A.G. Hamilton

LINEAR ALGEBRA

An introduction with concurrent examples

linear algebra

AN INTRODUCTION
WITH CONCURRENT EXAMPLES

A. G. HAMILTON
Department of Computing Science, University of Stirling

CAMBRIDGE
UNIVERSITY PRESS

Published by the Press Syndicate of the University of Cambridge
The Pitt Building, Trumpington Street, Cambridge CB2 1RP
40 West 20th Street, New York, NY 10011-4211 USA
10 Stamford Road, Oakleigh, Melbourne 3166, Australia

First published 1989
Reprinted 1992, 1994

British Library cataloguing in publication data
Hamilton, A. G. (Alan G.)
Linear algebra.
1. Linear algebra.
I. Title II. Hamilton, A. G. (Alan G.)
Linear algebra: an introduction with
concurrent examples
512'.5

Library of Congress cataloguing in publication data
Hamilton, A. G., 1943–
Linear algebra: an introduction with concurrent examples / A. G. Hamilton
p. cm.
Includes index.
1. Algebras, Linear. I. Title.
QA184.H362 1989
512.5—dc19 88-31177CIP

ISBN 0 521 32517 X hard covers
ISBN 0 521 31042 3 paperback

Transferred to digital printing 1999

Linear algebra

PREFACE

My earlier book, *A First Course in Linear Algebra with Concurrent Examples* (referred to below as the First Course), was an introduction to the use of vectors and matrices in the solution of sets of simultaneous linear equations and in the geometry of two and three dimensions. As its name suggests, that much is only a start. For many readers, such elementary material may satisfy the need for appropriate mathematical tools. But, for others, more advanced techniques may be required, or, indeed, further study of algebra for its own sake may be the objective.

This book is therefore in the literal sense an extension of the First Course. The first eleven chapters are identical to the earlier book. The remainder forms a sequel: a continuation into the next stage of the subject. This aims to provide a practical introduction to perhaps the most important applicable idea of linear algebra, namely eigenvalues and eigenvectors of matrices. This requires an introduction to some general ideas about vector spaces. But this is not a book about vector spaces in the abstract. The notions of subspace, basis and dimension are all dealt with in the concrete context of n-dimensional real Euclidean space. Much attention is paid to the diagonalisation of real symmetric matrices, and the final two chapters illustrate applications to geometry and to differential equations.

The organisation and presentation of the content of the First Course were unusual. This book has the same features, and for the same reasons. These reasons were described in the preface to the First Course in the following four paragraphs, which apply equally to this extended volume.

'Learning is not easy (not for most people, anyway). It is, of course, aided by being taught, but it is by no means only a passive exercise. One who hopes to learn must work at it actively. My intention in writing this book is not to teach, but rather to provide a stimulus and a medium through which a reader can learn. There are various sorts of textbook with widely differing approaches. There is the encyclopaedic sort, which tends to be unreadable but contains all of the information relevant to its subject. And at the other extreme there is the work-book, which leads the reader in a progressive series of exercises. In the field of linear algebra there are already enough books of the former kind, so this book is aimed away from that end of the spectrum. But it is not a work-book, neither is it comprehensive. It is a book to be worked through, however. It is intended to be read, not referred to.

'Of course, in a subject such as this, reading is not enough. Doing is also necessary. And doing is one of the main emphases of the book. It is about methods and their application. There are three aspects of this provided by this book: description, worked examples and exercises. All three are important, but I would stress that the most important of these is the exercises. You do not know it until you can do it.

'The format of the book perhaps requires some explanation. The worked examples are integrated with the text, and the careful reader will follow the examples through at the same time as reading the descriptive material. To facilitate this, the text appears on the right-hand pages only, and the examples on the left-hand pages. Thus the text and corresponding examples are visible simultaneously, with neither interrupting the other. Each chapter concludes with a set of exercises covering specifically the material of that chapter. At the end of the book there is a set of sample examination questions covering the material of the whole book.

'The prerequisites required for reading this book are few. It is an introduction to the subject, and so requires only experience with methods of arithmetic, simple algebra and basic geometry. It deliberately avoids mathematical sophistication, but it presents the basis of the subject in a way which can be built on subsequently, either with a view to applications or with the development of the abstract ideas as the principal consideration.'

Last, this book would not have been produced had it not been for the advice and encouragement of David Tranah of Cambridge University Press. My thanks go to him, and to his anonymous referees, for many helpful comments and suggestions.

Part 1

Examples

1.1 Simple elimination (two equations).

$$2x+3y=1$$
$$x-2y=4.$$

Eliminate x as follows. Multiply the second equation by 2:

$$2x+3y=1$$
$$2x-4y=8.$$

Now replace the second equation by the equation obtained by subtracting the first equation from the second:

$$2x+3y=1$$
$$-7y=7.$$

Solve the second equation for y, giving $y=-1$. Substitute this into the first equation:

$$2x-3=1,$$

which yields $x=2$. Solution: $x=2$, $y=-1$.

1.2 Simple elimination (three equations).

$$x-2y+z=5$$
$$3x+y-z=0$$
$$x+3y+2z=2.$$

Eliminate z from the first two equations by adding them:

$$4x-y=5.$$

Next eliminate z from the second and third equations by adding twice the second to the third:

$$7x+5y=2.$$

Now solve the two simultaneous equations:

$$4x-y=5$$
$$7x+5y=2$$

as in Example 1.1. One way is to add five times the first to the second, obtaining

$$27x=27,$$

so that $x=1$. Substitute this into one of the set of two equations above which involve only x and y, to obtain (say)

$$4-y=5,$$

so that $y=-1$. Last, substitute $x=1$ and $y=-1$ into one of the original equations, obtaining

$$1+2+z=5,$$

so that $z=2$. Solution: $x=1$, $y=-1$, $z=2$.

CONTENTS

Part 2

1

Gaussian elimination

We shall describe a standard procedure which can be used to solve sets of simultaneous linear equations, no matter how many equations. Let us make sure of what the words mean before we start, however. A *linear* equation is an equation involving unknowns called x or y or z, or x_1 or x_2 or x_3, or some similar labels, in which the unknowns all occur to the first degree, which means that no squares or cubes or higher powers, and no products of two or more unknowns, occur. To *solve* a set of simultaneous equations is to find all values or sets of values for the unknowns which satisfy the equations.

Given two linear equations in unknowns x and y, as in Example 1.1, the way to proceed is to *eliminate* one of the unknowns by combining the two equations in the manner shown.

Given three linear equations in three unknowns, as in Example 1.2, we must proceed in stages. First eliminate one of the unknowns by combining two of the equations, then similarly eliminate the same unknown from a different pair of the equations by combining the third equation with one of the others. This yields two equations with two unknowns. The second stage is to solve these two equations. The third stage is to find the value of the originally eliminated unknown by substituting into one of the original equations.

This general procedure will extend to deal with n equations in n unknowns, no matter how large n is. First eliminate one of the unknowns, obtaining $n-1$ equations in $n-1$ unknowns, then eliminate another unknown from these, giving $n-2$ equations in $n-2$ unknowns, and so on until there is one equation with one unknown. Finally, substitute back to find the values of the other unknowns.

There is nothing intrinsically difficult about this procedure. It consists of the application of a small number of simple operations, used repeatedly.

1.3 The Gaussian elimination process.

$$2x_1 - x_2 + 3x_3 = 1 \qquad (1)$$
$$4x_1 + 2x_2 - x_3 = -8 \qquad (2)$$
$$3x_1 + x_2 + 2x_3 = -1 \qquad (3)$$

Stage 1:
$$x_1 - \tfrac{1}{2}x_2 + \tfrac{3}{2}x_3 = \tfrac{1}{2} \qquad (1)\div 2$$
$$4x_1 + 2x_2 - x_3 = -8 \qquad (2)$$
$$3x_1 + x_2 + 2x_3 = -1 \qquad (3)$$

Stage 2:
$$x_1 - \tfrac{1}{2}x_2 + \tfrac{3}{2}x_3 = \tfrac{1}{2} \qquad (1)$$
$$4x_2 - 7x_3 = -10 \qquad (2) - 4\times(1)$$
$$\tfrac{5}{2}x_2 - \tfrac{5}{2}x_3 = -\tfrac{5}{2} \qquad (3) - 3\times(1)$$

Stage 3:
$$x_1 - \tfrac{1}{2}x_2 + \tfrac{3}{2}x_3 = \tfrac{1}{2} \qquad (1)$$
$$x_2 - \tfrac{7}{4}x_3 = -\tfrac{5}{2} \qquad (2)\div 4$$
$$\tfrac{5}{2}x_2 - \tfrac{5}{2}x_3 = -\tfrac{5}{2} \qquad (3)$$

Stage 4:
$$x_1 - \tfrac{1}{2}x_2 + \tfrac{3}{2}x_3 = \tfrac{1}{2} \qquad (1)$$
$$x_2 - \tfrac{7}{4}x_3 = -\tfrac{5}{2} \qquad (2)$$
$$\tfrac{15}{8}x_3 = \tfrac{15}{4} \qquad (3) - \tfrac{5}{2}\times(2)$$

Stage 5:
$$x_1 - \tfrac{1}{2}x_2 + \tfrac{3}{2}x_3 = \tfrac{1}{2} \qquad (1)$$
$$x_2 - \tfrac{7}{4}x_3 = -\tfrac{5}{2} \qquad (2)$$
$$x_3 = 2. \qquad (3)\div\tfrac{15}{8}$$

Now we may obtain the solutions. Substitute $x_3 = 2$ into the second equation.

$$x_2 - \tfrac{7}{2} = -\tfrac{5}{2}, \quad \text{so } x_2 = 1.$$

Finally substitute both into the first equation, obtaining

$$x_1 - \tfrac{1}{2} + 3 = \tfrac{1}{2}, \quad \text{so } x_1 = -2.$$

Hence the solution is $x_1 = -2$, $x_2 = 1$, $x_3 = 2$.

These include multiplying an equation through by a number and adding or subtracting two equations. But, as the number of unknowns increases, the length of the procedure and the variety of different possible ways of proceeding increase dramatically. Not only this, but it may happen that our set of equations has some special nature which would cause the procedure as given above to fail: for example, a set of simultaneous equations may be *inconsistent*, i.e. have no solution at all, or, at the other end of the spectrum, it may have many different solutions. It is useful, therefore, to have a standard routine way of organising the elimination process which will apply for large sets of equations just as for small, and which will cope in a more or less automatic way with special situations. This is necessary, in any case, for the solution of simultaneous equations using a computer. Computers can handle very large sets of simultaneous equations, but they need a routine process which can be applied automatically. One such process, which will be used throughout this book, is called *Gaussian elimination*. The best way to learn how it works is to follow through examples, so Example 1.3 illustrates the stages described below, and the descriptions should be read in conjunction with it.

Stage 1 Divide the first equation through by the coefficient of x_1. (If this coefficient happens to be zero then choose another of the equations and place it first.)

Stage 2 Eliminate x_1 from the second equation by subtracting a multiple of the first equation from the second equation. Eliminate x_1 from the third equation by subtracting a multiple of the *first* equation from the third equation.

Stage 3 Divide the second equation through by the coefficient of x_2. (If this coefficient is zero then interchange the second and third equations. We shall see later how to proceed if neither of the second and third equations contains a term in x_2.)

Stage 4 Eliminate x_2 from the third equation by subtracting a multiple of the second equation.

Stage 5 Divide the third equation through by the coefficient of x_3. (We shall see later how to cope if this coefficient happens to be zero.)

At this point we have completed the elimination process. What we have done is to find another set of simultaneous equations which have the same solutions as the given set, and whose solutions can be read off very easily. What remains to be done is the following.

> Read off the value of x_3. Substitute this value in the second equation, giving the value of x_2. Substitute both values in the first equation, to obtain the value of x_1.

1.4 Using arrays, solve the simultaneous equations:

$$\begin{aligned} x_1+x_2-\ x_3&=\ 4\\ 2x_1-x_2+3x_3&=\ 7\\ 4x_1+x_2+\ x_3&=15. \end{aligned}$$

First start with the array of coefficients:

$$\begin{array}{rrrr} 1 & 1 & -1 & 4\\ 2 & -1 & 3 & 7\\ 4 & 1 & 1 & 15\\ \hline 1 & 1 & -1 & 4\\ 0 & -3 & 5 & -1\\ 0 & -3 & 5 & -1\\ \hline 1 & 1 & -1 & 4\\ 0 & 1 & -\frac{5}{3} & \frac{1}{3}\\ 0 & -3 & 5 & -1\\ \hline 1 & 1 & -1 & 4\\ 0 & 1 & -\frac{5}{3} & \frac{1}{3}\\ 0 & 0 & 0 & 0\\ \hline \end{array}
\qquad
\begin{array}{l} \\ \\ \\ \\ (2)-2\times(1)\\ (3)-4\times(1)\\ \\ (2)\div-3\\ \\ \\ \\ (3)+3\times(2) \end{array}$$

See Chapter 2 for discussion of how solutions are obtained from here.

1.5 Using arrays, solve the simultaneous equations:

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

What follows is a full solution.

$$\begin{array}{rrrr} 3 & -3 & 1 & 1\\ -1 & 1 & 2 & 2\\ 2 & 1 & -3 & 0\\ \hline 1 & -1 & \frac{1}{3} & \frac{1}{3}\\ -1 & 1 & 2 & 2\\ 2 & 1 & -3 & 0\\ \hline 1 & -1 & \frac{1}{3} & \frac{1}{3}\\ 0 & 0 & \frac{7}{3} & \frac{7}{3}\\ 0 & 3 & -\frac{11}{3} & -\frac{2}{3}\\ \hline 1 & -1 & \frac{1}{3} & \frac{1}{3}\\ 0 & 3 & -\frac{11}{3} & -\frac{2}{3}\\ 0 & 0 & \frac{7}{3} & \frac{7}{3}\\ \hline \end{array}
\qquad
\begin{array}{l} \\ \\ \\ (1)\div 3\\ \\ \\ \\ (2)+(1)\\ (3)-2\times(1)\\ \\ \left.\begin{array}{l}\\ \\ \end{array}\right\}\text{interchange rows} \end{array}$$

Notice that after stage 1 the first equation is not changed, and that after stage 3 the second equation is not changed. This is a feature of the process, however many equations there are. We proceed downwards and eventually each equation is fixed in a new form.

Besides the benefit of standardisation, there is another benefit which can be derived from this process, and that is brevity. Our working of Example 1.3 includes much that is not essential to the process. In particular the repeated writing of equations is unnecessary. Our standard process can be developed so as to avoid this, and all of the examples after Example 1.3 show the different form. The sets of equations are represented by arrays of coefficients, suppressing the unknowns and the equality signs. The first step in Example 1.4 shows how this is done. Our operations on equations now become operations on the rows of the array. These are of the following kinds:

- interchange rows,
- divide (or multiply) one row through by a number,
- subtract (or add) a multiple of one row from (to) another.

These are called *elementary row operations*, and they play a large part in our later work. It is important to notice the form of the array at the end of the process. It has a triangle of 0s in the lower left corner and 1s down the diagonal from the top left.

Now let us take up two complications mentioned above. In stage 5 of the Gaussian elimination process (henceforward called the GE process) the situation not covered was when the coefficient of x_3 in the third equation (row) was zero. In this case we divide the third equation (row) by the number occurring on the right-hand side (in the last column), if this is not already zero. Example 1.4 illustrates this. The solution of sets of equations for which this happens will be discussed in the next chapter. What happens is that either the equations have no solutions or they have infinitely many solutions.

The other complication can arise in stage 3 of the GE process. Here the coefficient of x_2 may be zero. The instruction was to interchange equations (rows) in the hope of placing a non-zero coefficient in this position. When working by hand we may choose which row to interchange with so as to make the calculation easiest (presuming that there is a choice). An obvious way to do this is to choose a row in which this coefficient is 1. Example 1.5 shows this being done. When the GE process is formalised (say for computer application), however, we need a more definite rule, and the one normally adopted is called *partial pivoting*. Under this rule, when we interchange rows because of a zero coefficient, we choose to interchange with the row which has the coefficient which is numerically the *largest* (that

1	−1	$\frac{1}{3}$	$\frac{1}{3}$	
0	1	$-\frac{11}{9}$	$-\frac{2}{9}$	$(2)\div 3$
0	0	1	1	$(3)\div\frac{7}{3}$

From here, $x_3=1$, and substituting back we obtain

$$x_2-\tfrac{11}{9}=-\tfrac{2}{9},\quad \text{so } x_2=1.$$

Substituting again:

$$x_1-1+\tfrac{1}{3}=\tfrac{1}{3},\quad \text{so } x_1=1.$$

Hence the solution sought is: $x_1=1$, $x_2=1$, $x_3=1$.

1.6 Using arrays, solve the simultaneous equations:

$$\begin{aligned} x_1+x_2-x_3&=-3\\ 2x_1+2x_2+x_3&=0\\ 5x_1+5x_2-3x_3&=-8. \end{aligned}$$

Solution:

1	1	−1	−3	
2	2	1	0	
5	5	−3	−8	
1	1	−1	−3	
0	0	3	6	$(2)-2\times(1)$
0	0	2	7	$(3)-5\times(1)$
1	1	−1	−3	
0	0	1	2	$(2)\div 3$
0	0	2	7	
1	1	−1	−3	
0	0	1	2	
0	0	0	3	$(3)-2\times(2)$

Next, and finally, divide the last row by 3. How to obtain solutions from this point is discussed in Chapter 2. (In fact there are no solutions in this case.)

1.7 Solve the simultaneous equations:

$$\begin{aligned} 2x_1-2x_2+x_3-3x_4&=2\\ x_1-x_2+3x_3-x_4&=-2\\ -x_1-2x_2+x_3+2x_4&=-6\\ 3x_1+x_2-x_3-2x_4&=7. \end{aligned}$$

Convert to an array and proceed:

2	−2	1	−3	2
1	−1	3	−1	−2
−1	−2	1	2	−6
3	1	−1	−2	7

is, the largest when any negative signs are disregarded). This has two benefits. First, we (and more particularly, the computer) know precisely what to do at each stage and, second, following this process actually produces a more accurate answer when calculations are subject to rounding errors, as will always be the case with computers. Generally, we shall not use partial pivoting, since our calculations will all be done by hand with small-scale examples.

There may be a different problem at stage 3. We may find that there is no equation (row) which we can choose which has a non-zero coefficient in the appropriate place. In this case we do nothing, and just move on to consideration of x_3, as shown in Example 1.6. How to solve the equations in such a case is discussed in the next chapter.

The GE process has been described above in terms which can be extended to cover larger sets of equations (and correspondingly larger arrays of coefficients). We should bear in mind always that the form of the array which we are seeking has rows in which the first non-zero coefficient (if there is one) is 1, and this 1 is to the *right* of the first non-zero coefficient in the preceding row. Such a form for an array is called *row-echelon form*. Example 1.7 shows the process applied to a set of four equations in four unknowns.

Further examples of the GE process applied to arrays are given in the following exercises. Of course the way to learn this process is to carry it out, and the reader is recommended not to proceed to the rest of the book before gaining confidence in applying it.

Summary

The purpose of this chapter is to describe the Gaussian elimination process which is used in the solution of simultaneous equations, and the abbreviated way of carrying it out, using elementary row operations on rectangular arrays.