negative sign. That is, the point elasticity of supply is given by

$$\frac{p}{q}(q^S)'(p) = \frac{p}{q}\frac{dq^S}{dp}.$$

The effects of taxation

We now look at what happens in a market when a good is taxed. (This is not entirely a differentiation topic, but this is as appropriate a chapter as any in which to place this topic.)

When a fixed amount of tax is imposed on each unit of a good, we refer to this as an **excise tax** or **per-unit tax**. The following example illustrates how we can determine the new equilibrium price and quantity in the presence of such a tax.

Example: Suppose that the demand and supply functions for a good are given by

$$q^D(p) = 40 - 5p, \quad q^S(p) = \frac{15}{2}p - 10.$$

Then we can easily determine (by solving the equation $q = q^D(p) = q^S(p)$) that the equilibrium price is $p^* = 4$. Suppose that the government imposes an excise tax of T per unit. How does this affect the equilibrium price?

The answer is found by noting that, if the new selling price is p, then, from the supplier's viewpoint, *it is as if the price were* $p - T$, because the supplier's revenue per unit is not p, but $p - T$. In other words the supply function has changed: when the tax is T per unit, the new supply function q^{S_T} is given by

$$q^{S_T}(p) = q^S(p - T) = \frac{15}{2}(p - T) - 10.$$

Of course the demand function remains the same. Let us use q^T and p^T to denote the new equilibrium values in the presence of the tax T. Then q^T and p^T satisfy the equations

$$q^T = 40 - 5p^T \quad \text{and} \quad q^T = q^{S_T}(p^T) = \frac{15}{2}(p^T - T) - 10.$$

Eliminating q^T we get

$$40 - 5p^T = \frac{15}{2}(p^T - T) - 10.$$

Rearranging this equation, we obtain

$$\left(5 + \frac{15}{2}\right)p^T = 50 + \frac{15}{2}T,$$

and so we have a new equilibrium price of

$$p^T = 4 + \frac{3}{5}T.$$

For example, if $T = 1$, the equilibrium price rises from 4 to 4.6. Unsurprisingly, the selling price has risen (and, as you will discover from the Activity below, the quantity sold has decreased). But note that, although the tax is T per unit, the selling price has risen not by the full amount T, but by the fraction $3/5$ of T. In other words, not all of the tax is passed on to the consumer.

Activity 2.3 *Show that the new equilibrium quantity is* $q^T = 20 - 3T$.

If we encountered a **percentage** tax rather than a per-unit tax, a similar sort of analysis would apply. Again, the demand equation would remain unaltered, but the supply equation would change: we would replace p by $p(1 - t)$ if the tax is $100t\%$ of the selling price. (This is because, from the supplier's point of view, the revenue obtained from each item–in other words, the effective price–is not p, but p minus $100t\%$ of p, which is $p(1 - t)$.)

Suppose that a government wishes to raise revenue by imposing an excise tax on a good. Clearly, a small tax will bring in little revenue but, on the other hand, if the tax is too large consumption will fall dramatically and the revenue will also be hit. We can often use differentiation to determine which level of tax will maximise the tax revenue to the government. Suppose that, as in the example above, we have demand and supply functions given by

$$q^D(p) = 40 - 5p, \quad q^S(p) = \frac{15}{2}p - 10.$$

We have seen that the equilibrium price and quantity in the presence of an excise tax T are

$$p^T = 4 + \frac{3}{5}T, \quad q^T = 20 - 3T.$$

(The determination of q^T was Activity 2.3.) The revenue $R(T)$ is the product of the quantity sold q^T and the excise tax T. That is,

$$R(T) = q^T \times T = (20 - 3T)T = 20T - 3T^2.$$

To find the value of T_m where this is a maximum, we first set $R'(T) = 0$. We have $R'(T) = 20 - 6T$, so that $10/3$ is the only critical point. The second derivative is $R''(T) = -6$, which is negative, so $T_m = 10/3$ is the maximum point. The maximum revenue is $R(T_m) = 100/3$.

Revision of Mathematics 1 integration

We now very briefly review the key topics in integration that were covered in **Mathematics 1**.

Suppose the function f is given, and the function F is such that $F'(x) = f(x)$. Then we say that F is an **anti-derivative** of f. Any two anti-derivatives of a given function f differ only by a constant. The general form of the anti-derivative of f is called the **indefinite integral** of $f(x)$, and denoted by

$$\int f(x)\,dx.$$

Often we call it simply the **integral** of f. It is of the form $F(x) + c$, where F is **any particular** anti-derivative of f and c is an arbitrary constant, known as a **constant of integration**. The process of finding the indefinite integral of f is usually known as **integrating** f, and f is known as the **integrand**.

Just as for differentiation, we shall have a list of **standard integrals** and some rules for combining these.

$f(x)$	$\int f(x)\,dx$
x^n $(n \neq -1)$	$\dfrac{x^{n+1}}{(n+1)} + c$
$1/x$	$\ln\|x\| + c$
e^x	$e^x + c$
$\sin x$	$-\cos x + c$
$\cos x$	$\sin x + c$

Note that the integral of $1/x$ is $\ln|x| + c$ rather than $\ln x + c$ because one cannot take the logarithm of a negative integer.

We also have the simple rules

$$\int (f(x) + g(x))\, dx = \int f(x)\, dx + \int g(x)\, dx,$$

for any functions f and g, and

$$\int kf(x)\, dx = k \int f(x)\, dx,$$

for any constant k.

If f is a function with an anti-derivative F, then the **definite integral**[3] of the function f over the interval $[a, b]$ is **defined to be**

[3] *See Anthony and Biggs, Section 25.4.*

$$\int_a^b f(x)\, dx = F(b) - F(a).$$

In Mathematics 1, we met three important techniques for integration. First, we have the **substitution** or change of variable method. Formally, this uses the fact that when we change the variable by putting $x = x(u)$, we have

$$\int f(x)\, dx = \int f(x(u))\, x'(u) du.$$

The rule for **integration by parts** is:

$$\int u'(x)v(x)\, dx = u(x)v(x) - \int u(x)v'(x)\, dx.$$

The third technique is **partial fractions**. This involves rewriting integrands of the form $p(x)/q(x)$, where p and q are polynomials, in a simpler form which makes them easier to integrate. If $p(x)$ is linear (that is, of the form $ax + b$) and $q(x)$ is a quadratic with two different roots, then the method of partial fractions applies. Suppose that $q(x) = (q - a_1)(q - a_2)$, where $a_1 \neq a_2$ and C is some number. Then it is possible to write

$$\frac{p(x)}{q(x)} = \frac{p(x)}{(x - a_1)(x - a_2)} = \frac{A_1}{(x - a_1)} + \frac{A_2}{(x - a_2)} \qquad (*)$$

for some numbers A_1 and A_2. Cross-multiplying equation $(*)$, we get

$$p(x) = A_1(x - a_2) + A_2(x - a_1).$$

The numbers A_1 and A_2 may be found by substituting $x = a_1$, $x = a_2$ in turn into this identity. When $p(x)/q(x)$ is expressed in this way, it is easy to evaluate the integral. We have

$$\int \frac{p(x)}{q(x)}\, dx = \int \left(\frac{A_1}{(x - a_1)} + \frac{A_2}{(x - a_2)} \right) dx = A_1 \ln|x - a_1| + A_2 \ln|x - a_2| + c.$$

Definite integrals and areas

There is a useful relationship between the definite integral and the area under a curve. Suppose that $f(x) \geq 0$ for x in the interval $[a, b]$. Then the area enclosed by the curve $y = f(x)$, the x-axis and the vertical lines $x = a$ and $x = b$ is equal to $\int_a^b f(x)\, dx$.

The same result is essentially true even when the function is negative on part of the interval, but in this case the area enclosed by that part of the graph of the function and the x-axis is assigned a negative sign.

Example: What is the area under the curve $y = x^2$ between $x = 1$ and $x = 2$?

To answer this, we first note that an anti-derivative of x^2 is $x^3/3$. So the area is

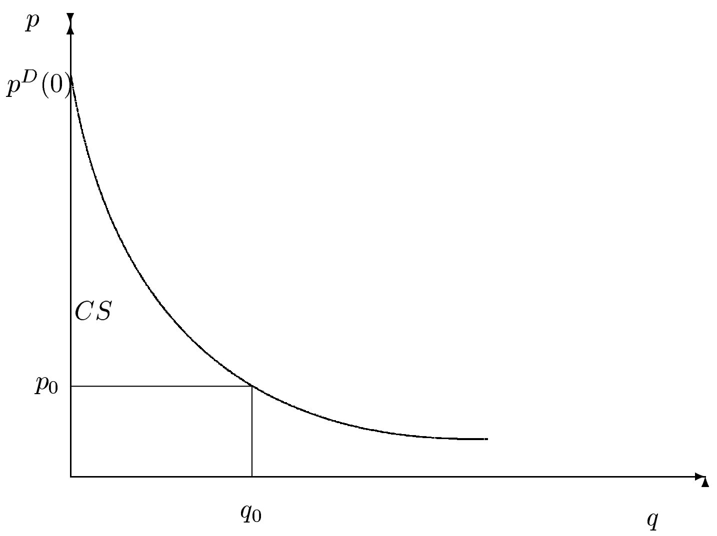
$$\int_1^2 x^2\, dx = \left[\frac{x^3}{3}\right]_1^2 = \frac{2^3}{3} - \frac{1^3}{3} = \frac{8}{3} - \frac{1}{3} = \frac{7}{3}.$$

> **Activity 2.4** *What is the area enclosed by the curve with equation $y = \sin x$, the x-axis, the y-axis and the line $x = \pi$?*

Consumer and producer surplus

Figure 2.1 illustrates a typical demand curve for a good. Here, the equilibrium price and quantity, p_0, q_0 are also indicated. (These are, of course, determined by the supply curve, which is not shown in this figure.)

Figure 2.1 Consumer surplus



At equilibrium, the consumers buy q_0 units of the good at price p_0 per unit, and hence the total amount they pay is $p_0 q_0$. However, it can be argued [4] that the total value the consumers place on q_0 items of the good is the area of the region bounded by the demand curve, the q-axis, $q = 0$, and $q = q_0$. The difference between this area and the amount they actually pay is called the **consumer surplus**. Since $p_0 q_0$ is

[4] *See Anthony and Biggs, Section 25.1*

the area of the rectangle of length q_0 and height p_0, the consumer surplus is the area of the region denoted CS. Now, the area under the demand curve from $q = 0$ to $q = q_0$ is $\int_0^{q_0} p^D(q)\, dq$ and hence the consumer surplus is given by the formula:

$$CS = \left(\int_0^{q_0} p^D(q)\, dq \right) - p_0 q_0.$$

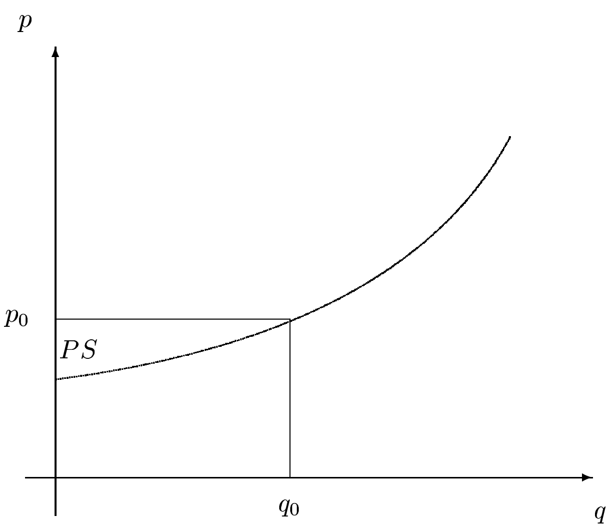
An alternative way of viewing the consumer surplus is as the region enclosed by the demand curve, the p-axis, $p = p_0$, and $p = p^D(0)$. Thus, we also have

$$CS = \int_{p_0}^{p^D(0)} q^D(p)\, dp.$$

Note that here we need to express q as a function of p on the demand curve, whereas the first formula expresses p as a function of q. Whichever formula you use, try not simply to memorise it, but rather remember what the relevant region is, and compute its area in the easiest way for a particular problem. (For example, if the demand curve is a straight line, then the region is question is triangular. Then, to compute its area and hence the consumer surplus, we don't need a complicated formula involving integrals: we would simply use the well-known fact that the area of a triangle is one half of the base times the height.)

Figure 2.2 illustrates a typical supply curve for a good. Here, the equilibrium price and quantity, p_0, q_0, are also indicated.

Figure 2.2 Producer surplus



There is a counterpart to the consumer surplus, known as the **producer surplus**. It can be argued that the total cost to the manufacturers of producing q_0 items is represented by the area under the supply curve between $q = 0$ and $q = q_0$. Yet they receive total receipts, in equilibrium, of $p_0 q_0$ for producing and selling q_0 units. The difference between these two values is known as the **producer surplus**. It is given by the formula

$$PS = p_0 q_0 - \int_0^{q_0} p^S(q)\, dq.$$

Activity 2.5 *The demand for a commodity is given by*

$$p(q+1) = 231,$$

and the supply is given by

$$p - q = 11.$$

Calculate the equilibrium price and quantity and the consumer surplus.

Learning outcomes

At the end of this chapter and the relevant reading, you should be able to:

- as in Mathematics 1, know what is meant by the derivative
- as in Mathematics 1, state the standard derivatives
- as in Mathematics 1, calculate derivatives using sum, product, quotient, and composite function (chain) rules
- as in Mathematics 1, calculate derivatives by taking logarithms
- as in Mathematics 1, establish the nature of the critical/stationary points of a function.
- as in Mathematics 1, use the derivative to help sketch functions
- as in Mathematics 1, know the terminology surrounding 'marginals' in economics, and be able to find fixed costs and marginal costs, given a total cost function
- as in Mathematics 1, know what is meant by the breakeven point and be able to determine this
- as in Mathematics 1, make use of the derivative in order to minimise or maximise functions, including profit functions
- use the derivative as a means of approximating small changes in a function
- quote Taylor's theorem, quote the standard Taylor (Maclaurin) series (for e^x, $\ln(1+x)$, $\sin x$, and $\cos x$, where the series are about $x = 0$), and know the ranges of x for which they are valid
- be able to determine power series approximations for functions using Taylor's theorem and by manipulating the standard Taylor series
- quote the definitions of point elasticity of demand and point elasticity of supply
- determine these elasticities given (respectively) the demand and supply equations or sets
- know what is meant by the terms elastic and inelastic, and be able to determine when demand is elastic or inelastic

- be able to determine the new resulting equilibrium price and quantity in the presence of an excise (per-unit) tax or a percentage tax

- as in Mathematics 1, understand the meaning of an (indefinite) integral and a definite integral

- as in Mathematics 1, state the standard integrals

- as in Mathematics 1, use integration by substitution

- as in Mathematics 1, use integration by parts

- as in Mathematics 1, integrate using partial fractions

- understand the connection between areas and definite integrals

- compute areas using definite integration

- know what is meant by consumer and producer surplus, and be able to calculate consumer and producer surpluses

Sample examination/practice questions

1. Use differentiation to find the approximate change in $\sqrt{x}$ as x increases from 100 to 101. Show, more generally, that when n is large, the change in $\sqrt{x}$ as x increases from n to $n+1$ is approximately $1/(2\sqrt{n})$.

2. Express $\ln\sqrt{\dfrac{1-x}{1+x}}$ as a series of terms in ascending powers of x up to and including x^5. Use your series to obtain an approximate value for $\ln(3/\sqrt{11})$, giving your answer correct to five decimal places.

3. Expand as a power series, in terms up to x^4, the function $f(x) = \cos(\sin x)$.

4. The demand quantity q for a good is given by

$$q(1+p^2) = 100,$$

where p is the price. Determine the point elasticity of demand as a function of p. For what values of p is the demand inelastic?

5. The supply and demand functions for a good are

$$q^S(p) = bp - a, \quad q^D(p) = c - dp,$$

where a, b, c, d are all positive, and $bc > ad$. Suppose the government wishes to raise as much money as possible by imposing an excise tax on the good. What should be the value of the excise tax? What is the resulting government revenue?

6. Find the area enclosed by the curves $y = 1/t^2$, $y = t^3$, the t-axis and the lines $t = 1/2$ and $t = 2$.

7. The demand relationship for a product is

$$p = \frac{50}{q+5}$$

and the supply relationship is

$$p = \frac{q}{10} + 4.5.$$

On the same diagram, produce graphs of the supply and demand curves. Determine the consumer and producer surpluses.

8. Suppose that the demand and supply functions for a good are

$$q^S(p) = bp - a, \quad q^D(p) = c - dp,$$

where a, b, c, d are positive constants. Find an expression for the consumer surplus.

Answers to activities

2.1 The demand function and its derivative are

$$q^D(p) = 6000p^{-2}, \quad q^{D'}(p) = 6000 \times (-2)p^{-3} = \frac{-12000}{p^3}.$$

Therefore when $p = 20$ and $\Delta p = 1$ we have

$$\Delta q \simeq (-12000/p^3)\Delta p = -12000/20^3 = -1.5.$$

So the price rise will result in approximately 1500 fewer items being sold.

2.2 You should obtain 0.995458. This compares well with the true value of 0.995461. (Of course, this needs a calculator, but I stress again that such calculations will not be required in the examinations, in which the use of calculators is not permitted.)

2.3 The corresponding new equilibrium quantity is

$$q^T = 40 - 5p^T = 40 - 5(4 + (3/5)T) = 20 - 3T.$$

2.4 The required area is the definite integral $\int_0^\pi \sin x \, dx$. This is calculated as follows:

$$\int_0^\pi \sin x \, dx = [-\cos x]_0^\pi = -\cos(\pi) - (-\cos(0)) = -(-1) + 1 = 2.$$

2.5 The inverse demand function is $p^D(q) = 231/(q+1)$. The equilibrium quantity q^* is the solution to the equation

$$\frac{231}{q+1} = q + 11,$$

obtained by equating $p^D(q)$ and $p^S(q)$. So, q^* satisfies the equation

$$(q+11)(q+1) = 231, \quad q^2 + 12q - 220 = 0, \quad (q+22)(q-10) = 0.$$

Since q^* cannot be negative, $q^* = 10$. The equilibrium price is $p^* = q^* + 11 = 21$. The consumer surplus is then

$$\begin{aligned} CS &= \int_0^{q^*} p^D(q)\,dq - p^*q^* \\ &= \int_0^{10} \frac{231}{q+1}\,dq - (21)(10) \\ &= [231\ln(q+1)]_0^{10} - 210 \\ &= 231\ln(11) - 231\ln(1) - 210 \\ &= 231\ln(11) - 210, \end{aligned}$$

which is approximately 343.9.

2.6 We have

$$\begin{aligned} \int_1^7 x\ln x\,dx &= \left[\frac{x^2}{2}\ln x\right]_1^7 - \int_1^7 \frac{x^2}{2}\frac{1}{x}\,dx \\ &= \left[\frac{x^2}{2}\ln x\right]_1^7 - \left[\frac{x^2}{4}\right]_1^7 \\ &= \frac{49}{2}\ln 7 - \frac{49}{4} + \frac{1}{4} = 35.675. \end{aligned}$$

Thus the approximate answer from Simpson's rule is indeed close to the true answer.

Answers to sample examination/practice questions

1. The derivative of $f(x) = \sqrt{x}$ is $f'(x) = 1/(2\sqrt{x})$. Therefore, by the approximation

$$\Delta f \simeq \frac{df}{dx}\Delta x,$$

we have

$$f(101) - f(100) \simeq f'(100)(1) = \frac{1}{2\sqrt{100}} = \frac{1}{20}.$$

The approximate change is therefore $1/20 = 0.05$. (Incidentally, the exact change is $\sqrt{101} - \sqrt{100}$, which is 0.04987562, so the approximation is good.)

Generally, if n is large and x increases from n to $n+1$, then the change is given approximately by

$$f(n+1) - f(n) \simeq f'(n)((n+1) - n) = \frac{1}{2\sqrt{n}},$$

as required.

2. We first note that

$$\ln\sqrt{\frac{1-x}{1+x}} = \frac{1}{2}\ln(1-x) - \frac{1}{2}\ln(1+x).$$

Now we use the following expansions:

$$\ln(1-x) = -x - \frac{x^2}{2} - \frac{x^3}{3} - \frac{x^4}{4} - \frac{x^5}{5},$$

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \frac{x^5}{5},$$

where we have omitted powers of x higher than x^6 (noting that the question requires us to work with powers up to x^5 only). It follows that we have the following Taylor (Maclaurin) series:

$$\begin{aligned}
\ln\sqrt{\frac{1-x}{1+x}} &\simeq \frac{1}{2}\left(-x - \frac{x^2}{2} - \frac{x^3}{3} - \frac{x^4}{4} - \frac{x^5}{5}\right) - \frac{1}{2}\left(x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \frac{x^5}{5}\right) \\
&= -x - \frac{x^3}{3} - \frac{x^5}{5}.
\end{aligned}$$

For the next part of the question, it would appear that we have to express $3/\sqrt{11}$ in the form $\sqrt{(1-x)/(1+x)}$. For this to be so, we need

$$(1-x)/(1+x) = (3/\sqrt{11})^2 = 9/11,$$

so $11(1-x) = 9(1+x)$, or $20x = 2$, so we take $x = 0.1$. Then,

$$\begin{aligned}
\ln\left(\frac{3}{\sqrt{11}}\right) &\simeq -(0.1) - \frac{(0.1)^3}{3} - \frac{(0.1)^5}{5} \\
&= -0.1 - \frac{0.001}{3} - \frac{0.00001}{5} \\
&= -0.1 - 0.000333 - 0.000002 = -0.100335.
\end{aligned}$$

(Incidentally, the true value is $-.1003353480$, so the approximation is correct to five decimal places.)

3. The series for $\sin x$ is

$$\sin x \simeq x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots,$$

but since we are only interested in powers of x up to x^4, we shall use the approximation

$$\sin x \simeq x - \frac{x^3}{3!} = x - \frac{x^3}{6}.$$

The series for $\cos y$ is

$$\cos y \simeq 1 - \frac{y^2}{2!} + \frac{y^4}{4!} - \frac{y^6}{6!} + \cdots.$$

To get an approximation to $\cos(\sin x)$ up to terms in x^4, we substitute

$$y = x - \frac{x^3}{6}$$

in the expansion for $\cos y$ and ignore any powers of x higher than x^4. (Ignored terms will be denoted by $\cdots$.) We get

$$\begin{aligned}
\cos(\sin x) &\simeq 1 - \frac{1}{2}\left(x - \frac{x^3}{6}\right)^2 + \frac{1}{24}\left(x - \frac{x^3}{6}\right)^4 + \cdots \\
&= 1 - \frac{x^2}{2}\left(1 - \frac{x^2}{6}\right)^2 + \frac{x^4}{24}\left(1 - \frac{x^2}{6}\right)^4 + \cdots \\
&= 1 - \frac{x^2}{2}\left(1 - \frac{x^2}{3} + \frac{x^4}{36}\right) + \frac{x^4}{24}(1 + \cdots) + \cdots \\
&= 1 - \frac{x^2}{2} + \frac{x^4}{6} + \frac{x^4}{24} + \cdots \\
&= 1 - \frac{1}{2}x^2 + \frac{5}{24}x^4 + \cdots.
\end{aligned}$$

Be careful about what can be ignored and what cannot. Consider, for example, the term

$$\frac{x^4}{2}\left(1-\frac{x^2}{3}+\frac{x^4}{36}\right).$$

In expanding this, since we are ignoring powers of x higher than 4, we can ignore all the terms inside the brackets apart from the first, since they lead to a power of x at least as high as x^6. But you also have to be careful to include every term that needs to be included.

4. We have

$$q = 100/(1+p^2) = 100(1+p^2)^{-1},$$

so

$$\frac{dq}{dp} = -100(2p)(1+p^2)^{-2} = -200p(1+p^2)^{-2},$$

and the elasticity of demand is

$$\begin{aligned}
\varepsilon &= -\frac{p}{q}\frac{dq}{dp} \\
&= -\frac{p}{q}\left(\frac{-200p}{(1+p^2)^2}\right) \\
&= \frac{200p^2}{q(1+p^2)^2} \\
&= \frac{200p^2}{(100/(1+p^2))\,(1+p^2)^2} \\
&= \frac{2p^2}{1+p^2}.
\end{aligned}$$

Note that in order to obtain the elasticity as a function of p, we needed to express q in terms of p in the second-last of these equations. Now, the demand is elastic when $\varepsilon > 1$, which means $2p^2/(1+p^2) > 1$, or $2p^2 > 1+p^2$. So we have elastic demand when $p^2 > 1$ which, since $p \geq 0$, means when $p > 1$.

5. The tax revenue is $R(T) = Tq^T$. In order to calculate q^T, we first note than when the excise tax is imposed, the selling price at equilibrium, p^T, is such that

$$q^T = b(p^T - T) - a = c - dp^T.$$

Solving for p^T, we obtain

$$p^T = \frac{c+a}{b+d} + \frac{bT}{b+d}.$$

Then

$$q^T = c - dp^T = \frac{bc-ad}{b+d} - \frac{bdT}{b+d},$$

so that

$$R(T) = \left(\frac{bc-ad}{b+d}\right)T - \left(\frac{bd}{b+d}\right)T^2.$$

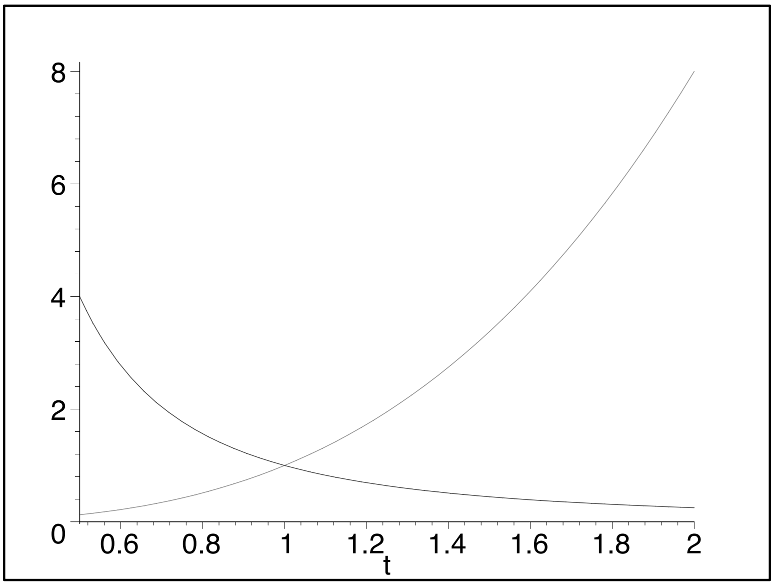
Setting $R'(T) = 0$, we discover that there is only one critical point,

$$T_m = \frac{bc-ad}{2bd}.$$

The second derivative $R''(T) = (-2bd)/(b+d)$ is constant, and negative (since b and d are positive). Hence $R''(T_m) < 0$ and T_m is a maximum point. Therefore, T_m is the level of excise tax the government should impose. The resulting government revenue is

$$R(T_m) = \frac{(bc-ad)^2}{4bd(b+d)}.$$

6. It is a good idea to sketch the region described. Here's what it looks like (where the part of the t axis shown is from $t = 1/2$ to $t = 2$).



Note that the curves $y = 1/t^2$ and $y = t^3$ intersect when $t^5 = 1$ which, in the positive quadrant, means $t = 1$. The region we're interested in then divides naturally into two parts. The first is for t from $1/2$ to 1, where the curve $1/t^2$ lies above t^3 and the region is therefore bounded by $t = 1/2$, $t = 1$ and $y = t^3$. The second part is for t ranging from $1/2$ to 1: here, the curve $y = t^3$ lies above $y = 1/t^2$, and so the region is bounded by $t = 1$, $t = 2$ and $y = 1/t^2$. The easiest way to compute the area, A, of the region is to calculate each of the areas of the two parts, which we'll call A_1 and A_2, separately. Then $A = A_1 + A_2$. We have

$$A_1 = \int_{1/2}^{1} t^3 \, dt = \left[\frac{t^4}{4}\right]_{1/2}^{1} = \frac{1}{4}(1)^4 - \frac{1}{4}\left(\frac{1}{2}\right)^4 = \frac{15}{64},$$

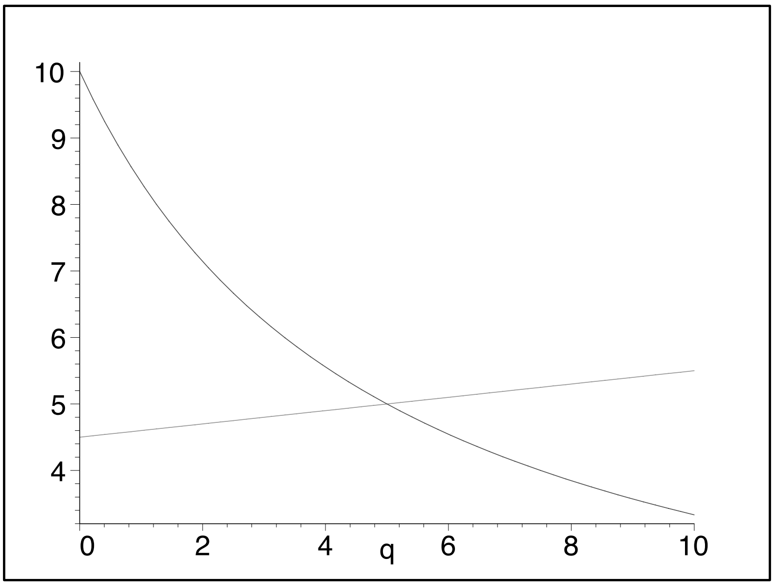
and

$$A_2 = \int_1^2 \frac{1}{t^2} \, dt = \left[-\frac{1}{t}\right]_1^2 = -\frac{1}{2} - \left(-\frac{1}{1}\right) = \frac{1}{2}.$$

It follows that

$$A = A_1 + A_2 = \frac{47}{64}.$$

7. The supply and demand curves are as follows. Note that the supply curve is an upward-sloping straight line.



We need to find the equilibrium price and quantity. To find the equilibrium quantity, we solve

$$\frac{50}{q+5} = \frac{q}{10} + 4.5.$$

Multiplying both sides by $10(q+5)$, we obtain

$$500 = q(q+5) + 45(q+5) = q^2 + 50q + 225,$$

so

$$q^2 + 50q - 275 = 0.$$

The solutions are 5 and -55 and clearly it is the positive solution we're interested in. So the equilibrium quantity is 5 and the equilibrium price is $50/(5+5) = 5$.

Now, the consumer surplus is

$$\begin{aligned} CS &= \int_0^5 \frac{50}{q+5}\,dq - (5)(5) \\ &= [50\ln(q+5)]_0^5 - 25 \\ &= 50\ln 10 - 25. \end{aligned}$$

The producer surplus is easy to calculate, since it is the area of the triangular region bounded by the p-axis, the line $p = 5$ and the portion of the supply curve between $q = 0$ and $q = 5$. Since the supply curve intersects the p-axis at 4.5, this triangle has height 0.5 and base 5, so its area is $(1/2)(0.5)(5)$, which is 1.25. So the producer surplus is 1.25.

8. We first calculate the consumer surplus by elementary methods, using only the fact that the area of a triangle is half its base times its height. As is easily seen, the equilibrium price is

$$p^* = \frac{c+a}{b+d},$$

and the equilibrium quantity is

$$q^* = c - dp^* = \frac{bc - ad}{b+d}.$$

The inverse demand function is

$$p^D(q) = \frac{c-q}{d}.$$

The consumer surplus is the area of the triangular region bounded by the lines $p = p^*$ and $q = 0$, and by the demand curve. (Sketch the curves to see that this is so! The fact that the supply demand curve is a straight line means the region in question is triangular.) Since the demand curve crosses the p-axis at $(0, p^D(0))$, this area is

$$CS = \frac{1}{2}(p^D(0) - p^*)q^* = \frac{1}{2}\left(\frac{c}{d} - \frac{c+a}{b+d}\right)\left(\frac{bc-ad}{b+d}\right).$$

That is,

$$CS = \frac{1}{2}\left(\frac{cb+cd-cd-ad}{d(b+d)}\right)\left(\frac{bc-ad}{b+d}\right) = \frac{(bc-ad)^2}{2d(b+d)^2}.$$

We can also calculate the consumer surplus by definite integration. (This turns out to be more difficult in this particular case, but definite integration really is needed for demand functions more complex than the simple linear one discussed here.) We have

$$CS = \int_0^{q^*} p^D(q)\,dq - p^*q^* = \int_0^{q^*} \frac{c-q}{d}\,dq - p^*q^*.$$

Now,

$$\int_0^{q^*} \frac{c-q}{d}\,dq = \int_0^{q^*} \left(\frac{c}{d} - \frac{1}{d}q\right)\,dq = \left[\frac{c}{d}q - \frac{1}{2d}q^2\right]_0^{q^*} = \frac{c}{d}q^* - \frac{1}{2d}(q^*)^2,$$

so

$$\begin{aligned}
CS &= \left(\frac{c}{d}q^* - \frac{1}{2d}(q^*)^2\right) - p^*q^* \\
&= \frac{q^*}{2d}\left(2c - q^* - 2p^*d\right) \\
&= \frac{q^*}{2d}\left(2c - \frac{(bc-ad)}{b+d} - 2\frac{(c+a)}{b+d}d\right) \\
&= \frac{q^*}{2d(b+d)}\left(2cb + 2cd - bc + ad - 2cd - 2ad\right) \\
&= \frac{(bc-ad)}{2d(b+d)^2}\left(bc-ad\right) \\
&= \frac{(bc-ad)^2}{2d(b+d)^2},
\end{aligned}$$

as above.

Chapter 3

Functions of several variables

Essential reading

(For full publication details, see Chapter 1.)

Anthony, M. and Biggs, N. *Mathematics for Economics and Finance.* Chapters 11, 12, 13, 21, and 22.

Further reading

Binmore, K. and Davies, J, *Calculus.* Sections 6.6 and 6.8.

Black, J. and Bradley, J.F. *Essential Mathematics for Economists.* Chapters 7 and 8.

Bradley, T., and Patton, P. *Essential Mathematics for Economics and Business.* Chapter 7.

Dowling, Edward T. *Introduction to Mathematical Economics.* Second Edition. Chapters 5 and 6.

Holden, K, and Pearson, A.W., *Introductory Mathematics for Economics and Business.* Chapter 7.

Ostaszewski, A. *Mathematics in Economics: Models and Methods.* Chapters 12 and 15.

Introduction

In this chapter of the guide, we briefly review the material from Mathematics 1 concerning functions of several variables, and we introduce some new topics and applications related to functions of several variables.

Partial derivatives

We briefly review some key ideas from Mathematics 1 (though in the slightly more general context of n-variable functions rather than 2-variable functions). A **function of n variables**, for $n \geq 2$, takes inputs $(x_1, x_2, \ldots, x_n)$ and returns an output value $f(x_1, x_2, \ldots, x_n)$. In Mathematics 1, we concentrated almost entirely on the case in which $n = 2$, and we would often use x, y to denote the variables, rather than x_1, x_2. In this chapter we shall sometimes consider specific examples where $n > 2$. In particular, when $n = 3$, we shall often use x, y, z to represent the three variables rather than x_1, x_2, x_3.

For a function $f(x_1, x_2, \ldots, x_n)$, the rate of change of f with respect to x_1, when the other variables are fixed, is the **partial derivative** of f with respect to x_1, denoted $\dfrac{\partial f}{\partial x_1}$. The partial derivatives with respect to the other variables are similarly defined.

As we saw in Mathematics 1, calculating partial derivatives is only slightly more difficult than calculating standard derivatives, and you should be proficient in this from your study of Mathematics 1. For example, to calculate the partial derivative of a function $f(x, y, z)$ with respect to x, you just treat y and z both as if they were fixed numbers, and differentiate with respect to x. Here is an example, to refresh your memory, and to indicate how to deal with 3-variable functions.

Example: Suppose $f(x, y, z) = x^2 y \sqrt{y^2 + z^2}$. Writing this in the form $x^2 y (y^2 + z^2)^{1/2}$ makes it slightly easier to work with. We regard this as a product of two functions, where the second one is a composition. Applying the product and chain rules appropriately, we have

$$\frac{\partial f}{\partial x} = 2xy(y^2 + z^2)^{1/2} = 2xy\sqrt{y^2 + z^2},$$

$$\frac{\partial f}{\partial y} = x^2(y^2 + z^2)^{1/2} + x^2 y(2y)\left(\frac{1}{2}\right)(y^2 + z^2)^{-1/2} = x^2\sqrt{y^2 + z^2} + \frac{x^2 y^2}{\sqrt{y^2 + z^2}},$$

$$\frac{\partial f}{\partial z} = x^2 y(2z)\left(\frac{1}{2}\right)(y^2 + z^2)^{-1/2} = \frac{x^2 yz}{\sqrt{y^2 + z^2}}.$$

Homogeneous functions and Euler's theorem

Suppose we have a function of two variables, such as $f(x, y) = 3x^2 y + 7xy^2$, and we multiply the 'inputs' x and y by a constant c. In this case we get 'output'

$$f(cx, cy) = 3(cx)^2(cy) + 7(cx)(cy)^2 = c^3(3x^2 y + 7xy^2) = c^3 f(x, y).$$

Thus, for this particular f, multiplying the inputs by c results in the 'output' being multiplied by c^3. In general, if a function h is such that

$$h(cx, cy) = c^D h(x, y),$$

then we say that h is **homogeneous of degree** D. The number D is called the **degree of homogeneity** of h. The function f given by the formula above is homogeneous of degree 3. **Note that many (indeed most) functions are not homogeneous.** The notion of a homogeneous function is related to the idea of 'returns to scale' in economics. In the case of a production function, we say that there are **constant returns to scale** if a proportional increase in k and l results in the same proportional increase in $q(k,l)$; that is, if

$$q(ck,cl) = cq(k,l).$$

This means, for example, that doubling both capital and labour doubles the production. This is the same as saying that q is homogeneous of degree 1. If q is homogeneous of degree $D > 1$ then the proportional increase in $q(k,l)$ will be larger than that in k and l, and we say that there are **increasing returns to scale**. On the other hand, if $D < 1$ we say that there are **decreasing returns to scale**.

Example: Consider the function $f(x,y) = x^2y^2 + y^2\sqrt{x^4+y^4}$. Then f is homogeneous of degree 4 because

$$\begin{aligned}
f(cx,cy) &= (cx)^2(cy)^2 + (cy)^2\sqrt{(cx)^4+(cy)^4} \\
&= c^2x^2c^2y^2 + c^2y^2\sqrt{c^4(x^4+y^4)} \\
&= c^4x^2y^2 + c^2y^2c^2\sqrt{x^4+y^4} \\
&= c^4x^2y^2 + c^4y^2\sqrt{x^4+y^4} \\
&= c^4 f(x,y).
\end{aligned}$$

Informally speaking, a function such as the one just considered, which involves only powers of the variables and powers of combinations of the variables, is homogeneous if the total 'degree' or 'power' of each constituent term of the function is the same. In the example we just saw, the total degree of the first part of the function, x^2y^2 is $2+2=4$, the sum of the degree of x and the degree of y. The degree of the second term is $2+(1/2)4=4$ (where the factor $1/2$ comes from the square root). Since these are the same, and both equal to 4, the function is homogeneous of degree 4. To see why this informal approach holds, just think about what happens when we substitute cx for x and cy for y. Perhaps it would also be useful to see why a function with parts of differing degrees cannot be homogeneous. Consider the function $f(x,y) = x+y^2$. (This has terms of differing degree 1 and 2.) Then $f(cx,cy) = cx + c^2y^2$. But there is no way in which $cx + c^2y^2$ can be written as $c^D(x+y^2)$, for all values of c, for some D.

Activity 3.1 *Which of the following functions are homogeneous? If homogeneous, what are their degrees?*

$$x^2y + \frac{x^4}{\sqrt{x^2+y^2}}, \quad \frac{x}{y} + 5 + e^{x^2/y^2}, \quad x^2 + x\sin y.$$

Example: Suppose we are told that the function

$$f(x,y) = \frac{x^\alpha y + x^2y^\beta}{xy^\gamma + y^2}$$

is homogeneous of degree 2. Then we can use this fact to determine the numbers α, β and γ. There are several ways to do this. Think about what happens when x, y are replaced by cx, cy. We have

$$f(cx,cy) = \frac{c^{\alpha+1}x^\alpha y + c^{2+\beta}x^2y^\beta}{c^{1+\gamma}xy^\gamma + c^2y^2}.$$

Now, if the function is homogeneous of degree 2, this should simply equal $c^2f(x,y)$. But for this to be true, three things must hold:

1. The degree of each of the two terms on the numerator (top line) should equal each other; let's say this common degree is d_1
2. The degree of each of the two terms on the denominator (bottom line) should equal each other; let's say this degree is d_2
3. The degrees d_1 and d_2 should cancel to give degree 2; that is, the c^{d_1} factor from the top and the c^{d_2} factor on the bottom should cancel to give c^2. This means that $d_1 - d_2 = 2$.

These conditions are equivalent to saying that

- The numerator must be homogeneous
- The denominator must be homogeneous
- The degree of the numerator must be 2 more than the degree of the denominator

Respectively, these three conditions mean:

1. $\alpha + 1 = 2 + \beta$
2. $1 + \gamma = 2$
3. We can write four possible equations here. For example, since the degree of the first term on the numerator is $\alpha + 1$ and the degree of the first term on the denominator is $1 + \gamma$, we have $\alpha + 1 - (1 + \gamma) = 2$. By considering other terms (such as the second term on the numerator, and the first on the denominator, and so on), we obtain three other equations: $\alpha + 1 - 2 = 2$, $2 + \beta - (1 + \gamma) = 2$, and $2 + \beta - 2 = 2$. But by the first and second observations above, these four equations are all equivalent and we need only use one of them.

Choosing the second of the four equations in 3., (because it is simple), we obtain $\alpha = 3$. Then the equation in 1. tells us that $\beta = 2$, and the equation in 2. tells us that $\gamma = 1$. So the function is

$$f(x, y) = \frac{x^3y + x^2y^2}{xy + y^2}.$$

There is a useful result about homogeneous functions known as **Euler's Theorem**. This states that a function $f(x, y)$ is homogeneous of degree D if and only if

$$x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} \;=\; Df.$$

Example: Consider the function $f(x, y) = (x^2 + y^2)^{3/2}x^{1/2}y^{1/2}$. We can show that this is homogeneous of degree 4, and verify that

$$x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} = 4f.$$

To show that the function is homogeneous of degree 4, we observe that

$$\begin{aligned} f(cx, cy) &= ((cx)^2 + (cy)^2)^{3/2}(cx)^{1/2}(cy)^{1/2} \\ &= (c^2)^{3/2}(x^2 + y^2))^{3/2}c^{1/2}x^{1/2}c^{1/2}y^{1/2} \\ &= c^4 f(x, y). \end{aligned}$$

Now,

$$\frac{\partial f}{\partial x} = \frac{3}{2}2x(x^2 + y^2)^{1/2}(x^{1/2}y^{1/2}) + (x^2 + y^2)^{3/2}\left(\frac{1}{2}x^{-1/2}y^{1/2}\right).$$

$$\frac{\partial f}{\partial y} = \frac{3}{2}2y(x^2 + y^2)^{1/2}(x^{1/2}y^{1/2}) + (x^2 + y^2)^{3/2}\left(\frac{1}{2}x^{1/2}y^{-1/2}\right).$$

It follows that

$$\begin{aligned} x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} &= (3x^2 + 3y^2)(x^2 + y^2)^{1/2}(x^{1/2}y^{1/2}) + (x^{1/2}y^{1/2})(x^2 + y^2)^{3/2} \\ &= 3(x^2 + y^2)^{3/2}(x^{1/2}y^{1/2}) + (x^2 + y^2)^{3/2}(x^{1/2}y^{1/2}) \\ &= 4(x^2 + y^2)^{3/2}(x^{1/2}y^{1/2}) = 4f(x, y), \end{aligned}$$

as predicted by Euler's Theorem.

Activity 3.2 *That was a difficult example. Try an easier one for yourself, by verifying Euler's theorem for the function* $f(x, y) = x^2y^3 + x^4y$.

More generally, a function $f(x_1, x_2, \ldots, x_n)$ of n variables is said to be homogeneous of degree D if for all c,

$$f(cx_1, cx_2, \ldots, cx_n) = c^D f(x_1, x_2, \ldots, x_n).$$

Euler's theorem in this case is: $f(x_1, x_2, \ldots, x_n)$ is homogeneous of degree D if and only if

$$x_1\frac{\partial f}{\partial x_1} + x_2\frac{\partial f}{\partial x_2} + \cdots + x_n\frac{\partial f}{\partial x_n} = Df(x_1, x_2, \ldots, x_n).$$

Optimisation

Introduction

We now briefly review the optimisation techniques we met in Mathematics 1, and again we shall describe them a little more generally, in the context of n-variable functions. For this subject, Mathematics 2, we will look at constrained optimisation of three-variable—rather than simply two-variable—functions. For unconstrained optimisation of such functions, there are techniques for testing whether a critical point is a maximum, a minimum, or a saddle point. However, for functions of more than two variables, these techniques are outside the scope of this subject. Also not covered in this subject are tests for ensuring that the solution to a constrained optimisation problem is a maximum or a minimum.

Unconstrained optimisation

The local maxima and minima of a function $f(x_1, x_2, \ldots, x_n)$ occur at points where all the partial derivatives $\partial f / \partial x_i$ (for $i = 1, 2, \ldots, n$) are equal to 0. Such points are called **critical points** or **stationary points**. A critical point which is neither a local maximum nor a local minimum is a **saddle point**. If $n = 2$, then there is a fairly simple test to determine the nature of a critical point: Suppose that (a, b) is a critical point of $f(x, y)$. Then the following statements are true, where the partial derivatives are evaluated at (a, b):

- If $\dfrac{\partial^2 f}{\partial x^2}\dfrac{\partial^2 f}{\partial y^2} - \left(\dfrac{\partial^2 f}{\partial x \partial y}\right)^2 > 0$ and $\dfrac{\partial^2 f}{\partial x^2} < 0$, it is a maximum.
- If $\dfrac{\partial^2 f}{\partial x^2}\dfrac{\partial^2 f}{\partial y^2} - \left(\dfrac{\partial^2 f}{\partial x \partial y}\right)^2 > 0$ and $\dfrac{\partial^2 f}{\partial x^2} > 0$, it is a minimum.
- If $\dfrac{\partial^2 f}{\partial x^2}\dfrac{\partial^2 f}{\partial y^2} - \left(\dfrac{\partial^2 f}{\partial x \partial y}\right)^2 < 0$, it is a saddle point.

Applications of unconstrained optimisation

In Mathematics 1, we described the problem of maximising profit for a firm making two products, X and Y. Generally, if p_X and p_Y are the selling prices of one unit of X and one unit of Y, then the total revenue obtained by producing amounts x and y is

$$TR(x, y) = xp_X + yp_Y.$$

The **joint total cost function** $TC(x, y)$ will tell us how much it costs the manufacturer to produce x units of X and y of Y. Then, the profit function is

$$\Pi(x, y) = TR(x, y) - TC(x, y) = xp_X + yp_Y - TC(x, y)$$

and we maximise this function of x and y using the techniques described above. [1]

[1] *See Anthony and Biggs, Chapter 13, for many examples*

There are many other interesting applications of unconstrained optimisation in Economics. We now consider the case of a firm producing just one good, but selling that good in two markets, such as the domestic market and the export market.

Let us suppose that the firm is a monopoly in each of the two markets, and that the demand curves for its product are different in the two markets. The firm may decide to set different prices in the domestic market and in the export market, or it may set the same price in both. In the former case, we say that we have **price discrimination**. The following example demonstrates how to deal with such problems.

Example: Suppose the demand functions for a firm's domestic and foreign markets are given by

$$\begin{aligned} P_1 &= 30 - 4Q_1 \\ P_2 &= 50 - 5Q_2. \end{aligned}$$

and the total cost function is

$$TC = 10 + 10Q$$

where $Q = Q_1 + Q_2$. We shall determine the prices which maximise profit when we have price discrimination, and when we have no price discrimination. First, we note that the total revenue is $P_1Q_1 + P_2Q_2$, and hence the profit function is

$$\begin{aligned}\Pi(Q_1, Q_2) &= (P_1Q_1 + P_2Q_2) - (10 + 10Q) \\ &= (30 - 4Q_1)Q_1 + (50 - 5Q_2)Q_2 - (10 + 10Q_1 + 10Q_2) \\ &= 20Q_1 + 40Q_2 - 4Q_1^2 - 5Q_2^2 - 10.\end{aligned}$$

To find the maximum, we solve

$$\begin{aligned}\frac{\partial \Pi}{\partial Q_1} = 20 - 8Q_1 &= 0 \\ \frac{\partial \Pi}{\partial Q_2} = 40 - 10Q_2 &= 0,\end{aligned}$$

obtaining $Q_1 = 5/2$ and $Q_2 = 4$, and hence $P_1 = 20$, $P_2 = 30$. We can see that this is indeed a maximum of profit by using the second derivative test (check this!). The maximum profit with price discrimination is therefore $\Pi(5/2, 4) = 95$. (Here, we have found an expression for profit in terms of Q_1 and Q_2. An alternative approach would be to determine Π as a function of P_1 and P_2 and maximise the resulting function of P_1 and P_2.) Now, when there is no price discrimination, we have $P_1 = P_2 = P$, say. From the demand equations, we have

$$P = 30 - 4Q_1 = 50 - 5Q_2,$$

so

$$Q_1 = \frac{15}{2} - \frac{P}{4}, \quad Q_2 = 10 - \frac{P}{5},$$

and

$$Q = Q_1 + Q_2 = \frac{35}{2} - \frac{9P}{20}.$$

It follows that the profit, as a function of P, is given by

$$\begin{aligned}\Pi(P) &= P(Q_1 + Q_2) - (10Q + 10) \\ &= PQ - 10Q - 10 \\ &= P\left(\frac{35}{2} - \frac{9P}{20}\right) - 10\left(\frac{35}{2} - \frac{9P}{20}\right) - 10 \\ &= 22P - \frac{9}{20}P^2 - 185.\end{aligned}$$

This is a function of the single variable P. (We could just as easily have expressed Π in terms of the single variable Q: both approaches are fine.) To maximise, we set $d\Pi/dP = 0$. This yields

$$22 - \frac{9}{10}P = 0,$$

so $P = 220/9$. This gives a maximum of the profit function, since $d^2\Pi/dP^2 = -(9/10) < 0$. The corresponding profit is $\Pi(220/9) = 755/9$, which is 83.88 to two decimal places. This is, as we would expect, less than the maximum profit when price discrimination is allowed.

It is certainly possible to use Lagrangean methods, with the constraint $P_1 - P_2 = 0$ expressed in terms of Q_1 and Q_2, to deal with the case of no price discrimination. However, it is probably easier to take the approach indicated in the example just given, where the equation $P_1 = P_2$ is used to reduce the optimisation problem to a one-variable problem.

Activity 3.3 *Follow up the comment just made by using the Lagrangean method for the last part of the example, in which there is no price discrimination.*

Constrained optimisation

Suppose that $f(x_1, x_2, \ldots, x_n)$ has to be minimised or maximised **subject to the constraint** $g(x_1, x_2, \ldots, x_n) = 0$. This means we want to find the maximum (or minimum) value of the function f at points $(x_1, x_2, \ldots, x_n)$ which satisfy the condition $g(x_1, x_2, \ldots, x_n) = 0$. Then we may use the method of **Lagrange multipliers**.[2] To find the optimal points, we first find the critical points of the $(n+1)$-variable function

$$L(x_1, x_2, \ldots, x_n, \lambda) = f(x_1, x_2, \ldots, x_n) - \lambda g(x_1, x_2, \ldots, x_n).$$

[2] *See Anthony and Biggs, Chapters 21 and 22.*

The function L is known as the **Lagrangean** (sometimes spelt Lagrangian) and λ is known as the **Lagrange multiplier**. In other words, we find the points at which the **first-order conditions**

$$\frac{\partial L}{\partial x_i} = 0 \text{ (for } i = 1, 2, \ldots, n), \quad \frac{\partial L}{\partial \lambda} = 0.$$

Then the theory of Lagrange multipliers asserts that the required optimal points of f, subject to the constraint, are to be found among these critical points.

In Mathematics 1, you studied the Lagrangean method quite extensively for two-variable functions. Here, we don't introduce much new material on this subject, but for Mathematics 2, you should be able to deal with more challenging 2-variable constrained optimisation problems, and with Lagrangean problems involving more than two variables. The following example serves as an illustration of the technique for such functions.

Example: Use Lagrange Multipliers to find the minimum value of

$$\frac{1}{2x^2} + \frac{1}{3y^2} + \frac{1}{6z^2}$$

for $x, y, z > 0$ subject to

$$\frac{1}{2}x + \frac{1}{3}y + \frac{1}{6}z = c,$$

where $c > 0$ is a constant.

To solve this, we form the Lagrangean

$$L = \frac{1}{2x^2} + \frac{1}{3y^2} + \frac{1}{6z^2} - \lambda\left(\frac{1}{2}x + \frac{1}{3}y + \frac{1}{6}z - c\right)$$

and we solve the equations

$$\begin{aligned}
\frac{\partial L}{\partial x} &= -\frac{1}{4x^3} - \frac{1}{2}\lambda = 0 \\
\frac{\partial L}{\partial y} &= -\frac{1}{6y^3} - \frac{1}{3}\lambda = 0 \\
\frac{\partial L}{\partial z} &= -\frac{1}{12z^3} - \frac{1}{6}\lambda = 0 \\
\frac{\partial L}{\partial \lambda} &= -\frac{1}{2}x - \frac{1}{3}y - \frac{1}{6}z + c = 0
\end{aligned}$$

The first three equations give three expressions for λ:

$$\lambda = -\frac{1}{2x^3} = -\frac{1}{2y^3} = -\frac{1}{2z^3},$$

from which it follows that $x^3 = y^3 = z^3$ and hence $x = y = z$. Then, setting $y = z = x$ in the final equation (which is simply the constraint) gives

$$\frac{1}{2}x + \frac{1}{3}x + \frac{1}{6}x = c,$$

which is $x = c$. Therefore the optimum (in this case minimum) is when $x = y = z = c$, and the optimum value is

$$\frac{1}{2c^2} + \frac{1}{3c^2} + \frac{1}{6c^2} = \frac{1}{c^2}.$$

Learning outcomes

At the end of this chapter and the relevant reading, you should be able to:

- as in Mathematics 1, understand the concept of a function of many variables
- as in Mathematics 1, calculate partial derivatives, and use implicit differentiation
- state precisely what is meant by a homogeneous function of n variables, and be able to show that a given function is homogeneous
- state Euler's theorem, and be able to verify Euler's equation for given homogeneous functions
- as in Mathematics 1, find and classify critical points
- as in Mathematics 1, solve optimisation and constrained optimisation problems: additionally, solve constrained optimisation problems with three or more variables
- as in Mathematics 1, explain, and make use of, the meaning of the Lagrange multiplier

Sample examination/practice questions

1. The production function $Q(K, L)$ of a firm is given by

$$Q(K, L) = \frac{5L^{\alpha}K^{\beta}}{L^2 + K^{2\alpha}L}.$$

If the firm has constant returns to scale, what are the values of α and β?

2. If the following function is homogeneous of degree 1 determine the values of α, β, γ:

$$g(x, y, z) = \frac{3x^{\gamma}y^{\beta} - 4x^{\alpha+3\gamma}}{y^{3/2}\left(x^{\beta+\gamma} + 2z^{2\alpha+\gamma}\right)^{1/2}}.$$

3. Show that the function

$$f(x, y) = 2y^3x + 5y^4 - (y^{3/4} - 2x^{3/4})^4x$$

is homogeneous of degree 4. Verify that

$$x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} = 4f.$$

4. A monopolist has a total cost function of the form $TC = 20Q + 20$ where Q is the demand. He is aware of the possibility of separating his customers into two distinct markets with the following demand equations:

$$\text{Market 1}: \; Q_1 = 9 - 0.05P_1,$$

$$\text{Market 2}: \; P_2 = 80 - 5Q_2,$$

where $Q_1 + Q_2 = Q$ and P_i is the price in market i $(i = 1, 2)$. If the monopolist wishes to maximise his total profits, determine the prices he will charge in the two markets (i) if he is permitted to choose different prices for the two markets, and, (ii) if he has to choose the same price for the two markets.

5. It is thought that a consumer measures the utility u of possessing a quantity x of apples and a quantity of y of oranges by the formula:

$$u = u(x, y) = x^{\alpha} y^{1-\alpha}.$$

It is known that when the consumer's budget for apples and oranges is \$1 he will buy 1 apple and 2 oranges when they are equally priced. Find α. [Hint: First solve the utility maximisation problem with α as a parameter.]

The price of apples falls to half that of oranges, with the price of oranges unchanged. How many apples and oranges will the consumer buy for \$10?

6. Company headquarters holds 100 units of a raw material; it proposes to divide this into three lots of x, y and z units of raw material to be sent to three area factories in the North, the Midlands and the South of the country and there they are turned into a finished product. Although the product sells at the same price in all three regions, the factories are not equally productive and in fact their respective outputs are:

$$\begin{aligned} \text{North } &= 3\sqrt{x} \\ \text{Midlands } &= 4\sqrt{y} \\ \text{South } &= 5\sqrt{z}. \end{aligned}$$

How should the headquarters allocate x, y, z to maximize the revenue? [Hint: Use the Lagrange multiplier method.]

7. Use the Lagrange Multiplier Method to minimize

$$\frac{1}{x} + \frac{1}{y} + \frac{1}{z}$$

subject to

$$x + y + z = c,$$

for $x, y, z > 0$, where c is a *positive* constant. Use your result to deduce that for $x, y, z > 0$

$$\frac{1}{3}(x + y + z) \geq \left\{ \frac{1}{3}\left(\frac{1}{x} + \frac{1}{y} + \frac{1}{z}\right) \right\}^{-1}.$$

Answers to activities

3.1 If $f(x,y) = x^2y + \dfrac{x^4}{\sqrt{x^2+y^2}}$, then

$$f(cx,cy) = c^3x^2y + \frac{c^4x^4}{\sqrt{c^2(x^2+y^2)}} = c^3x^2y + \frac{c^4x^4}{c\sqrt{x^2+y^2}} = c^3x^2y + c^3\frac{x^4}{\sqrt{x^2+y^2}},$$

which is $c^3f(x,y)$, so the function is homogeneous of degree 3.

If $f(x,y) = \dfrac{x}{y} + 5 + e^{x^2/y^2}$, then

$$f(cx,cy) = \frac{cx}{cy} + 5 + e^{c^2x^2/(c^2y^2)} = \frac{x}{y} + 5 + e^{x^2/y^2}.$$

In other words, $f(cx,cy) = f(x,y)$, which means that f is homogeneous of degree 0 (since $f(cx,cy) = c^0f(x,y)$).

For $f(x,y) = x^2 + x\sin y$, we have $f(cx,cy) = c^2x^2 + cx\sin(cy)$. Now, this cannot be written as $c^Df(x,y)$ for some D. For, if this were the case then, clearly, from the x^2 part, D would have to be 2. Then we'd have to have $\sin(cy) = c\sin y$ for all y, which simply isn't true. So the function is not homogeneous.

3.2 Since

$$f(cx,cy) = c^2x^2c^3y^3 + c^4x^4cy = c^5(x^2y^3 + x^4y) = c^5f(x,y),$$

the function is homogeneous of degree 5. To verify Euler's theorem, we need to check that

$$x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} = 5f.$$

Now,

$$\frac{\partial f}{\partial x} = 2xy^3 + 4x^3y, \quad \frac{\partial f}{\partial y} = 3x^2y^2 + x^4,$$

so

$$\begin{aligned} x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} &= x(2xy^3 + 4x^3y) + y(3x^2y^2 + x^4) \\ &= 2x^2y^3 + 4x^4y + 3x^2y^3 + x^4y \\ &= 5x^2y^3 + 5x^4y \\ &= 5f(x,y), \end{aligned}$$

as required.

3.3 The constraint is $P_1 = P_2$, which is

$$30 - 4Q_1 = 50 - 5Q_2,$$

so we take as the constraint equation $20 + 4Q_1 - 5Q_2 = 0$, and the Lagrangean is

$$L = 20Q_1 + 40Q_2 - 4Q_1^2 - 5Q_2^2 - 10 - \lambda(20 + 4Q_1 - 5Q_2).$$

We then solve the equations

$$\begin{aligned} \frac{\partial L}{\partial Q_1} &= 20 - 8Q_1 - 4\lambda = 0 \\ \frac{\partial L}{\partial Q_2} &= 40 - 10Q_2 + 5\lambda = 0 \\ \frac{\partial L}{\partial \lambda} &= 20 + 4Q_1 - 5Q_2 = 0. \end{aligned}$$

From the first two equations,

$$\lambda = \frac{20 - 8Q_1}{4} = \frac{10Q_2 - 40}{5},$$

so

$$100 - 40Q_1 = 40Q_2 - 160,$$

and $Q_2 = (13/2) - Q_1$. Substituting this into the third equation gives

$$20 + 4Q_1 - 5\left(\frac{13}{2} - Q_1\right) = 0,$$

which implies $Q_1 = 25/18$. The corresponding value of Q_2 is $46/9$. To find the price, we can use the equation for P_1 or P_2. (Of course, the prices P_1 and P_2 are equal, to P, say: that was our constraint.) Using the formula for P_1, we get

$$P_1 = 30 - 4\left(\frac{25}{18}\right) = \frac{220}{9},$$

as before.

Answers to sample examination/practice questions

1. To say the firm has constant returns to scale means that the function Q is homogeneous of degree 1. The total degree of the numerator is $\alpha + \beta$. Since the terms on the denominator must have the same degree as each other, we must have $2 = 2\alpha + 1$, which means $\alpha = 1/2$. Now, since Q is homogeneous of degree 1, the degree of the numerator must be one more than that of the denominator. So, $\alpha + \beta = 1 + 2$, which, since $\alpha = 1/2$, means $\beta = 5/2$.

2. Since the function is homogeneous of degree 1, we can make the following three observations:

- The numerator must be homogeneous, so $\gamma + \beta = \alpha + 3\gamma$
- The denominator must be homogeneous, so $\beta + \gamma = 2\alpha + \gamma$
- the degree of the numerator must be one more than that of the denominator. Now, because the bracketed term is raised to the power $1/2$ in the denominator, the degree of the denominator is $3/2 + (1/2)(\beta + \gamma)$ (or, equivalently, given the previous observation, $3/2 + (1/2)(2\alpha + \gamma)$). The degree of the numerator is $\gamma + \beta$ (or, equivalently, $\alpha + 3\gamma$). So we can say

  $$\gamma + \beta = 1 + \frac{3}{2} + \frac{1}{2}(2\alpha + \gamma).$$

  In fact, given the equivalences from the first two observations, there are four different equations we could write down, and all would say the same thing.)

So we have the system

$$\gamma + \beta = \alpha + 3\gamma, \;\; \beta + \gamma = 2\alpha + \gamma, \;\; \gamma + \beta = \frac{5}{2} + \alpha + \frac{1}{2}\gamma,$$

which simplifies to

$$\beta = \alpha + 2\gamma, \;\; \beta = 2\alpha, \;\; \frac{1}{2}\gamma + \beta = \frac{5}{2} + \alpha.$$

There are many ways to solve this. From the first two equations, $\alpha + 2\gamma = 2\alpha$, so $\alpha = 2\gamma$. Then, $\beta = 2\alpha = 4\gamma$. Using these in the third equation, we obtain

$$\frac{1}{2}\gamma + 4\gamma = \frac{5}{2} + 2\gamma,$$

so $\gamma = 1$ and, then, $\beta = 4$ and $\alpha = 2$.

3. We have

$$\begin{aligned} f(cx, cy) &= 2c^3y^3cx + 5c^4y^4 - \left(c^{3/4}y^{3/4} - 2c^{3/4}x^{3/4}\right)^4 cx \\ &= 2c^4y^3x + 5c^4y^4 - c^4\left(y^{3/4} - 2x^{3/4}\right)^4 x \\ &= c^4 f(x, y), \end{aligned}$$

so the function is homogeneous of degree 4.

Now,

$$\begin{aligned} \frac{\partial f}{\partial x} &= 2y^3 - \left(-2\frac{3}{4}x^{-1/4}\right)4(y^{3/4} - 2x^{3/4})^3x - (y^{3/4} - 2x^{3/4})^4 \\ &= 2y^3 + 6(y^{3/4} - 2x^{3/4})^3x^{3/4} - (y^{3/4} - 2x^{3/4})^4 \end{aligned}$$

and

$$\begin{aligned} \frac{\partial f}{\partial y} &= 6y^2x + 20y^3 - \left(\frac{3}{4}y^{-1/4}\right)4(y^{3/4} - 2x^{3/4})^3x \\ &= 6y^2x + 20y^3 - 3y^{-1/4}(y^{3/4} - 2x^{3/4})^3x. \end{aligned}$$

So,

$$x\frac{\partial f}{\partial x} = 2y^3x + 6(y^{3/4} - 2x^{3/4})^3x^{7/4} - (y^{3/4} - 2x^{3/4})^4x$$

and

$$y\frac{\partial f}{\partial y} = 6y^3x + 20y^4 - 3y^{3/4}(y^{3/4} - 2x^{3/4})^3x.$$

Therefore,

$$\begin{aligned} x\frac{\partial f}{\partial x} + y\frac{\partial f}{\partial y} &= 8y^3x + 20y^4 - (y^{3/4} - 2x^{3/4})^4x + 6(y^{3/4} - 2x^{3/4})^3x^{7/4} - 3y^{3/4}(y^{3/4} - 2x^{3/4})^3x \\ &= 8y^3x + 20y^4 - (y^{3/4} - 2x^{3/4})^4x - 3x(y^{3/4} - 2x^{3/4})^3(y^{3/4} - 2x^{3/4}) \\ &= 8y^3x + 20y^4 - (y^{3/4} - 2x^{3/4})^4x - 3x(y^{3/4} - 2x^{3/4})^4 \\ &= 8y^3x + 20y^4 - 4(y^{3/4} - 2x^{3/4})^4x \\ &= 4f(x, y), \end{aligned}$$

as required.

4. With a view to making everything a function of Q_1 and Q_2, we note first that $P_1 = 180 - 20Q_1$. The profit, as a function of Q_1 and Q_2 is then

$$\begin{aligned} \Pi(Q_1, Q_2) &= (P_1Q_1 + P_2Q_2) - (20Q + 20) \\ &= (180 - 20Q_1)Q_1 + (80 - 5Q_2)Q_2 - (20Q_1 + 20Q_2 + 20) \\ &= 160Q_1 + 60Q_2 - 20Q_1^2 - 5Q_2^2 - 20. \end{aligned}$$

To find the maximum, we solve

$$\begin{aligned} \frac{\partial \Pi}{\partial Q_1} = 160 - 40Q_1 &= 0 \\ \frac{\partial \Pi}{\partial Q_2} = 60 - 10Q_2 &= 0, \end{aligned}$$

obtaining $Q_1 = 4$ and $Q_2 = 6$, and hence $P_1 = 100$, and $P_2 = 50$. We can see that this is indeed a maximum of profit by using the second derivative test. We have

$$\frac{\partial^2 \Pi}{\partial Q_1^2} = -40, \quad \frac{\partial^2 \Pi}{\partial Q_2^2} = -10, \quad \frac{\partial^2 \Pi}{\partial Q_1 \partial Q_2} = 0,$$

so

$$\frac{\partial^2 \Pi}{\partial Q_1^2} < 0, \quad \left(\frac{\partial^2 \Pi}{\partial Q_1^2}\right)\left(\frac{\partial^2 \Pi}{\partial Q_2^2}\right) - \left(\frac{\partial^2 \Pi}{\partial Q_1 \partial Q_2}\right)^2 > 0,$$

so the critical point is indeed a maximum.

Now, when there is no price discrimination, we have $P_1 = P_2 = P$, say. From the demand equations, we have

$$Q_1 = 9 - 0.05P, \quad Q_2 = 16 - 0.2P$$

so

$$Q = Q_1 + Q_2 = 25 - 0.25P.$$

It follows that the profit, as a function of P, is given by

$$\begin{aligned}
\Pi(P) &= P(Q_1 + Q_2) - (20Q + 20) \\
&= PQ - 20Q - 20 \\
&= P(25 - 0.25P) - 20(25 - 0.25P) - 20 \\
&= 30P - 0.25P^2 - 520.
\end{aligned}$$

To maximise, we set $d\Pi/dP = 0$. This yields

$$30 - 0.5P = 0,$$

so $P = 60$. This gives a maximum of the profit function, since $d^2\Pi/dP^2 = -0.5 < 0$.

5. Suppose that oranges and apples are equally priced, at p. The budget equation is then $px + py = 1$. The Lagrangean for the utility maximisation problem, subject to the budget equation, is

$$L = x^{\alpha} y^{1-\alpha} - \lambda(px + py - 1),$$

and the first-order conditions are

$$\begin{aligned}
\frac{\partial L}{\partial x} &= \alpha x^{\alpha-1} y^{1-\alpha} - p\lambda = 0 \\
\frac{\partial L}{\partial y} &= (1-\alpha) x^{\alpha} y^{-\alpha} - p\lambda = 0 \\
\frac{\partial L}{\partial \lambda} &= 1 - px - py = 0
\end{aligned}$$

From the first two equations,

$$\lambda = \frac{1}{p}\alpha x^{\alpha-1} y^{1-\alpha} = \frac{1}{p}(1-\alpha) x^{\alpha} y^{-\alpha},$$

so

$$y^{1-\alpha} y^{\alpha} = \frac{(1-\alpha)}{\alpha} x^{\alpha} x^{1-\alpha},$$

so

$$y = \frac{(1-\alpha)}{\alpha} x$$

when utility is maximised. Now, we're told that when utility is maximised, we have $x = 1$ and $y = 2$, so we must have $(1-\alpha)/\alpha = 2$, which means $1 - \alpha = 2\alpha$, or $\alpha = 1/3$. The budget equation tells us what p is, since $p(1) = p(2) = 1$, so $p = 1/3$.

Now, suppose the price of apples is halved, to $1/6$. Then the Lagrangean is

$$L = x^{1/3}y^{2/3} - \lambda\left(\frac{1}{6}x + \frac{1}{3}y - 1\right),$$

and the first-order conditions are

$$\begin{aligned}
\frac{\partial L}{\partial x} &= (1/3)x^{-2/3}y^{2/3} - (1/6)\lambda = 0\\
\frac{\partial L}{\partial y} &= (2/3)x^{1/3}y^{-1/3} - (1/3)\lambda = 0\\
\frac{\partial L}{\partial \lambda} &= 1 - (1/6)x - (1/3)y = 0.
\end{aligned}$$

The first two equations show that

$$\lambda = 2x^{-2/3}y^{2/3} = 2x^{1/3}y^{-1/3},$$

so $x = y$. By the third equation, $(1/2)x = 1$, so $x = 2$ and $y = 2$.

6. The constraint is given by the fact that there are only 100 units of the raw material, so that $x + y + z = 100$. The function to be maximised is the revenue, which is equal to the selling price (which is fixed across all three sectors) times the quantity produced. Thus, the revenue is proportional to the quantity produced, which is $3\sqrt{x} + 4\sqrt{y} + 5\sqrt{z}$, so it suffices to maximise this function subject to $x + y + z - 100 = 0$. The Lagrangean is

$$L = 3\sqrt{x} + 4\sqrt{y} + 5\sqrt{z} - \lambda(x + y + z - 100),$$

and the equations to be solved are

$$\begin{aligned}
\frac{\partial L}{\partial x} &= \frac{3}{2\sqrt{x}} - \lambda = 0\\
\frac{\partial L}{\partial y} &= \frac{2}{\sqrt{y}} - \lambda = 0\\
\frac{\partial L}{\partial z} &= \frac{5}{2\sqrt{z}} - \lambda = 0\\
\frac{\partial L}{\partial \lambda} &= -x - y - z + 100 = 0
\end{aligned}$$

The first three equations give three expressions for λ:

$$\lambda = \frac{3}{2\sqrt{x}} = \frac{2}{\sqrt{y}} = \frac{5}{2\sqrt{z}}.$$

These relationships enable us to express all three variables x, y, z in terms of any given one of them. Let's express y and z in terms of x. We have $3/(2\sqrt{x}) = 2/\sqrt{y}$, so $\sqrt{y} = (4/3)\sqrt{x}$ which, on squaring both sides, tells us that $y = (16/9)x$. Similarly, from $3/(2\sqrt{x}) = 5/(2\sqrt{z})$, we obtain $\sqrt{z} = (5/3)\sqrt{x}$ and $z = (25/9)x$. Then the constraint equation $x + y + z = 100$ becomes

$$x + \frac{16}{9}x + \frac{25}{9}x = 100,$$

so $(50/9)x = 100$ and $x = 18$. Then, $y = (16/9)(18) = 32$ and $z = (25/9)(18) = 50$. So the optimal values of x, y, z are:

$$x = 18,\ y = 32,\ z = 50.$$

7. The Lagrangean is

$$L = \frac{1}{x} + \frac{1}{y} + \frac{1}{z} - \lambda\left(x + y + z - c\right).$$

The first order conditions are

$$\begin{aligned}
\frac{\partial L}{\partial x} &= -\frac{1}{x^2} - \lambda = 0 \\
\frac{\partial L}{\partial y} &= -\frac{1}{y^2} - \lambda = 0 \\
\frac{\partial L}{\partial z} &= -\frac{1}{z^2} - \lambda = 0 \\
\frac{\partial L}{\partial \lambda} &= -x - y - z + c = 0
\end{aligned}$$

From the first three equations,

$$\lambda = -\frac{1}{x^2} = -\frac{1}{y^2} = -\frac{1}{z^2}.$$

So, $x = y = z$.

By the third equation, $3x = c$, so the function is minimised when $x = y = z = c/3$. The minimum value is

$$\frac{1}{c/3} + \frac{1}{c/3} + \frac{1}{c/3} = \frac{9}{c}.$$

Now, given any $x, y, z > 0$, let $c = x + y + z$. By what has just been shown, for any X, Y, Z, with $X + Y + Z = c$,

$$\frac{1}{X} + \frac{1}{Y} + \frac{1}{Z} \geq \frac{9}{c} = \frac{9}{x + y + z}.$$

In particular, this must be true for $X = x$, $Y = y$ and $Z = z$. Therefore

$$\frac{1}{x} + \frac{1}{y} + \frac{1}{z} \geq \frac{9}{c} = \frac{9}{x + y + z},$$

and so

$$\frac{x + y + z}{3} \geq 3\left(\frac{1}{1/x + 1/y + 1/z}\right),$$

or, in other words,

$$\frac{1}{3}(x + y + z) \geq \left\{\frac{1}{3}\left(\frac{1}{x} + \frac{1}{y} + \frac{1}{z}\right)\right\}^{-1}.$$

Chapter 4

Linear Algebra and Applications

Essential reading

(For full publication details, see Chapter 1.)

Anthony, M. and Biggs, N. *Mathematics for Economics and Finance.* Chapters 15–20.

Recommended reading

Black, J. and Bradley, J.F. *Essential Mathematics for Economists.* Chapters 10 and 11.

Bradley, T. and Patton, P. *Essential Mathematics for Economics and Finance.* Chapter 9, Sections 9.3 and 9.4.

Dowling, Edward T. *Introduction to Mathematical Economics.* Chapters 10 and 11 and Chapter 12, Section 12.7.

Holden, K, and Pearson, A.W., *Introductory Mathematics for Economics and Business*: Chapter 1, Section 1.13. Chapter 2, sections 2.1, 2.2, 2.6, 2.12, 2.13.

Ostaszewski, A. *Mathematics in Economics: Models and Methods.* Chapter 6, Sections 6.5, 6.7, 6.9, 6.10, 6.11.

Simon, C.P. and Blume, L., *Mathematics for Economists*: Chapter 7, Sections 7.1–7.3, Chapter 8, Sections 8.1–8.5, Chapter 9, Section 9.1, and Chapter 23, Sections 23.1 and 23.3.

Introduction

You will have seen from Mathematics 1 that matrices and linear equations play an important role in analysing economics mathematically. In this chapter we develop further the theory of matrices and linear equations and present some applications.

Revision of Mathematics 1

Matrices

You should recall how to add and multiply vectors and matrices. I will not re-iterate all the definitions here, but if you need to refresh your memory, consult the Mathematics 1 subject guide or the textbooks (particularly the Anthony and Biggs book).

Linear equations

The most important use of matrices we met in Mathematics 1 was their application to solving systems of linear equations.

Recall that a **system of m linear equations in n unknowns** $x_1, x_2, \ldots, x_n$ is a set of m equations of the form

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

The numbers a_{ij} are usually known as the **coefficients** of the system. We say that $(x_1^*, x_2^*, \ldots, x_n^*)$ is a **solution** of the system if **all** m equations hold true when $x_1 = x_1^*$, $x_2 = x_2^*$ and so on. Sometimes a system of linear equations is known as a set of **simultaneous** equations; such terminology emphasises that a solution is an assignment of values to each of the n unknowns such that **each and every** equation holds with this assignment.

In order to deal with large systems of linear equations we usually write them in matrix form. First we observe that vectors are just special cases of matrices: a row vector or list of n numbers is simply a matrix of size $1 \times n$, and a column vector is a matrix of size $n \times 1$. The rule for multiplying matrices tells us how to calculate the product $A\mathbf{x}$ of an $m \times n$ matrix A and an $n \times 1$ column vector $\mathbf{x}$. According to the rule, $A\mathbf{x}$ is

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \qquad\qquad \vdots \\ a_{m1}x_1 + a_{n2}x_2 + \cdots + a_{mn}x_n \end{pmatrix}.$$

Note that $A\mathbf{x}$ is a column vector with m rows, these being the left-hand sides of our system of linear equations. If we define another column vector $\mathbf{b}$, whose components are the right-hand sides b_i, the system is equivalent to the matrix equation

$$A\mathbf{x} = \mathbf{b}.$$

We often use the phrase **linear system** to mean 'system of linear equations' and we say that a linear system is **square** if the number of equations is the same as the number of unknowns; that is, if the matrix A is square.

In Mathematics 1, we used the method of row operations (or, as it is also known, the Gauss-Jordan elimination method) to solve some systems of linear equations. In this chapter of the guide, we shall use this method again to solve many more types of systems of linear equations. (Revision of the row operations method is deferred until later in this chapter, since we shall want to develop the method further than we did in Mathematics 1.)

Matrix inverses

Recall that a matrix A is square if it has the same number of rows as columns. We say that A has an **inverse matrix**[1] if there is a square matrix B such that

$$AB = BA = I,$$

the identity matrix. It turns out that a square matrix A either may have no inverse at all. But, if it does, then it has **only one**, which we denote by A^{-1}. We say that A is **invertible** or **nonsingular** if it has an inverse. We say that A is **non-invertible** or **singular** if it has no inverse.

[1] *See Anthony and Biggs, Section 18.2.*

For example, if

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \text{where} \quad ad - bc \neq 0,$$

then A has an inverse and

$$A^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Activity 4.1 *Check that this is indeed the inverse of A by showing that its products with A are the identity matrix I.*

The number $ad - bc$ is called the **determinant** of A and is denoted $|A|$. (I shall say more about determinants very soon.) Thus, a 2×2 matrix has an inverse if and only if its determinant is not 0.

Activity 4.2 *Let* $A = \begin{pmatrix} 3 & 2 \\ 5 & 1 \end{pmatrix}$. *Show that*

$$A^{-1} = \frac{1}{7}\begin{pmatrix} -1 & 2 \\ 5 & -3 \end{pmatrix}.$$

Inverse matrices and linear equations

Suppose we have to solve the system of linear equations $A\mathbf{x} = \mathbf{b}$, where A is a square matrix. If A has an inverse then

$$A^{-1}A\mathbf{x} = A^{-1}\mathbf{b},$$

that is, $\mathbf{x} = A^{-1}\mathbf{b}$. Clearly, $A^{-1}\mathbf{b}$ is a solution since $A(A^{-1}\mathbf{b}) = (AA^{-1})\mathbf{b} = I\mathbf{b} = \mathbf{b}$. Thus, if A is square and has an inverse then the system of linear equations $A\mathbf{x} = \mathbf{b}$ has the **unique** solution $\mathbf{x} = A^{-1}\mathbf{b}$.

Example: Consider the following very simple system of linear equations.

$$3x + 2y = 2 \quad (1)$$

$$5x + y = 2. \quad (2)$$

This can be written as

$$\begin{pmatrix} 3 & 2 \\ 5 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \end{pmatrix},$$

i.e.,

$$A\mathbf{x} = \mathbf{b}$$

where

$$\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}, \ \mathbf{b} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}.$$

Now, $|A| = 3 \times 1 - 2 \times 5 = -7 \neq 0$, so the linear system has precisely one solution,

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

Now, by the Activity above,

$$A^{-1} = \frac{1}{7} \begin{pmatrix} -1 & 2 \\ 5 & -3 \end{pmatrix}$$

and the solution is therefore

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{7} \begin{pmatrix} -1 & 2 \\ 5 & -3 \end{pmatrix} \begin{pmatrix} 2 \\ 2 \end{pmatrix} = \begin{pmatrix} 2/7 \\ 4/7 \end{pmatrix}.$$

That is, $x = 2/7$ and $y = 4/7$.

Determinants

The determinant

The **determinant**[2] of a square matrix A is a particular number associated with A, written $\det A$ or $|A|$. When A is a 2×2 matrix, the determinant is given by the formula

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

For example,

$$\det \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} = 1 \times 3 - 2 \times 2 = -1.$$

[2] *See Chapter 20 of Anthony and Biggs*

For a 3×3 matrix, one way of calculating the determinant is as follows:

$$\begin{aligned} \det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} &= \begin{vmatrix} a & b & c \\ d & e & f \\ g & h & i \end{vmatrix} \\ &= a \begin{vmatrix} e & f \\ h & i \end{vmatrix} - b \begin{vmatrix} d & f \\ g & i \end{vmatrix} + c \begin{vmatrix} d & e \\ g & h \end{vmatrix} \\ &= aei - afh + bfg - bdi + cdh - ceg. \end{aligned}$$

We have used **expansion by first row** to express the determinant in terms of 2×2 determinants. Note how this works and note, particularly, the minus sign in front of the second 2×2 determinant. (Expansion by other rows or by columns is also possible: see the textbooks for details.)

Example:

$$\begin{aligned} \begin{vmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 3 & 1 & 2 \end{vmatrix} &= 1 \begin{vmatrix} 1 & 1 \\ 1 & 2 \end{vmatrix} - 1 \begin{vmatrix} 0 & 1 \\ 3 & 2 \end{vmatrix} + 2 \begin{vmatrix} 0 & 1 \\ 3 & 1 \end{vmatrix} \\ &= 1(2-1) - 1(0-3) + 2(0-3) = -2. \end{aligned}$$

Determinants of larger matrices (4×4 and so on) can be evaluated recursively in terms of 3×3 determinants, but we present an easier method below. We refer to computing determinants in this way as computation **by expansion**.

Activity 4.3 *Calculate* $\begin{vmatrix} 1 & -1 & 0 \\ 2 & 0 & 1 \\ 3 & 1 & 1 \end{vmatrix}$.

Calculating determinants using row operations

Determinants of large matrices can often be more easily calculated by using a technique based on row operations. Recall that there are three basic types of row operation (which are often combined together):

- multiply every entry of a row of the matrix by a non-zero constant [Row operation of type (i).]
- add a multiple of one row of the matrix to another [Row operation of type (ii).]
- interchange two rows of the matrix. [Row operation of type (iii).]

It turns out that performing row operations on a matrix has the following effects on its determinant:

- If any row of a matrix is multiplied by a constant c, its determinant is also multiplied by c.
- If a multiple of one row is added to another, the determinant is unchanged.
- If two rows are interchanged, the determinant is multiplied by -1.

These observations, combined with the fact that the identity matrix has determinant equal to 1, will enable us to compute determinants.

Note first that if a matrix is a **diagonal matrix**, by which we mean it has nonzero entries only on its main diagonal, then its determinant is the product of these numbers. That is, if

$$C = \begin{pmatrix} c_1 & 0 & 0 & \dots & 0 \\ 0 & c_2 & 0 & \dots & 0 \\ 0 & 0 & c_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & c_n \end{pmatrix},$$

then $\det C = c_1 c_2 \dots c_n$. This is because C is obtained from I by multiplying the first row by c_1, the second row by c_2, and so on. Hence, using the observations above,

$$\det C = c_1 c_2 \dots c_n \det I = c_1 c_2 \dots c_n.$$

Now suppose we have a matrix in which every entry below the main diagonal is zero, known as an **upper triangular** matrix. That is, the matrix takes the form

$$C = \begin{pmatrix} c_1 & * & * & \dots & * \\ 0 & c_2 & * & \dots & * \\ 0 & 0 & c_3 & \dots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & c_n \end{pmatrix},$$

where the $*$'s are any numbers. Such a matrix can be reduced to the diagonal matrix

$$D = \begin{pmatrix} c_1 & 0 & 0 & \dots & 0 \\ 0 & c_2 & 0 & \dots & 0 \\ 0 & 0 & c_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & c_n \end{pmatrix},$$

by adding multiples of one row to another, *without altering the diagonal entries.* (Note that adding a negative multiple of a row is equivalent to subtracting a multiple of the row.) Hence, again by the observations above,

$$\begin{vmatrix} c_1 & * & * & \dots & * \\ 0 & c_2 & * & \dots & * \\ 0 & 0 & c_3 & \dots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & c_n \end{vmatrix} = \begin{vmatrix} c_1 & 0 & 0 & \dots & 0 \\ 0 & c_2 & 0 & \dots & 0 \\ 0 & 0 & c_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & c_n \end{vmatrix} = c_1 c_2 \dots c_n.$$

In other words, we have proved that

- the determinant of an upper triangular matrix is equal to the product of its diagonal entries.

It follows that we can calculate the determinant of a matrix by reducing it to upper triangular form, using row operations of types (ii) and (iii).

Example: The matrix

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

is reduced to an upper triangular matrix T as follows:

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

$$\to \begin{pmatrix} 1 & 3 & 2 & 5 \\ 0 & 2 & 2 & -5 \\ 0 & 0 & 3 & 2 \\ 0 & 0 & 0 & -\frac{13}{2} \end{pmatrix} = T.$$

Here, the second step is an interchange of rows (rule (iii)) and the other steps are covered by rule (ii). It follows that

$$\det A = -\det T = -(1 \times 2 \times 3 \times (-13/2)) = 39.$$

Activity 4.4 *Calculate* $\begin{vmatrix} 1 & -1 & 0 \\ 2 & 0 & 1 \\ 3 & 1 & 1 \end{vmatrix}$ *by using the method just described.*

Cramer's Rule

In Mathematics 1, you met the row operations method for solving systems of linear equations. An alternative technique for solving $A\mathbf{x} = \mathbf{b}$, where A is an invertible $n \times n$ matrix, is **Cramer's Rule**. Consider first a 2×2 system. Let us write the general system of two equations in two unknowns in the matrix form $A\mathbf{x} = \mathbf{b}$, that is,

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}.$$

Using the formula for the inverse we get

$$\begin{aligned} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &= \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}^{-1} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \\ &= \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \\ &= \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{pmatrix} a_{22}b_1 - a_{12}b_2 \\ -a_{21}b_1 + a_{11}b_2 \end{pmatrix}. \end{aligned}$$

After a minor rearrangement, this tells us that the solutions x_1 and x_2 can each be written in terms of determinants:

$$x_1 = \frac{\begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}, \qquad x_2 = \frac{\begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix}}{\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}}.$$

Cramer's rule in the 2×2 case can be stated in the following way: $x_i = \Delta_i/\Delta$, where Δ is the determinant of A and Δ_i is the determinant of the matrix obtained by replacing the ith column of A by $\mathbf{b}$. The result in the $n \times n$ case is just the same, so it is a neat way of writing down the solution.

Example: Let us find the solution of the linear system

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

by using Cramer's rule. First, we write it in matrix form $A\mathbf{x} = \mathbf{b}$ as

$$\begin{pmatrix} 1 & 2 & 1 \\ 2 & 2 & 0 \\ 3 & 5 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix},$$

Using the notation above, we have

$$\begin{aligned} \Delta = |A| = \begin{vmatrix} 1 & 2 & 1 \\ 2 & 2 & 0 \\ 3 & 5 & 4 \end{vmatrix} &= 1 \begin{vmatrix} 2 & 0 \\ 5 & 4 \end{vmatrix} - 2 \begin{vmatrix} 2 & 0 \\ 3 & 4 \end{vmatrix} + 1 \begin{vmatrix} 2 & 2 \\ 3 & 5 \end{vmatrix} \\ &= 1 \times 8 - 2 \times 8 + 1 \times 4 = -4. \end{aligned}$$

The other relevant determinants are as follows. (Check the calculations.)

$$\Delta_1 = \begin{vmatrix} 1 & 2 & 1 \\ 2 & 2 & 0 \\ 1 & 5 & 4 \end{vmatrix} = 0, \ \Delta_2 = \begin{vmatrix} 1 & 1 & 1 \\ 2 & 2 & 0 \\ 3 & 1 & 4 \end{vmatrix} = -4, \ \Delta_3 = \begin{vmatrix} 1 & 2 & 1 \\ 2 & 2 & 2 \\ 3 & 5 & 1 \end{vmatrix} = 4.$$

So the solution is given by

$$x_1 = \frac{\Delta_1}{\Delta} = 0, \ x_2 = \frac{\Delta_2}{\Delta} = 1, \ x_3 = \frac{\Delta_2}{\Delta} = -1.$$

It is very simple to verify that we have the right answer just by checking that all three equations hold with these values of x_1, x_2, x_3. It is a good idea always to check your answer in this way.

Activity 4.5 *Check that we do indeed have the correct answer by substituting the values back into the original equations.*

The square linear system when A is not invertible

It is the case that a matrix is invertible if and only if it has nonzero determinant. When the matrix A has nonzero determinant, the system $A\mathbf{x} = \mathbf{b}$ has just one solution, $\mathbf{x} = A^{-1}\mathbf{b}$. When the matrix A has zero determinant, the system $A\mathbf{x} = \mathbf{b}$ might have no solutions at all, or infinitely many solutions.[3] For example, the matrix

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

has determinant 0. Consider the system

$$\begin{aligned} x_1 + x_2 &= 1 \\ 2x_1 + 2x_2 &= 2. \end{aligned}$$

Clearly the second equation, being just the first equation doubled, adds no new restriction on x_1 and x_2, so this system is equivalent to the single equation

[3] *See Anthony and Biggs, Section 16.1 and 17.1.*

$x_1 + x_2 = 1$. But this has infinitely many solutions; let α be any number: then $x_1 = \alpha, x_2 = 1 - \alpha$ is a solution. On the other hand, consider the system

$$\begin{aligned} x_1 + x_2 &= 1 \\ 2x_1 + 2x_2 &= 1. \end{aligned}$$

This has **no** solutions, for the second equation is equivalent to $x_1 + x_2 = 1/2$ and this is in contradiction with the first equation, since $1 \neq 1/2$.

The examples just given show that we have to be careful if a square linear system is defined by a matrix which is not invertible. But things are far simpler in the very special case in which $\mathbf{b}$ is the **zero vector**,

$$\mathbf{0} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}.$$

This always has at least one solution (namely the zero vector itself). The following simple fact is important:

The square linear system $A\mathbf{x} = \mathbf{0}$ has solutions other than $\mathbf{x} = \mathbf{0}$ precisely when $|A| = 0$.

Row operations

The inverse matrix method and Cramer's rule for solving square linear systems can be generalised to square linear systems of any size. But they only work when A is invertible. We need a different technique for dealing with systems where A is not invertible. It is here that we find row operations again to be very useful. By way of revision, we briefly review the reasoning behind row operations.

It is a simple observation that the set of solutions of a system of linear equations is unaltered by the following three operations, since the restrictions on the variables $x_1, x_2, \ldots$ given by the new equations imply, and are implied by, the restrictions given by the old ones (that is, we can undo the manipulations made on the old system):

- multiply both sides of an equation by a non-zero constant
- add a multiple of one equation to another
- interchange two equations.

These observations form the motivation behind the row operations method[4] to solve linear equations.

[4] *See Mathematics 1 and Anthony and Biggs, Chapters 16 and 17.*

To solve a linear system $A\mathbf{x} = \mathbf{b}$ (where A need not be square), we first form the augmented matrix $(A\mathbf{b})$, which is A with column $\mathbf{b}$ tagged on. For example, if (as in our Example of Cramer's rule) the system is

$$\begin{pmatrix} 1 & 2 & 1 \\ 2 & 2 & 0 \\ 3 & 5 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix},$$

then the augmented matrix is the 3×4 matrix

$$(A\mathbf{b}) = \begin{pmatrix} 1 & 2 & 1 & 1 \\ 2 & 2 & 0 & 2 \\ 3 & 5 & 4 & 1 \end{pmatrix}.$$

We use this form because of the important fact that elementary operations on the equations of the system correspond to the same operations on the rows of the augmented matrix. The method now proceeds as follows: we use a sequence of elementary row operations on the augmented matrix until we have changed it into a matrix of the form

$$(C\mathbf{d}) = \begin{pmatrix} 1 & * & * & * & \dots & * & * \\ 0 & 0 & 1 & * & \dots & * & * \\ 0 & 0 & 0 & 1 & \dots & * & * \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix},$$

which is said to be in **echelon form**. Here, the $*$ symbols merely indicates some numbers. Note that in an echelon matrix, the first non-zero entry in each row is 1 (we call this the **leading** 1), the position of the leading 1 moves to the right as we go down the rows, and any rows which consist entirely of zeros are located at the bottom of the matrix. To see why this will be useful, let us carry out the procedure for the linear system given above. We have, by the use of certain row operations,

$$\begin{aligned} \begin{pmatrix} 1 & 2 & 1 & 1 \\ 2 & 2 & 0 & 2 \\ 3 & 5 & 4 & 1 \end{pmatrix} &\to \begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & -2 & -2 & 0 \\ 0 & -1 & 1 & -2 \end{pmatrix} \\ &\to \begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & 1 & -2 \end{pmatrix} \\ &\to \begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 2 & -2 \end{pmatrix} \\ &\to \begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}. \end{aligned}$$

Therefore the initial system has the same set of solutions as the system

$$\begin{pmatrix} 1 & 2 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

This system of equations is

$$x_1 + 2x_2 + x_3 = 1, \quad x_2 + x_3 = 0, \quad x_3 = -1.$$

But it is easy to solve these equations by working backwards from the third equation to the first one. Immediately, we have $x_3 = -1$. The second equation then gives $x_2 = 1$, and then the first gives $x_1 = 1 - 2x_2 - x_3 = 0$.

The method based on row operations is more useful than Cramer's rule, since, for instance, it applies to systems of linear equations in which the matrix A is not square. Moreover, even when A is square, this method may work when Cramer's rule does not. The following two examples demonstrate this [5].

[5] *See Anthony and Biggs, Chapter 17, for general discussion and further examples*

Example: Consider

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

As usual, we form the matrix

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

and apply elementary row operations to reduce it to echelon form. (Cramer's rule cannot be applied here, because as one would quickly discover, the determinant of the coefficient matrix is $|A| = 0$.)

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

$$\to \begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & -2 & -2 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

This last matrix represents the system

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

The third of these equations conveys no information at all about x_1, x_2 and x_3, for it simply tells us that $0 = 0$. Consequently, the original system has the same solutions as the following system of two equations in three unknowns.

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

These equations tell us first that, given x_3, we have $x_2 + x_3 = 0$, so $x_2 = -x_3$. Then we have $x_1 = 1 - 2x_2 - x_3 = 1 + 2x_3 - x_3 = 1 + x_3$, so both x_1 and x_2 are determined in terms of x_3. Indeed, if we let x_3 be any real number s, then

$$x_1 = 1 + s, \quad x_2 = -s, \quad x_3 = s$$

is a solution to the system. So there are infinitely many solutions in this example.

Example: Consider the system of equations

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

Using row operations to reduce the augmented matrix to echelon form, we obtain

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

$$\to \begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & -2 & -2 & -1 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

Thus the original system of equations is equivalent to the system

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 1 \\ x_2 + x_3 &= 0 \\ 0x_1 + 0x_2 + 0x_3 &= -1. \end{aligned}$$

But this system has no solutions, since there are no values of x_1, x_2, x_3 which satisfy the last equation. It reduces to the false statement '$0 = -1$', whatever values we give the unknowns. We deduce, therefore, that the original system has no solutions. Such a system is said to be **inconsistent**.

Activity 4.6 *Use row operations to show that the system*

$$\begin{aligned} x_1 - 2x_2 + 3x_3 &= 5 \\ 2x_1 + x_2 - x_3 &= 1 \\ 4x_1 - 3x_2 + 5x_3 &= 11 \end{aligned}$$

has infinitely many solutions. Find a formula for the general solution. Use row operations to show that if the number '11' on the right-hand side of the third equation is replaced by 2, then the resulting system has no solutions.

Calculating inverses using determinants

We already have noted a formula for the inverse of an invertible 2×2 matrix: if $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and $\det(A) = ad - bc \neq 0$, then

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

There is also a formula for the inverse of an invertible 3×3 matrix, though it is much more complicated.[6] Suppose we have the 3×3 matrix

[6] *See Ostaszewski, Section 6.11.4 for an explanation.*

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}.$$

Then A is invertible precisely when $\det(A) \neq 0$. For each i and j between 1 and 3, let A_{ij} be the determinant of the 2×2 matrix obtained from A by deleting row i and column j. For example,

$$A_{23} = \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix}.$$

Then it turns out that if $\det(A) \neq 0$, then A is invertible and

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} +A_{11} & -A_{21} & +A_{31} \\ -A_{12} & +A_{22} & -A_{32} \\ +A_{13} & -A_{23} & +A_{33} \end{pmatrix}.$$

Note carefully the 'alternating' pattern of signs attached to the entries of the matrix.

The matrix

$$\begin{pmatrix} +A_{11} & -A_{12} & +A_{13} \\ -A_{21} & +A_{22} & -A_{23} \\ +A_{31} & -A_{32} & +A_{33} \end{pmatrix}$$

is given a special name. Note that it is the **transpose** of the matrix appearing on the right in the formula above; that is, it is obtained by interchanging the rows and columns. It is called the **adjugate** of the matrix A, denoted $\operatorname{adj}(A)$. For this reason you might sometimes see the formula

$$A^{-1} = \frac{1}{\det(A)} \left(\operatorname{adj}(A)\right)^T .$$

Calculating the inverse this way requires care, since we have to compute a 3×3 determinant ($\det(A)$), then 9 different 2×2 determinants, then we have to attach the correct signs to these, and place them in the right positions. Here is an example.

Example: Let

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

We shall calculate its inverse using the method just described. First, we find its determinant:

$$\begin{aligned} \det(A) &= -2\begin{vmatrix} -1 & 1 \\ 2 & 0 \end{vmatrix} - 1\begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} + 3\begin{vmatrix} 0 & -1 \\ 1 & 2 \end{vmatrix} \\ &= -2(-2) - 1(-1) + 3(1) = 8 \end{aligned}$$

Since this is nonzero, we know that the inverse exists. We now calculate the relevant 2×2 determinants.

$$A_{11} = \begin{vmatrix} -1 & 1 \\ 2 & 0 \end{vmatrix} = -2,$$

$$A_{12} = \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = -1$$

$$A_{13} = \begin{vmatrix} 0 & -1 \\ 1 & 2 \end{vmatrix} = 1,$$

$$A_{21} = \begin{vmatrix} 1 & 3 \\ 2 & 0 \end{vmatrix} = -6,$$

$$A_{22} = \begin{vmatrix} -2 & 3 \\ 1 & 0 \end{vmatrix} = -3,$$

$$A_{23} = \begin{vmatrix} -2 & 1 \\ 1 & 2 \end{vmatrix} = -5,$$

$$A_{31} = \begin{vmatrix} 1 & 3 \\ -1 & 1 \end{vmatrix} = 4,$$

$$A_{32} = \begin{vmatrix} -2 & 3 \\ 0 & 1 \end{vmatrix} = -2,$$

$$A_{33} = \begin{vmatrix} -2 & 1 \\ 0 & -1 \end{vmatrix} = 2.$$

Then,

$$\operatorname{adj}(A) = \begin{pmatrix} +A_{11} & -A_{12} & +A_{13} \\ -A_{21} & +A_{22} & -A_{23} \\ +A_{31} & -A_{32} & +A_{33} \end{pmatrix} = \begin{pmatrix} -2 & 1 & 1 \\ 6 & -3 & 5 \\ 4 & 2 & 2 \end{pmatrix}$$

and hence

$$A^{-1} = \frac{1}{\det(A)} \left(\operatorname{adj}(A)\right)^T = \frac{1}{8}\begin{pmatrix} -2 & 6 & 4 \\ 1 & -3 & 2 \\ 1 & 5 & 2 \end{pmatrix}.$$

You can easily check that this is correct by checking that the product of A with this matrix is the identity matrix.

Activity 4.7 *Check that this answer is correct.*

Activity 4.8 *Use this method to determine the inverse of the matrix*

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

Calculating inverses using row operations

Calculation of inverses using determinants is a rather difficult technique, easy to get wrong, and very impractical for large matrices. Instead we show how one can use elementary row operations to find the inverse of a matrix.[7] We start with the matrix A and we form a new, larger, matrix by placing the identity matrix to the right of A, obtaining the matrix denoted $(A\ I)$. We then use row operations to reduce this to $(I\ B)$. If this is not possible (which will become apparent) then the matrix is not invertible. If it can be done, then A is invertible and $A^{-1} = B$.

[7] *See Anthony and Biggs, Section 18.3 for an explanation of why this technique works.*

Example: We use the same matrix as in the previous example. In order to determine whether the matrix

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

is invertible and, if so, to determine its inverse, we form the matrix

$$(A\ I) = \left(\begin{array}{ccc|ccc} -2 & 1 & 3 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 & 0 & 1 \end{array}\right).$$

(We have separated A from I by a vertical line just to emphasise how this matrix is formed.) Then we carry out elementary row operations.

$$\left(\begin{array}{ccc|ccc} -2 & 1 & 3 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 & 1 & 0 \\ 1 & 2 & 0 & 0 & 0 & 1 \end{array}\right) \to \left(\begin{array}{ccc|ccc} -2 & 1 & 3 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 & 1 & 0 \\ 0 & \frac{5}{2} & \frac{3}{2} & \frac{1}{2} & 0 & 1 \end{array}\right)$$

$$\to \left(\begin{array}{ccc|ccc} -2 & 1 & 3 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 4 & \frac{1}{2} & \frac{5}{2} & 1 \end{array}\right)$$

$$\to \left(\begin{array}{ccc|ccc} 1 & -\frac{1}{2} & -\frac{3}{2} & -\frac{1}{2} & 0 & 0 \\ 0 & 1 & -1 & 0 & -1 & 0 \\ 0 & 0 & 1 & \frac{1}{8} & \frac{5}{8} & \frac{1}{4} \end{array}\right)$$

$$\to \left(\begin{array}{ccc|ccc} 1 & -\frac{1}{2} & -\frac{3}{2} & -\frac{1}{2} & 0 & 0 \\ 0 & 1 & 0 & \frac{1}{8} & -\frac{3}{8} & \frac{1}{4} \\ 0 & 0 & 1 & \frac{1}{8} & \frac{5}{8} & \frac{1}{4} \end{array}\right)$$

$$\to \left(\begin{array}{ccc|ccc} 1 & 0 & -\frac{3}{2} & -\frac{7}{16} & -\frac{3}{16} & \frac{1}{8} \\ 0 & 1 & 0 & \frac{1}{8} & -\frac{3}{8} & \frac{1}{4} \\ 0 & 0 & 1 & \frac{1}{8} & \frac{5}{8} & \frac{1}{4} \end{array}\right)$$

$$\rightarrow \quad \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -\frac{1}{4} & \frac{3}{4} & \frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{8} & -\frac{3}{8} & \frac{1}{4} \\ 0 & 0 & 1 & \frac{1}{8} & \frac{5}{8} & \frac{1}{4} \end{array}\right).$$

This is now in the form $(I\ B)$, so we deduce that A is invertible and that

$$A^{-1} = B = \frac{1}{8}\begin{pmatrix} -2 & 6 & 4 \\ 1 & -3 & 2 \\ 1 & 5 & 2 \end{pmatrix}.$$

Activity 4.9 *Look carefully at the calculation just carried out, and write down which row operations have been used at each stage.*

Activity 4.10 *Use the row operations method to determine the inverse of the matrix*

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

(This is the matrix in Activity 4.7 above, whose inverse you should have found using the other method.)

Application: input-output analysis

Suppose an economy has n interdependent production processes, manufacturing commodities C_1, C_2, ..., C_n. The production process for any one of the commodities requires an input, which uses part of the output of some of the others. In addition, there is an external demand for each commodity. The problem is to determine the production schedule which enables each process to meet all the demands for its product. The following example illustrates how such a problem can be described as a system of linear equations. [8]

[8] *See Anthony and Biggs, Sections 19.1 and 19.2 for a general analysis.*

Example: The production processes for three goods, C_1, C_2, C_3 are interlinked, as follows:

- To produce one unit of C_1 requires the input of 0.2 of C_1, 0.4 of C_2 and 0.1 of C_3.
- To produce one unit of C_2 requires 0.3 of C_1, 0.1 of C_2 and 0.3 worth of C_3.
- To produce one unit of C_3 requires 0.2 of each of C_1, C_2 and C_3.

Suppose that, in a given time period, there is an external demand for d_1 of C_1, d_2 of C_2 and d_3 of C_3. We wish to know the production levels x_1, x_2, x_3 of C_1, C_2, C_3 required to satisfy all demands in the given period. Consider first the total demand for C_1. This is d_1, the external demand, plus the quantity required to produce C_1, C_2 and C_3. Each unit of C_1 requires 0.2 units of C_1, each unit of C_2 requires 0.3

of C_1 and each unit of C_3 requires 0.2 of C_1. Since the quantities of C_1, C_2, C_3 being produced are x_1, x_2, x_3, the total demand for C_1 is therefore

$$x_1 = d_1 + 0.2x_1 + 0.3x_2 + 0.2x_3.$$

Similarly, considering the total demands for C_2 and C_3 shows that

$$\begin{aligned} x_2 &= d_2 + 0.4x_1 + 0.1x_2 + 0.2x_3 \\ x_3 &= d_3 + 0.1x_1 + 0.3x_2 + 0.2x_3. \end{aligned}$$

The system of equations is simply

$$\mathbf{x} = \mathbf{d} + A\mathbf{x},$$

where

$$\mathbf{d} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix}, \quad A = \begin{pmatrix} 0.2 & 0.3 & 0.2 \\ 0.4 & 0.1 & 0.2 \\ 0.1 & 0.3 & 0.2 \end{pmatrix}.$$

A is known as the **technology matrix**. Rearranging we get $\mathbf{x} - A\mathbf{x} = \mathbf{d}$, and using the fact that $I\mathbf{x} = \mathbf{x}$, where I is the identity matrix, we see that we have to solve the linear system

$$(I - A)\mathbf{x} = \mathbf{d}.$$

Eigenvalues and eigenvectors

One of the most useful techniques in applications of matrices and linear algebra is **diagonalisation**. Before discussing this, we have to look at the topic of **eigenvalues and eigenvectors**. In this subject, we shall be primarily interested in eigenvalues, eigenvectors and diagonalisation in the case of 2×2 matrices. Those who subsequently take the subject **Further Mathematics for Economists** will find there a more extensive investigation of these topics and their applications.

Definitions

Suppose that A is a square matrix. The number λ is said to be an **eigenvalue** of A if for some **non-zero** vector $\mathbf{x}$, $A\mathbf{x} = \lambda\mathbf{x}$. Any **non-zero** vector $\mathbf{x}$ for which this equation holds is called an **eigenvector for eigenvalue** λ or an **eigenvector of** A **corresponding to eigenvalue** λ**.**

Finding eigenvalues and eigenvectors

To determine whether λ is an eigenvalue of A, we need to determine whether there are any non-zero solutions to the matrix equation $A\mathbf{x} = \lambda\mathbf{x}$. Note that the matrix equation $A\mathbf{x} = \lambda\mathbf{x}$ is not of the standard form, since the right-hand side is not a fixed vector $\mathbf{b}$, but depends explicitly on $\mathbf{x}$. However, we can rewrite it in standard form. Note that $\lambda\mathbf{x} = \lambda I\mathbf{x}$, where I is, as usual, the identity matrix. So, the equation is equivalent to $A\mathbf{x} = \lambda I\mathbf{x}$, or $A\mathbf{x} - \lambda I\mathbf{x} = \mathbf{0}$, which is equivalent to $(A - \lambda I)\mathbf{x} = \mathbf{0}$. Now, a square linear system $B\mathbf{x} = \mathbf{0}$ has solutions other than $\mathbf{x} = \mathbf{0}$ precisely when $|B| = 0$. Therefore, taking $B = A - \lambda I$, λ is an eigenvalue if and only if the determinant of the matrix $A - \lambda I$ is zero. This determinant, $p(\lambda) = |A - \lambda I|$, is known as the **characteristic polynomial** of A, since it turns out to be a polynomial in the variable λ. To find the eigenvalues, we solve the equation $|A - \lambda I| = 0$. Let us illustrate with a very simple 2×2 example.

Example: Let

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}.$$

Then

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

and the characteristic polynomial is

$$\begin{aligned} |A - \lambda I| &= \begin{vmatrix} 1-\lambda & 1 \\ 2 & 2-\lambda \end{vmatrix} \\ &= (1-\lambda)(2-\lambda) - 2 \\ &= \lambda^2 - 3\lambda + 2 - 2 = \lambda^2 - 3\lambda. \end{aligned}$$

So the eigenvalues are the solutions of $\lambda^2 - 3\lambda = 0$. To solve this, one could use either the formula for the solutions to a quadratic, or simply observe that the equation is $\lambda(\lambda - 3) = 0$ with solutions $\lambda = 0$ and $\lambda = 3$. Hence the eigenvalues of A are 0 and 3.

To find an eigenvector for eigenvalue λ, we have to find a solution to $(A - \lambda I)\mathbf{x} = \mathbf{0}$, **other than the zero vector**. (I stress the fact that eigenvectors cannot be the zero vector because this is a mistake many students make.) This is easy, since for a particular value of λ, all we need to do is solve a simple linear system We illustrate by finding the eigenvectors for the matrix of the example just given.

Example: We find eigenvectors of

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}.$$

We have seen that the eigenvalues are 0 and 3. To find an eigenvector for eigenvalue 0 we solve the system $(A - 0I)\mathbf{x} = \mathbf{0}$: that is, $A\mathbf{x} = \mathbf{0}$, or

$$\begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

This could be solved using row operations. (Note that it cannot be solved by using inverse matrices since A is not invertible. In fact, inverse matrix techniques or Cramer's rule will never be of use here since λ being an eigenvalue means that $A - \lambda I$ is not invertible.) However, we can solve this fairly directly just by looking at the equations. We have to solve

$$x_1 + x_2 = 0, \;\; 2x_1 + 2x_2 = 0.$$

Clearly both equations are equivalent. From either one, we obtain $x_1 = -x_2$. We can choose x_2 to be any number we like. Let's take $x_2 = 1$; then we need $x_1 = -x_2 = -1$. It follows that an eigenvector for 0 is

$$\mathbf{x} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

The choice $x_2 = 1$ was arbitrary; we could have chosen any non-zero number, so, for example, the following are eigenvectors for 0:

$$\begin{pmatrix} 2 \\ -2 \end{pmatrix}, \begin{pmatrix} 5.2 \\ -5.2 \end{pmatrix}.$$

There are infinitely many eigenvectors for 0: for each $\alpha \neq 0$,

$$\begin{pmatrix} \alpha \\ -\alpha \end{pmatrix}$$

is an eigenvector for 0. But be careful not to think that you can choose $\alpha = 0$; for then $\mathbf{x}$ becomes the zero vector, and this is never an eigenvector, simply by definition. To find an eigenvector for 3, we solve $(A - 3I)\mathbf{x} = \mathbf{0}$, which is

$$\begin{pmatrix} -2 & 1 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

This is equivalent to the equations

$$-2x_1 + x_2 = 0, \;\; 2x_1 - x_2 = 0,$$

which are together equivalent to the single equation $x_2 = 2x_1$. If we choose $x_1 = 1$, we obtain the eigenvector

$$\mathbf{x} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

(Again, any non-zero scalar multiple of this vector is also an eigenvector for eigenvalue 3.)

Diagonalisation of a square matrix

Square matrices A and B are **similar** if there is an invertible matrix P such that $P^{-1}AP = B$.

The matrix A is **diagonalisable** if it is similar to a diagonal matrix; in other words, if there is a diagonal matrix D and an invertible matrix P such that $P^{-1}AP = D$.

Suppose that matrix A is diagonalisable, and that $P^{-1}AP = D$, where D is a diagonal matrix

$$D = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n) = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix}.$$

(Note the useful notation for describing the diagonal matrix D.) Then we have $AP = PD$. If the columns of P are the vectors $\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n$, then

$$AP = A\,(\mathbf{v}_1 \ldots \mathbf{v}_n) = (A\mathbf{v}_1 \ldots A\mathbf{v}_n),$$

and

$$PD = (\mathbf{v}_1 \ldots \mathbf{v}_n) \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} = (\lambda_1\mathbf{v}_1 \ldots \lambda_n\mathbf{v}_n).$$

So this means that

$$A\mathbf{v}_1 = \lambda_1\mathbf{v}_1, A\mathbf{v}_2 = \lambda_2\mathbf{v}_2, \ldots, A\mathbf{v}_n = \lambda_n\mathbf{v}_n.$$

The fact that P^{-1} exists means that none of the vectors $\mathbf{v}_i$ is the zero vector. So this means that (for $i = 1, 2, \ldots, n$) λ_i is an eigenvalue of A and $\mathbf{v}_i$ is a corresponding eigenvector.

So we have established that an $n \times n$ matrix A is diagonalisable if we can find an invertible matrix P whose columns are eigenvectors of A. It is not always the case that such a matrix can be found, but in the 2×2 case the issue is much simpler: if $\mathbf{v}$ and $\mathbf{w}$ are eigenvectors which are not multiples of each other (which will certainly be true if A has two different eigenvalues and these eigenvectors correspond to exactly one of these eigenvalues each), then the matrix P with columns $\mathbf{v}$ and $\mathbf{w}$ will be as required.

Example: Consider $A = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}$. We have seen above that the eigenvalues of A are 0 and 3 and that examples of corresponding eigenvectors are, respectively

$$\begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

If we let

$$P = \begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix},$$

then P^{-1} exists and we should have $P^{-1}AP = D$ where

$$D = \begin{pmatrix} 0 & 0 \\ 0 & 3 \end{pmatrix}.$$

It is easy to check that this is so, by calculating P^{-1} and then evaluating $P^{-1}AP$, seeing that this product equals D.

Activity 4.11 *Check this: determine P^{-1} and calculate $P^{-1}AP$.*

Not all $n \times n$ matrices have n linearly independent eigenvalues, as the following example shows.

Example: The 2×2 matrix

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

has characteristic polynomial $\lambda^2 - 6\lambda + 9 = (\lambda - 3)^2$, so there is only one eigenvalue, $\lambda = 3$. The eigenvectors are the non-zero solutions to $(A - 3I)\mathbf{x} = \mathbf{0}$: that is,

$$\begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

This is equivalent to the single equation $x_1 + x_2 = 0$, with general solution $x_1 = -x_2$. Setting $x_2 = r$, we see that the solution set of the system consists of all vectors of the form $\begin{pmatrix} -r \\ r \end{pmatrix}$ as r runs through all non-zero real numbers. So the eigenvectors are precisely the non-zero scalar multiples of the fixed vector $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$. Any two eigenvectors are therefore multiples of each other and hence we cannot form an *invertible* matrix P with eigenvectors as its columns. So the matrix is not diagonalisable.

Learning outcomes

At the end of this chapter and the relevant reading, you should be able to:

- demonstrate the learning outcomes of the Matrices and Linear Equations section of **Mathematics 1**

- calculate determinants of 2×2 and 3×3 matrices by 'expansion'

- calculate determinants by using row operations to reduce a matrix to upper triangular form

- know what is meant by the inverse of a matrix, and what it means to say that a matrix is invertible

- understand why when a system of linear equations has an invertible coefficient matrix, then there is a unique solution to the system

- state and use the formula for the inverse of an invertible 2×2 matrix

- use Cramer's rule

- solve linear systems using row operations

- determine, using row operations, when a system is inconsistent

- solve, using row operations, systems with infinitely many solutions

- calculate inverses of 3×3 matrices by computing the adjugate

- calculate matrix inverses using row operations

- solve input-output problems

- understand what is meant by an eigenvalue and a corresponding eigenvector

- determine eigenvalues and eigenvectors of 2×2 matrices

- understand what is meant by saying that a matrix is diagonalisable

- diagonalise a 2×2 matrix, or determine that it cannot be diagonalised

Sample examination/practice questions

1. Suppose

$$C = C_0 + bY, \quad I = I_0 - ar \quad (a, b, C_0, I_0 \text{ positive constants})$$

and that $Y = C + I$. Suppose also that

$$M_d = M_0 + fY - gr, \quad (f, g, M_0 \text{ positive constants})$$

and $M_s = M_d$. Show that

$$fY - gr = M_s - M_0$$

$$(1-b)Y + ar = C_0 + I_0 + G.$$

Use matrix algebra to prove that the solutions Y^*, r^* to these equations are

$$\begin{aligned} Y^* &= \frac{a(M_s - M_0) + g(C_0 + I_0 + G)}{fa + g(1-b)} \\ r^* &= \frac{-(1-b)(M_s - M_0) + f(C_0 + I_0 + G)}{fa + g(1-b)}. \end{aligned}$$

2. For which value of c is the following system of equations consistent? Find all the solutions when c has this value.

$$\begin{aligned} x_1 + x_2 + x_3 &= 2 \\ 2x_1 + x_2 + 2x_3 &= 5 \\ 4x_1 + 3x_2 + 4x_3 &= c. \end{aligned}$$

3. A firm manufactures 3 different types of chocolate bar, 'Abracadabra', 'Break' and 'Choca-mocha'. The main ingredients in each are cocoa, milk and coffee. To produce 1000 Abracadabra bars requires 5 units of cocoa, 3 units of milk and 2 units of coffee. To produce 1000 Break bars requires 5 units of cocoa, 4 of milk and 1 of coffee, and the production of 1000 Choca-mocha bars requires 5 units of cocoa, 2 of milk and 3 of coffee. The firm has supplies of 250 units of cocoa, 150 of milk and 100 of coffee each week (and as much as it wants of the other ingredients, such as sugar). Show that if the firm uses up its supply of cocoa, milk and coffee, then the number of Break bars produced each week equals the number of Choca-mocha bars produced. How does the number of Abracadabra bars produced relate to the production level of the other two bars? Find the maximum possible weekly production of Choca-Mocha bars.

4. Determine the inverse of the matrix

$$A = \begin{pmatrix} 3 & 2 & -1 \\ 1 & 7 & 5 \\ -1 & 0 & 1 \end{pmatrix}.$$

5. Determine whether the following matrix is invertible, and find its inverse if it is:

$$\begin{pmatrix} 1 & 1 & 5 \\ 2 & 3 & 1 \\ 1 & 0 & 14 \end{pmatrix}.$$

6. Consider an economy with three industries: coal, electricity, railways. To produce \$1 of coal requires \$0.25 worth of electricity and \$0.25 rail costs for transportation. To produce \$1 of electricity requires \$0.65 worth of coal for fuel, \$0.05 of electricity for the auxiliary equipment, and \$0.05 for transport. To provide \$1 worth of transport, the railway requires \$0.55 coal for fuel and \$0.10 electricity. Each week the external demand for coal is \$50 000 and the external demand for electricity is \$25 000. There is no external demand for the railway. Find a system of linear equations which determines the weekly production schedule for each of the three industries. Express the system in matrix form.

7. Find the eigenvalues of the matrix $\begin{pmatrix} 2 & 1 \\ 3 & 1 \end{pmatrix}$.

8. Find the eigenvalues of the matrix $A = \begin{pmatrix} 2 & 4 \\ 3 & 3 \end{pmatrix}$ and determine an eigenvector corresponding to each eigenvalue. Find an invertible matrix P and a diagonal matrix D such that $P^{-1}AP = D$.

Answers to activities

4.2 We have $\det(A) = 3(1) - 2(5) = -7$. This is nonzero, so the inverse exists. Using the formula for the inverse of a 2×2 matrix, we have

$$A^{-1} = \frac{1}{-7}\begin{pmatrix} 1 & -2 \\ -5 & 3 \end{pmatrix} = \frac{1}{7}\begin{pmatrix} -1 & 2 \\ 5 & -3 \end{pmatrix}.$$

4.3 We have

$$\begin{aligned} \begin{vmatrix} 1 & -1 & 0 \\ 2 & 0 & 1 \\ 3 & 1 & 1 \end{vmatrix} &= 1\begin{vmatrix} 0 & 1 \\ 1 & 1 \end{vmatrix} - (-1)\begin{vmatrix} 2 & 1 \\ 3 & 1 \end{vmatrix} + 0\begin{vmatrix} 2 & 0 \\ 3 & 1 \end{vmatrix} \\ &= 1(-1) - (-1)(-1) = -2. \end{aligned}$$

4.4 Using row operations, we have

$$\begin{pmatrix} 1 & -1 & 0 \\ 2 & 0 & 1 \\ 3 & 1 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & -1 & 0 \\ 0 & 2 & 1 \\ 0 & 4 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & -1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & -1 \end{pmatrix}.$$

In the first step, we subtracted multiples of the first row from the second and third and in the second step we subtracted twice the second row from the third. These operations do not change the determinant, so the determinant is the determinant of the resulting upper triangular matrix, which is $1(2)(-1) = -2$.

4.6 The augmented matrix is

$$\begin{pmatrix} 1 & -2 & 3 & 5 \\ 2 & 1 & -1 & 1 \\ 4 & -3 & 5 & 11 \end{pmatrix}.$$

We reduce this to echelon form as follows:

$$\begin{pmatrix} 1 & -2 & 3 & 5 \\ 2 & 1 & -1 & 1 \\ 4 & -3 & 5 & 11 \end{pmatrix} \to \begin{pmatrix} 1 & -2 & 3 & 5 \\ 0 & 5 & -7 & -9 \\ 0 & 5 & -7 & -9 \end{pmatrix} \to \begin{pmatrix} 1 & -2 & 3 & 5 \\ 0 & 5 & -7 & -9 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

It follows that the system is consistent, and equivalent to the system

$$\begin{aligned} x_1 - 2x_2 + 3x_3 &= 5 \\ 5x_2 - 7x_3 &= -9. \end{aligned}$$

(The final row translates simply into the statement $0 = 0$.) There are infinitely many solutions. Letting $x_3 = r$, where r can be any number at all, we have $5x_2 - 7x_3 = -9$, so $x_2 = (7r - 9)/5$. From the first equation,

$$x_1 = 5 + 2x_2 - 3x_3 = 5 + \frac{2}{5}(7r - 9) - 3r = \frac{7}{5} - \frac{r}{5}.$$

So the general solution is

$$x_1 = \frac{7}{5} - \frac{r}{5}, \quad x_2 = \frac{7}{5}r - \frac{9}{5}, \quad x_3 = r.$$

When '11' on the right-hand side of the third equation is replaced by 2, the augmented matrix we start with is

$$\begin{pmatrix} 1 & -2 & 3 & 5 \\ 2 & 1 & -1 & 1 \\ 4 & -3 & 5 & 2 \end{pmatrix}.$$

It is easy to see that this reduces to

$$\begin{pmatrix} 1 & -2 & 3 & 5 \\ 0 & 5 & -7 & -9 \\ 0 & 0 & 0 & -9 \end{pmatrix}.$$

The last row shows that the system is inconsistent in this case (for if it was not, then we'd have $0x_1 + 0x_2 + 0x_3 = -9$, or $0 = -9$, which can't be).

4.8 The determinant of the matrix is 1 and the adjugate is

$$\text{adj}(A) = \begin{pmatrix} +(-2) & -(-1) & +(-1) \\ -(7) & +(3) & -(1) \\ +(5) & -(2) & +(1) \end{pmatrix},$$

so

$$A^{-1} = \frac{1}{\det(A)}\,(\text{adj}(A))^T = \begin{pmatrix} -2 & -7 & 5 \\ 1 & 3 & -2 \\ -1 & -1 & 1 \end{pmatrix}.$$

(It can easily be checked that this is correct by calculating AA^{-1} and noting that the product equals the identity matrix.)

4.9 The sequence of row operations performed (in each step) is:

$$R_3 \to R_3 + \frac{1}{2}R_1,$$
$$R_3 \to R_3 + \frac{5}{2}R_2,$$
$$R_1 \to R_1/2, \;\; R_3 \to R_3/4, \;\; R_2 \to -R_2,$$
$$R_2 \to R_2 + R_3,$$
$$R_1 \to R_1 + \frac{1}{2}R_2,$$
$$R_1 \to R_1 + \frac{3}{2}R_3.$$

(Here, $R_2 \to R_2 + R_1$, for example, means that the first row has been added to the second; in other words, the second row R_2 has become what was R_2, with R_1 added.)

4.10

$$\left(\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 1 & 3 & 1 & 0 & 1 & 0 \\ 2 & 5 & 1 & 0 & 0 & 1 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & 2 & -1 & 1 & 0 \\ 0 & 1 & 3 & -2 & 0 & 1 \end{array}\right)$$

$$\to \left(\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & 2 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 & -1 & 1 \end{array}\right)$$
$$\to \left(\begin{array}{ccc|ccc} 1 & 2 & -1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 3 & -2 \\ 0 & 0 & 1 & -1 & -1 & 1 \end{array}\right)$$
$$\to \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 0 & -1 & 1 \\ 0 & 1 & 0 & 1 & 3 & -2 \\ 0 & 0 & 1 & -1 & -1 & 1 \end{array}\right)$$
$$\to \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -2 & -7 & 5 \\ 0 & 1 & 0 & 1 & 3 & -2 \\ 0 & 0 & 1 & -1 & -1 & 1 \end{array}\right).$$

It follows that
$$A^{-1} = \begin{pmatrix} -2 & -7 & 5 \\ 1 & 3 & -2 \\ -1 & -1 & 1 \end{pmatrix},$$
in agreement with the answer obtained using the other method.

4.11 We easily get
$$P^{-1} = \frac{1}{3}\begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix}$$
using the formula for the inverse. Then,
$$\begin{aligned} P^{-1}AP &= \frac{1}{3}\begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -1 & 2 \end{pmatrix} \\ &= \frac{1}{3}\begin{pmatrix} 2 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 0 & 3 \\ 0 & 6 \end{pmatrix} \\ &= \frac{1}{3}\begin{pmatrix} 0 & 0 \\ 0 & 9 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & 3 \end{pmatrix}. \end{aligned}$$

Answers to sample examination/practice questions

1. Because $Y = C + I$, we have
$$Y = (C_0 + bY) + (I_0 - ar) + G, \quad \text{or} \quad (1-b)Y + ar = C_0 + I_0 + G.$$
Because $M_s = M_d$,
$$M_s = M_0 + fY - gr, \quad \text{or} \quad fY - gr = M_s - M_0.$$
We therefore have
$$\begin{aligned} fY - gr &= M_s - M_0 \\ (1-b)Y + ar &= C_0 + I_0 + G. \end{aligned}$$
In matrix terms this system is
$$\begin{pmatrix} f & -g \\ 1-b & a \end{pmatrix}\begin{pmatrix} Y \\ r \end{pmatrix} = \begin{pmatrix} M_s - M_0 \\ C_0 + I_0 + G \end{pmatrix}.$$

and the solution is

$$\begin{pmatrix} Y \\ r \end{pmatrix} = \begin{pmatrix} f & -g \\ 1-b & a \end{pmatrix}^{-1} \begin{pmatrix} M_s - M_0 \\ C_0 + I_0 + G \end{pmatrix}$$
$$= \frac{1}{fa + g(1-b)} \begin{pmatrix} a & g \\ -(1-b) & f \end{pmatrix} \begin{pmatrix} M_s - M_0 \\ C_0 + I_0 + G \end{pmatrix}.$$

Thus we have explicit formulae for the solutions Y^* and r^*:

$$Y^* = \frac{a(M_s - M_0) + g(C_0 + I_0 + G)}{fa + g(1-b)}$$
$$r^* = \frac{-(1-b)(M_s - M_0) + f(C_0 + I_0 + G)}{fa + g(1-b)}.$$

2. The augmented matrix is

$$\begin{pmatrix} 1 & 1 & 1 & 2 \\ 2 & 1 & 2 & 5 \\ 4 & 3 & 4 & c \end{pmatrix}.$$

Performing row operations to reduce this to echelon form,

$$\begin{pmatrix} 1 & 1 & 1 & 2 \\ 2 & 1 & 2 & 5 \\ 4 & 3 & 4 & c \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 1 & 2 \\ 0 & -1 & 0 & 1 \\ 0 & -1 & 0 & c-8 \end{pmatrix}$$
$$\to \begin{pmatrix} 1 & 1 & 1 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & 8-c \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 1 & 2 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 9-c \end{pmatrix}.$$

Looking at the bottom row of the last matrix, we see that the system of equations is inconsistent unless $9 - c = 0$; that is, it is consistent only when $c = 9$. When $c = 9$, the system is equivalent to

$$x_1 + x_2 + x_3 = 2$$
$$x_2 = -1.$$

Setting $x_3 = s$, we have $x_2 = -1$ and

$$x_1 = 2 - x_2 - x_3 = 3 - s.$$

The general solution when $c = 9$ is therefore

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 3-s \\ -1 \\ s \end{pmatrix}.$$

3. Let us denote the weekly production levels of Abracadabra, Break and Choca-mocha bars by a, b, c (respectively), measured in thousands of bars. Then, since 5 units of cocoa are needed to produce one thousand bars of each type, and since 250 units of cocoa are used each week, we must have

$$5a + 5b + 5c = 250.$$

By considering the distribution of the 150 available units of milk,

$$3a + 4b + 2c = 150.$$

Similarly, since 100 units of coffee are used, it must be the case that

$$2a + b + 3c = 100.$$

In other words, we have the system of equations

$$\begin{aligned} 5a + 5b + 5c &= 250 \\ 3a + 4b + 2c &= 150 \\ 2a + b + 3c &= 100. \end{aligned}$$

We solve this using row operations, starting with the augmented matrix, as follows.

$$\begin{aligned} \begin{pmatrix} 5 & 5 & 5 & 250 \\ 3 & 4 & 2 & 150 \\ 2 & 1 & 3 & 100 \end{pmatrix} &\to \begin{pmatrix} 1 & 1 & 1 & 50 \\ 3 & 4 & 2 & 150 \\ 2 & 1 & 3 & 100 \end{pmatrix} \\ &\to \begin{pmatrix} 1 & 1 & 1 & 50 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \end{pmatrix} \\ &\to \begin{pmatrix} 1 & 1 & 1 & 50 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

Therefore, the system is equivalent to

$$\begin{aligned} a + b + c &= 50 \\ b - c &= 0, \end{aligned}$$

from which we obtain $b = c$ and $a = 50 - b - c = 50 - 2c$. In other words, the weekly production levels of Break and Choca-mocha bars are equal, and the production of Abracadabra is (in thousands) $50 - 2c$ where c is the production of Choca-mocha (and Break) bars. Clearly, none of a, b, c can be negative, so the production level, c, of Choca-mocha bars must be such that $a = 50 - 2c \geq 0$; that is, $c \leq 25$. Therefore, the maximum number of Choca-mocha bars which it is possible to manufacture in a week is 25000 (in which case the same number of Break bars are produced and no Abracadabra bars will be manufactured).

4. We start with the matrix

$$(A \mid I) = \left(\begin{array}{ccc|ccc} 3 & 2 & -1 & 1 & 0 & 0 \\ 1 & 7 & 5 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 & 0 & 1 \end{array}\right)$$

and reduce this, using row operations, to $(I \mid A^{-1})$. We have

$$\left(\begin{array}{ccc|ccc} 3 & 2 & -1 & 1 & 0 & 0 \\ 1 & 7 & 5 & 0 & 1 & 0 \\ -1 & 0 & 1 & 0 & 0 & 1 \end{array}\right) \to \left(\begin{array}{ccc|ccc} -1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 7 & 5 & 0 & 1 & 0 \\ 3 & 2 & -1 & 1 & 0 & 0 \end{array}\right)$$

$$\to \left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 0 & 0 & -1 \\ 1 & 7 & 5 & 0 & 1 & 0 \\ 3 & 2 & -1 & 1 & 0 & 0 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 0 & 0 & -1 \\ 0 & 7 & 6 & 0 & 1 & 1 \\ 0 & 2 & 2 & 1 & 0 & 3 \end{array}\right)$$

$$\to \left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 0 & 0 & -1 \\ 0 & 2 & 2 & 1 & 0 & 3 \\ 0 & 7 & 6 & 0 & 1 & 1 \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 0 & 0 & -1 \\ 0 & 1 & 1 & \frac{1}{2} & 0 & \frac{3}{2} \\ 0 & 7 & 6 & 0 & 1 & 1 \end{array}\right)$$

$$\to \left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 0 & 0 & -1 \\ 0 & 1 & 1 & \frac{1}{2} & 0 & \frac{3}{2} \\ 0 & 0 & -1 & -\frac{7}{2} & 1 & -\frac{19}{2} \end{array}\right) \to \left(\begin{array}{ccc|ccc} 1 & 0 & -1 & 0 & 0 & -1 \\ 0 & 1 & 1 & \frac{1}{2} & 0 & \frac{3}{2} \\ 0 & 0 & 1 & \frac{7}{2} & -1 & \frac{19}{2} \end{array}\right)$$

$$\rightarrow \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{7}{2} & -1 & \frac{17}{2} \\ 0 & 1 & 0 & -3 & 1 & -8 \\ 0 & 0 & 1 & \frac{7}{2} & -1 & \frac{19}{2} \end{array}\right).$$

So

$$A^{-1} = \begin{pmatrix} \frac{7}{2} & -1 & \frac{17}{2} \\ -3 & 1 & -8 \\ \frac{7}{2} & -1 & \frac{19}{2} \end{pmatrix}.$$

(Of course, there is nothing special about the sequence of row operations used here: there are many possible routes to the required final matrix.) Another approach is to determine $\det(A)$ and the adjugate matrix $\text{adj}(A)$ and then use the fact that

$$A^{-1} = \frac{1}{\det(A)}\text{adj}(A).$$

This will, of course, give the same answer!

5. Suppose we were to try to use the determinant-based method to determine the inverse. The first thing we'd do (before finding the adjugate matrix) is to calculate the determinant of the matrix. Let's call the matrix A. Then

$$\begin{aligned} \det(A) &= 1\begin{vmatrix} 3 & 1 \\ 0 & 14 \end{vmatrix} - (1)\begin{vmatrix} 2 & 1 \\ 1 & 14 \end{vmatrix} + 5\begin{vmatrix} 2 & 3 \\ 1 & 0 \end{vmatrix} \\ &= 1(42) - (27) + 5(-3) \\ &= 0. \end{aligned}$$

We can stop now and conclude that the matrix has no inverse, because we know that a matrix is invertible if and only if it has non-zero determinant.

Alternatively, we can use row operations. We start with the 3×6 matrix

$$(A \mid I) = \left(\begin{array}{ccc|ccc} 1 & 1 & 5 & 1 & 0 & 0 \\ 2 & 3 & 1 & 0 & 1 & 0 \\ 1 & 0 & 14 & 0 & 0 & 1 \end{array}\right)$$

and we reduce this as follows:

$$\left(\begin{array}{ccc|ccc} 1 & 1 & 5 & 1 & 0 & 0 \\ 2 & 3 & 1 & 0 & 1 & 0 \\ 1 & 0 & 14 & 0 & 0 & 1 \end{array}\right) \rightarrow \left(\begin{array}{ccc|ccc} 1 & 1 & 5 & 1 & 0 & 0 \\ 0 & 1 & -9 & -2 & 1 & 0 \\ 0 & -1 & 9 & -1 & 0 & 1 \end{array}\right)$$

$$\rightarrow \left(\begin{array}{ccc|ccc} 1 & 1 & 5 & 1 & 0 & 0 \\ 0 & 1 & -9 & -2 & 1 & 0 \\ 0 & 0 & 0 & -3 & 1 & 1 \end{array}\right).$$

There is no need to proceed further. Since this last matrix has a row whose first half is all-zero, we deduce that A is not invertible.

6. Let x_1 be the amount of coal production required (in dollars), x_2 the amount of electricity and x_3 the amount of transportation. Then, by considering in turn the total demand for each of the three commodities,

$$\begin{aligned} x_1 &= 50000 + 0.65x_2 + 0.55x_3 \\ x_2 &= 25000 + 0.25x_1 + 0.05x_2 + 0.10x_3 \\ x_3 &= 0.25x_1 + 0.05x_2. \end{aligned}$$

Therefore, we have

$$\begin{pmatrix} 1 & -0.65 & -0.55 \\ -0.25 & 0.95 & -0.10 \\ -0.25 & -0.05 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_2 \end{pmatrix} = \begin{pmatrix} 50000 \\ 25000 \\ 0 \end{pmatrix}.$$

(The question did not ask you to solve the equations (which is a fairly messy business), but you can find a solution in Anthony and Biggs, Worked Example 19.3 of Chapter 19.)

7. Denoting the matrix by A, we have

$$\begin{aligned} |A - \lambda I| &= \begin{vmatrix} 2-\lambda & 1 \\ 3 & 1-\lambda \end{vmatrix} \\ &= (2-\lambda)(1-\lambda) - 3 \\ &= \lambda^2 - 3\lambda + 2 - 3 \\ &= \lambda^2 - 3\lambda - 1. \end{aligned}$$

To determine the eigenvalues, we therefore seek the solutions to $\lambda^2 - 3\lambda - 1 = 0$. By the formula for the roots of a quadratic, we have eigenvalues

$$\lambda = \frac{3 \pm \sqrt{13}}{2},$$

so the eigenvalues are $(3 + \sqrt{13})/2$ and $(3 - \sqrt{13})/2$.

8. Denoting the matrix by A, we have

$$\begin{aligned} |A - \lambda I| &= \begin{vmatrix} 2-\lambda & 4 \\ 3 & 3-\lambda \end{vmatrix} \\ &= (2-\lambda)(3-\lambda) - 12 \\ &= \lambda^2 - 5\lambda - 6. \end{aligned}$$

To determine the eigenvalues, we therefore seek the solutions to $\lambda^2 - 3\lambda - 1 = 0$. You can use the formula for the solutions of a quadratic, but in this case it's quite easy to spot that the quadratic factorises as $(\lambda - 6)(\lambda + 1)$, so that the eigenvalues are 6 and -1.

To find an eigenvector for eigenvalue 6 we solve the system $(A - 6I)\mathbf{x} = \mathbf{0}$: that is,

$$\begin{pmatrix} -4 & 4 \\ 3 & -3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

This is clearly equivalent to the single equation $x_1 = x_2$, so a suitable eigenvector is $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

To find an eigenvector for eigenvalue -1 we solve the system $(A - (-1)I)\mathbf{x} = \mathbf{0}$: that is,

$$\begin{pmatrix} 3 & 4 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

This is equivalent to the single equation $3x_1 + 4x_2 = 0$, so a suitable eigenvector is (taking $x_1 = 4$) $\begin{pmatrix} 4 \\ -3 \end{pmatrix}$.

To diagonalise A, we let P be the matrix whose columns are these eigenvectors,

$$P = \begin{pmatrix} 1 & 4 \\ 1 & -3 \end{pmatrix}.$$

Then $P^{-1}AP = D$, where D is the diagonal matrix

$$D = \begin{pmatrix} 6 & 0 \\ 0 & -1 \end{pmatrix}.$$