Basic Linear Algebr

Thomas S. Blyth, Edmund F. Robertson

T. S. BIyth and E. F. Robertson Basic Linear Algebra Second Edition SPRINGER UNDERGRADUATE MATHEMATICS 5 Springer V* SERIES

Springer Undergraduate Mathematics Series

Calculus P.C. Matthews Worlds Out of Nothing: A Course in the History of Geometry in the 19th Century J. Gray

T.S. Blyth and E.F. Robertson Basic Linear Algebra Second Edition Springer

Preface The word 'basic* in the title of this text could be substituted by 'elementary' or by 'an introduction to'; such are the contents. We have chosen the word 'basic' in order to emphasise our objective, which is to provide in a reasonably compact and readable form a rigorous first course that covers all of the material on linear algebra to which every student of mathematics should be exposed at an early stage. By developing the algebra of matrices before proceeding to the abstract notion of a vector space, we present the pedagogical progression as a smooth transition from the computational to the general, from the concrete to the abstract. In so doing we have included more than 125 illustrative and worked examples, these being presented immediately following definitions and new results. We have also included more than 300 exercises. In order to consolidate the student's understanding, many of these appear strategically placed throughout the text. They are ideal for self-tutorial purposes. Supplementary exercises are grouped at the end of each chapter. Many of these are 'cumulative' in the sense that they require a knowledge of material covered in previous chapters. Solutions to the exercises are provided at the conclusion of the text. In preparing this second edition we decided to take the opportunity of including, as in our companion volume Further Linear Algebra in this series, a chapter that gives a brief introduction to the use of MAPLE1 in dealing with numerical and algebraic problems in linear algebra.

We have also included some additional exercises at the end of each chapter. No solutions are provided for these as they are intended for assignment purposes. T.S.B., E.F.R. 1 MAPLE™ is a registered trademark of Waterloo Maple Inc., 57 Erb Street West. Waterloo. Ontario. Canada. N2L6C2. www.maplesoft.com

Foreword The early development of matrices on the one hand, and linear spaces on the other, was occasioned by the need to solve specific problems, not only in mathematics but also in other branches of science. It is fair to say that the first known example of matrix methods is in the text Nine Chapters of the Mathematical Art written during the Han Dynasty. Here the following problem is considered: There are three types of com, of which three bundles of the first, two bundles of the second, and one of the third make 39 measures. Two of the first, three of the second, and one of the third make 34 measures. And one of the first, two of the second, and three of the third make 26 measures. How many measures of com are contained in one bundle of each type! In considering this problem the author, writing in 200BC, does something that is quite remarkable. He sets up the coefficients of the system of three linear equations in three unknowns as a table on a 'counting board' 1 2 3 2 3 2 3 1 1 26 34 39 the middle column by 3 and subtract the right The same instruction applied in respect of the 0 0 3 4 5 2 8 1 1 39 24 39 Next, the leftmost column is multiplied by 5 and the middle column subtracted from and instructs the reader to multiply column as many times as possible. first column gives

viii Basic Linear Algebra it as many times as possible, giving 0 0 3 0 5 2 36 1 1 99 24 39 from which the solution can now be found for the third type of corn, then for the second and finally the first by back substitution. This method, now sometimes known as gaussian elimination, would not become well-known until the 19th Century. The idea of a determinant first appeared in Japan in 1683 when Seki published his Method of solving the dissimulated problems which contains matrix methods written as tables like the Chinese method described above. Using his 'determinants' (he had no word for them), Seki was able to compute the determinants of 5 x 5 matrices and apply his techniques to the solution of equations. Somewhat remarkably, also in 1683, Leibniz explained in a letter to de 1'Hopital that the system of equations $10 + 11x + 12? = 0$ $20 + 2\backslash x + 22? = 0$ $30 + 3\backslash x + 32? = 0$ has a solution if $10.21.32 + 11.22.30 +$

12.20.31 = 10.22.31 + 11.20.32+ 12.21.30. Bearing in mind that Leibniz was not using numerical coefficients but rather two characters, the first marking in which equation it occurs, the second marking which letter it belongs to we see that the above condition is precisely the condition that the coefficient matrix has determinant 0. Nowadays we might write, for example, a2x for 21 in the above. The concept of a vector can be traced to the beginning of the 19th Century in the work of Bolzano. In 1804 he published Betrachtungen iiber einige Gegenstande der Elementargeometrie in which he considers points, lines and planes as undefined objects and introduces operations on them. This was an important step in the axiomati- sation of geometry and an early move towards the necessary abstraction required for the later development of the concept of a linear space. The first axiomatic definition of a linear space was provided by Peano in 1888 when he published Calcolo geo- metrico secondo I'Ausdehnungslehre de H. Grassmann preceduto dalle operazioni della logica deduttiva. Peano credits the work of Leibniz, Mobius, Grassmann and Hamilton as having provided him with the ideas which led to his formal calculus. In this remarkable book, Peano introduces what subsequently took a long time to become standard notation for basic set theory.

Foreword IX Peano's axioms for a linear space are 1. a = b if and only if b = a, if a = b and b = c then a = c. 2. The sum of two objects a and b is defined, i.e. an object is defined denoted by a + b, also belonging to the system, which satisfies Ifa-b then a + c = b + c,a + b = b + a, a + (b + c)={a + b) + c, and the common value of the last equality is denoted by a + b + c. 3. If a is an object of the system and m a positive integer, then we understand by ma the sum of m objects equal to a. It is easy to see that for objects a,b,... of the system and positive integers m, n,... one has Ifa-b then ma - mb, m(a + b) = ma + mb, (m + n)a = ma + na, m(na) = mna, \a = a. We suppose that for any real number m the notation ma has a meaning such that the preceding equations are valid. Peano also postulated the existence of a zero object 0 and used the notation a-b for a + (-b). By introducing the notions of dependent and independent objects, he defined the notion of dimension, showed that finite-dimensional spaces have a basis and gave examples of infinite-dimensional linear spaces. If one considers only functions of degree n, then these functions form a linear system with n + 1 dimensions, the entire functions of arbitrary degree form a linear system with infinitely many dimensions. Peano also introduced linear operators on a linear space and showed that by using coordinates one obtains a

matrix. With the passage of time, much concrete has set on these foundations. Techniques and notation have become more refined and the range of applications greatly enlarged. Nowadays Linear Algebra, comprising matrices and vector spaces, plays a major role in the mathematical curriculum. Notwithstanding the fact that many important and powerful computer packages exist to solve problems in linear algebra, it is our contention that a sound knowledge of the basic concepts and techniques is essential.

Contents Preface Foreword

1 The Algebra of Matrices If m and n are positive integers then by a matrix of size m by n, or an m x n matrix, we shall mean a rectangular array consisting of mn numbers in a boxed display consisting of m rows and n columns. Simple examples of such objects are the following: size 1x5 : [lO 9 8 7 6] size 3x2: size 4x4 12 3 4 2 3 4 5 3 4 5 6 4 5 6 7 size 3 x 1 1 2 3 4 5 6 In general we shall display anmxn matrix as *11 *21 *31 *m\ x\2 *13 • *22 *23 • *32 *33 • Xm2 Xm3 • •• *1,, • *2n • *3n • *mn. • Note that the first suffix gives the number of the row and the second suffix that of the column, so that xtj appears at the intersection of the i-th row and the j-\h column. We shall often find it convenient to abbreviate the above display to simply and refer to xi} as the (i,;)-th element or the (i,;")-th entry of the matrix. • Thus the expression X = [Xy]mxn will be taken to mean that 'X is the m x n matrix whose (ij)-th element is x^.

2 Basic Linear Algebra Example 1.1 The 3x3 matrix X = Example 1.2 The 3 x 3 matrix X = n 2 3 1 22 32 r 23 33 = /I can be expressed as X = [Xy]3x3 where x^ - i a 0 0 a a 0 a a a ;■{ can be expressed as X = [x^]3x3 where a ifi<;; 0 otherwise. Example 1.3 The n x n matrix X = 10 0. e 1.0. e2 <? 1 . en-\ en-2 en-l ..0 ..0 ..0 .. 1 can be expressed as X = [jty],,™ where fe'-' ifi>;; \ 0 otherwi EXERCISES 1.1 Write out the 3 x 3 matrix whose entries are given by xkj = i +;. 1.2 Write out the 3 x 3 matrix whose entries are given by {1 if i +j is even; 0 otherwise. 1.3 Write out the 3 x 3 matrix whose entries are given by xv

= (-1)'"1'. 1.4 Write out the n x n matrix whose entries are given by f 1 T • -l ifi>;; x{j = < 0 ifi=;; I 1 if i<j.

1. The Algebra of Matrices

1.5 Write out the 6 x 6 matrix A = [af|-] in which a^ is given by (1) the least common multiple of i and ;; (2) the greatest common divisor of i andy. 1.6 Given the n x n matrix A = [a^], describe the n x n matrix B = [b^] which is such that bi} = fl,ilI+i_y. Before we can develop an algebra for matrices, it is essential that we decide what is meant by saying that two matrices are equal. Common sense dictates that this should happen only if the matrices in question are of the same size and have corresponding entries equal. Definition If A = [fl/yLx,, and B = [bjj]pxq then we shall say that A and B are equal (and write A = B) if and only if (1) m = pandn= q\ (2) atj = bu for al\i,j. The algebraic system that we shall develop for matrices will have many of the familiar properties enjoyed by the system of real numbers. However, as we shall see, there are some very striking differences. Definition Given m x n matrices A = [a^] and B = [b^], we define the sum A + B to be the mx n matrix whose (ij)-th element is ai} + bVy Note that the sum A + B is defined only when A and B are of the same size; and to obtain this sum we simply add corresponding entries, thereby obtaining a matrix again of the same size. Thus, for instance, -1 [2 0' -2 J + ' 1 2 [-1 Oj ,, 0 [l 2" -2 J Theorem 1.1 Addition of matrices is (1) commutative [in the sense that if A, B are of the same size then we have A + B = B + A); (2) associative [in the sense that if A, B, C are of the same size then we have A + {B + C)={A + B) + C]. Proof (1) If A and B are each of size mx n then A + B and B + A are also of size mx n and by the above definition we have A + B = [aij + bu], B + A = [bij + atj).

4 Basic unearMiueuio Since addition of numbers is commutative we have ai} + bVi = bVi + fly for all ij and so, by the definition of equality for matrices, we conclude that A + B = B + A. (2) If A, B, C are each of size m x n then so are A + (B + C) and (A + B) + C. Now the (/ j)-th element of A + (B + C) is fly + (¾ + ci}) whereas that of {A + fi) + C is (fly + by) + cy. Since addition of numbers is associative we have fly + (¾ + c^,) = (ay+fcy)+cy for all 1,7 and so, by the definition of equality for matrices, we conclude lhalA + {B + C) = {A + B) + C. D Because of Theorem 1.1(2) we agree, as with numbers, to write A + B + C for either A + (5 + C) or (A + B) + C. Theorem 1.2 77i£re « a unique mx n matrix M such that, for every m x n matrix A, A + M = A. Proof Consider the matrix M = [wyLxl, all of whose entries are 0; i.e. mit - 0 for all

ij. For every matrix A = [at-[mytH we have A + M = [fly + mij]mxn = [fl,v + 0]mxn = [flylmxn = A. To establish the uniqueness of this matrix My suppose that B = [bl})mxn is also such that A + B - A for every mx n matrix A. Then in particular we have M + B = M. But, taking fi instead of A in the property for A/, we have fi + M = B. It now follows by Theorem 1.1(1) that B=M. D Definition The unique matrix arising in Theorem 1.2 is called the m x n zero matrix and will be denoted by 0mxn, or simply by 0 if no confusion arises. Theorem 1.3 For every m x n matrix A there is a unique mxn matrix B such that A + B = 0. Proof Given A = latj]mxnt consider B = [-fly-Lxl,, ie. the matrix whose (/,;)-th element is the additive inverse of the (ij)-lh element of A. Clearly, we have >4 + B = [fly+(-fl,7)Lxn = 0. To establish the uniqueness of such a matrix B, suppose that C = [cy]mx,, is also such that A + C = 0. Then for all ij we have fly + Cy = 0 and consequently cl} = -fly which means, by the above definition of equality, that C-B. □ Definition The unique matrix B arising in Theorem 1.3 is called the additive inverse of A and will be denoted by -A. Thus -A is the matrix whose elements are the additive inverses of the corresponding elements of A.

1. The Algebra of Matrices 5 Given numbers jc, y the difference jc - y is defined to be x + (-y). For matrices v4, B of the same size we shall similarly write A - B for A + [-B), the operation '-* so defined being called subtraction of matrices. EXERCISES 1.7 Show that subtraction of matrices is neither commutative nor associative. 1.8 Prove that if A and B are of the same size then -(A + B) = -A - B. x -w y-x z-y y-x z-y w-z So far our matrix algebra has been confined to the operation of addition. This is a simple extension of the same notion for numbers, for we can think of 1 x 1 matrices as behaving essentially as numbers. We shall now investigate how the notion of multiplication for numbers can be extended to matrices. This, however, is not quite so straightforward. There are in fact two distinct 'multiplications' that can be defined. The first 'multiplies' a matrix by a number, and the second 'multiplies' a matrix by another matrix. Definition Given a matrix A and a number \ we define the product of A by \ to be the matrix, denoted by \A, that is obtained from A by multiplying every element of A by \. Thus, if A = [flyLxn then \A = [\au]mxn. This operation is traditionally called multiplying a matrix by a scalar (where the word scalar is taken to be synonymous with number). Such multiplication by scalars may also be thought of as scalars acting on matrices. The principal properties of this action are as follows. Theorem 1.4 If

A and B are m x n matrices then, for any scalars \ and /i, (1) \(A + B)=\A + \B\ (2) (\ +fi)A = \A + iiA; (3) MM) = (\MM; (4) (-1),4 =-* (5)0A = 0mxn. Proof Let A = [fljmxn and B = [bij]mxn. Then the above equalities follow from the observations (1) \(a,,+bu) = \au + \bu. (2) (X + n)ai} = \Oij + naij. 1.9 Simplify x-y y-z z-w w-x x-y y-z

6 Basic Linear Algepra (3) M^;)=(VK- (4) (-\)au = -fl,y. (5)0fl,v = 0. D Observe that for every positive integer n we have nA = A + A + -- + A (n terms). This follows immediately from the definition of the product \A\ for the (i,;)-th element of nA is nau = fly + fly + - • + fly, there being n terms in the summation. EXERCISES 1.10 Given any m x n matrices A and B, solve the matrix equation 3{X + \A) = 5(X - \B). 1.11 Given the matrices "l 0 o" 0 1 0 0 0 1 , *= "l 1 f 1 1 1 1 I 1 solve the matrix equation X + A = 2[X - B). We shall now describe the operation that is called matrix multiplication This is the 'multiplication' of one matrix by another. At first sight this concept (due originally to Cayley) appears to be a most curious one. Whilst it has in fact a very natural interpretation in an algebraic context that we shall see later, we shall for the present simply accept it without asking how it arises. Having said this, however, we shall illustrate its importance in Chapter 2, particularly in the applications of matrix algebra. Definition Let A = [fly]mxn and B = [bij]nxp (note the sizes!). Then we define the product AB to be the m x p matrix whose (i,,/)-th element is [ABlj = aixbXj + a^by + aaby + • ■ + ainbnJ- In other words, the (/',;)-th element of the product AB is obtained by summing the products of the elements in the z'-th row of A with the corresponding elements in the y-th column of B. The above expression for [AB)ti can be written in abbreviated form using the so-called £-notation:

1. The Algebra of Matrices The process for computing products can be pictorially summarised as follows j-th column ofB I'M ** h .»«. —» 0/1*1/ > aabn > <"i3fc3/ ► ainhj n i-th row of A It is important to note that, in forming these sums of products, there are no elements that are *Ieft over' since in the definition of the product AB the number n of columns of A is the same as the number of rows of B. Example 1.4 Consider the matrices A = 0 1 0 2 3 t # = 20 1 2 1 1 The product AB is defined since A is of size 2x3 and B is of size 3x2; moreover, AB is of size 2 x 2. We have AB = 0-2 + 11 +01 00 + 12+0-1 2-2+3-1 + 11 20+3-2 + 11 1 2 8 7 Note that in this case the product BA is also defined (since B has the same number of columns as A has rows). The product

BA is of size 3x3: BA = 2-0+0-2 21 +0-3 20+01 1-0+2-2 1-1+23 10+21 10 + 1-2 11 + 1-3 10 + 11 = "0 2 0" 4 7 2 2 4 1 The above example exhibits a curious fact concerning matrix multiplication, namely that if AB and BA are defined then these products need not be equal. Indeed, as we have just seen, AB and BA need not even be of the same size. It is also possible for AB and BA to be defined and of the same size and still be not equal:

8 Basic Linear AigeDra Example 1.5 The matrices are such that AB = 0 and BA = A. We thus observe that in general matrix multiplication is not commutative. A = ro ii [o oj B = [1 01 [o oj EXERCISES 1.12 Compute the matrix product 3 1 -2 2 -2 -1 1 1.13 Compute the matrix product "l 1 0 1 l" -1 1 1 2 "l 1 l" 0 1 1 0 0 1 1 0 0 1 1 0 1 1 1 1.14 Compute the matrix products 1 2 3 4 1.15 Given the matrices A = 12 3 4 12 3 4 " 3 0" -1 2 1 1 , B = [4 0 -11 2 , c = [l 4 2] 3 1 5 J compute the products (>4£)C and A(BC). We now consider the basic properties of matrix multiplication Theorem 1.5 Matrix multiplication is associative [in the sense that, when the products are defined, A(BC) = (AB)C). Proof For A{BC) to be defined we require the respective sizes to be m x «, n x p, p x q in which case the product A(BC) is also defined, and conversely. Computing the

The Algebra of Matrices (ij)-th element of this product, we obtain A-I *= I 1= I ii /i A=l 1=1 If we now compute the (i',y)-th clement of {AB)C, wc obtain the same. /' . « [(ab)c],j = 5>*Un = E (E «A/k /= I /= I A=l = EErtA/<v /= I A=l Consequently we see that A(BC) = (AB)C. D Because of Theorem 1.5 we shall write ABC for either A(BC) or {AB)C. Also, for every positive integer n we shall write A" for the product AA -A{n terms). EXERCISES 1.16 Compute the matrix product xy\ a h g h b f g f c X y l Hence express in matrix notation the equations (1) jc2 + 9*y + y2 + 8* + 5y + 2 = 0; 2 2 JC V (3) jcy = a2; (4) y2 = 4ax. 1.17 Compute v42 and A3 where >4 = Matrix multiplication and matrix addition are connected by the following distributive laws. Theorem 1.6 When the relevant sums and products are defined, we have A{B + C) = AB + AC, (B + C)A = BA + CA. "o 0 0 a 0 0 a2] a 0

10 Basic Linear Algebra Proof For the first equality we require A to be of size mxn and B, C to be of size n x p, in which case [A(B + Q)lj = £ aik[B + C]kj = £ aik(bkJ + ckJ) n n = £ aiAj + E fl«c*y *=i *=i = [^Ly + [*C]/y = [>4B + >4C]/; and it follows that A(B + C) = AB + AC. For the second equality, in

which we require B, C to be of size mxn and A to be of size n x p, a similar argument applies. D Matrix multiplication is also connected with multiplication by scalars. Theorem 1.7 IfAB is defined then for all scalars \ we have \(AB) = (\A)B = A(\B). Proof The (/,;)-th elements of the three mixed products are A=l *=l *=l from which the result follows. D EXERCISES 1.18 Consider the matrices o r 0 1 , B = -l -i" 0 0 Prove that (A + B)2?A2 + 2AB + B2, but that (A + B)3 = A3 + 3A2B + 3AB2 + B3. Definition A matrix is said to be square if it is of size n x n\ i.e. has the same number of rows and columns.

• The Algebra of Matrices Our next result is the multiplicative analogue of Theorem 1.2, but the reader should note that it applies only in the case of square matrices. Theorem 1.8 There is a unique n x n matrix M with the property that, for every n x n matrix A, AM = A = MA. Proof Consider the n x n matrix 100 ...0 0 1 0 ...0 M= 0 0 1 ...0 .000 ... 1 More precisely, if we define the Kronecker symbol 6^ by ifi=;; 0 otherwise, ".-{I then we have M = [6ij]nxn. If A = [fliylnxiI then [AM]tj = "£aik6kj = aijt the last equality following from the fact that every term in the summation is 0 except that in which k = ;, and this term is a^l = air We deduce, therefore, that AM - A. Similarly, we can show that MA = A. This then establishes the existence of a matrix M with the stated property. To show that such a matrix M is unique, suppose that P is also annxn matrix such that AP = A = PA for every n x n matrix A. Then in particular we have MP = M = PM. But, by the same property for Mt we have PM - P = MP. Thus we see that P = M. D Definition The unique matrix M described in Theorem 1.8 is called the n x n identity matrix and will be denoted by /,,. Note that /,, has all of its 'diagonal' entries equal to 1 and all other entries 0. This is a special case of the following important type of square matrix. Definition A square matrix D - [djj]nxn is said to be diagonal if dti = 0 whenever i fj. Less formally, D is diagonal when all the entries off the main diagonal are 0. EXERCISES 1.19 If A and B are n x n diagonal matrices prove that so also is AB.

12 Basic Linear Algebra 1.20 If A and B are n x n diagonal matrices prove that so also is MB* for all positive integers p,q. There is no multiplicative analogue of Theorem 1.3; for example, if then we have a b' c d "0 1' 0 0 = "0 a 0 c so there is no matrix A/ such that MA = /2. Note also that several of the familiar laws of high-school algebra break down in this new algebraic system. This is illustrated in particular in Example 1.5 and Exercise 1.14 above.

Definition We say that matrices A, B commute if AB = BA. Note that for A B to commute it is necessary that they be square and of the same size. EXERCISES 1.21 If A and B commute, prove that so also do -4''' and B'' for all positive integers m and /?. 1.22 Using an inductive argument, prove that if A and B commute then the usual binomial expansion is valid for $(A + B)n$. There are other curious properties of matrix multiplication. We mention in particular the following examples, which illustrate in a very simple way the fact that matrix multiplication has to be treated with some care. Example 1.6 If X is a non-zero real number then the equality 0 X' x-' 0 0 X'' X''1 0 shows that /2 has infinitely many square roots! Example 1.7 It follows by Theorem 1.4(5) that the matrix 0WXW has the property that 0/IXM>4 = °«xn = A0nxn for every n x n matrix A. In Example 1.5 we have seen that it is possible for the product of two non-zero matrices to be the zero matrix.

1. The Algebra of Matrices 13 Definition If A is an m x n matrix then by the transpose of A we mean the n x m matrix A' whose (i,;)-th element is the {jti)~\h element of A. More precisely, if A = [flf/]/IIK/l then A' = [a^m- The principal properties of transposition of matrices are summarised in the following result. Theorem 1.9 When the relevant sums and products are defined, we have (A')' = A, (A + B)' = A' + B\ (\A)' = \A\ (AB)' = B'A'. Proof The first three equalities follow immediately from the definitions. To prove that [AB]' = B'A' (note the reversal!), suppose that A = [a^)mxll and B = [b ^)^,. Then (AB)' and B'A' are each of size p x m. Since we deduce that (AB)' = B'A'. D EXERCISES 1.23 Prove that, when either expression is meaningful, [A(B + C)]' = B'A'+CA'. 1.24 Prove by induction that $(An)' = (A')n$ for every positive integer n. 1.25 If A and B commute, prove that so also do A' and B'. 1.26 LetX=[fl b c] and 0 -c b c 0 -a -b a 0 A = where $a2 + b2 + c2 = 1$. (1) Show that $y*2 = X'X-/2$. (2) Prove that A3 = -A. (3) Find A* in terms of X. Definition A square matrix A is symmetric if A - A'\ and skew-symmetric if A = -A'.

14 Basic Linear Algebra Example 1.8 For every square matrix A the matrix A + A' is symmetric, and the matrix A - A' is skew-symmetric. In fact, by Theorem 1.9, we have {A+A')' = A' + {A')' = A'+A\ (A -A')' = A'-{A')' = A' - A = -{A -A'). Theorem 1.10 Every square matrix can be expressed uniquely as the sum of a symmetric matrix and a skew-symmetric matrix. Proof The equality A=\{A + A')+\{A-A% together with Example 1.8, shows that such an

expression exists. As for uniqueness, suppose that A - B + C where B is symmetric and C is skew-symmetric. Then A' = ff + C = B -C. It follows from these equations that B = i(/\ + A') and C=\{A-A'). ☐ EXERCISES 1.27 Prove that the zero square matrices are the only matrices that are both symmetric and skew-symmetric. 1.28 Let A,B be of size n x n with A symmetric and B skew-symmetric. Determine which of the following are symmetric and which are skew- symmetric: AB + BA, AB-BA, A2, B2, AP&AP {pt q positive integers). SUPPLEMENTARY EXERCISES 1.29 Let x and y be n x 1 matrices. Show that the matrix A = xy' - yx' is of size n x n and is skew-symmetric. Show also that x'y and y'x are of size 1 x 1 and are equal. If x'x = yy = [1J and x'y = y'x = [*], prove that A3 = (it2 - 1)A. 1.30 Show that if A and B are 2 x 2 matrices then the sum of the diagonal elements of AB - BA is zero. If £ is a 2 x 2 matrix and the sum of the diagonal elements of E is zero, show that E2 = \l2 for some scalar X. Deduce from the above that if A, B, C are 2 x 2 matrices then {AB-BA)2C = C{AB-BA)2.

1. The Algebra of Matrices 15 1.31 Determine all 2 x 2 matrices X with real entries such that X2 = I2. 1.32 Show that there are infinitely many real 2x2 matrices A with A2 = -I2. 1.33 Let A be the matrix TO a a2 fl31 0 0 a a2 000 a .0 0 0 0 . Define the matrix B by B = A-]1A2 + \A3-±A4 + --. Show that this series has only finitely many terms different from zero and calculate B. Show also that the series has only finitely many non-zero terms and that its sum is A. 1.34 UA = cos i? sin t? -sin t? cost? andfl = cos <p sin <p —sin <p cos <p prove that 1.35 HA = T cos(t? + p) sin(t? + p) AB = ... v . „ . -sin(t7 + p) cos(t? + p) prove that An = cos t? sin t? -sin t? cost? 1.36 Prove that, for every positive integer/i, cos /it? sin /it? -sin/it? cosnt? a I 0" 0 a 1 0 0a n = /la a1 0 n-l 5-/1(/1 - l)a na' a n-2 1.37 If A and B are /1 x n matrices, define the Lie product [AB] = AB-BA. Establish the following identities : (1) [[AB]C]+ [[BC]A] + [[CA]B) = 0-t (2)[(A + B)C) = [AC]+[BC]- (3) [[[AB]C)D] + [[[BC]D)A] + [[[CD)A)B] + [[[DA)B)C] = 0. Show by means of an example that in general [[>4B]Cl f [A[BC\[. fliy + fl2 . b.z + b2 , c.z+c2 1.38 Given that x = — and y = r1 ^ prove that x = where a3y + a4 cx c2 .C3 C4. b3z + b4 a, fl2] r^i a3 a4\ [b3 c3z + c4

16 Basic Linear Algebra ASSIGNMENT EXERCISES {I) For each x £ IR define [cosh x sinh x " sinh .v cosh x Prove that, for all .v.y 6 IR, AXAX = Al+V. (2) For each real polynomial /?(X) = «„ + axX + • • • + rt^" denote its

derivative by D/?(X) and let = [/'(°) ° ' " [Dp(0) />(0) ' Prove that, for all polynomials p(X) and q{X), Mp + A/,, = M,,+v and A/;|A/,, = A/w. (3) A square matrix A is nilpotent if/\" = 0 for some positive integer/?; and unipotent if l,, -A is nilpotent. If N is nilpotent and U is unipotcnt deline ex p N = l,, + N + i N + • • • + £ A/* + • • • logi/ = -(/,,-¢/)-1 {/,,-1/)2 \(in-u)k- • [Here the notation reflects that for the series expansions of the functions ex and log(1 + a) from analysis, bul in this situation there is no question of 'convergence' since each expression is in fact a finite sum.] For the matrices n = verify that exp log U - U and log exp N - N. For each real number t deline U(t) = exp tM where 0 12 3 0 a b 0 0c 0 0 0 . u = 1 X V 0 1 z 0 0 1 M = 00 4 5 00 0 6 00 00 Determine 1/(/) and verify that U{s)U(t) = U{s + t) for all s, t 6 IR.

2 Some Applications of Matrices We shall now give brief descriptions of some situations to which matrix theory finds a natural application, and some problems to which the solutions are determined by the algebra that we have developed. Some of these applications will be dealt with in greater detail in later chapters. 1. Analytic geometry In analytic geometry, various transformations of the coordinate axes may be described using matrices. By way of example, suppose that in the two-dimensional cartesian plane we rotate the coordinate axes in an anti-clockwise direction through an angle t?, as illustrated in the following diagram: Let us compute the new coordinates (jc', /) of the point P whose old coordinates were (*,)')•

18 Basic Linear Algebra From the diagram we have x = rcos a and y = rsin a so x' - rcos(a -t?) = rcosacost? + rsin a sin t? = x cos t? + y sin t?; / = rsin (a -1?) = rsin a cos t? - rcos a sin t? = y cost?-* sin t?. These equations give x'ty' in terms of xty and t?. They can be expressed in the matrix form cost? sin t?"| pel -sin i? cost?J [y\' The 2x2 matrix _ T cost? sin t?1 * -sin t? cost?J is called the rotation matrix associated with t?. It has the following property: RjRj = /2 — /?fl/?fl. In fact, we have _ I" cost? sin t?l Tcost? —sin 1?" * *~ [-sin 1? cost?J [sini? cost? _ ' cos2t? + sin2t? -cost?sin t? + sin t?cost?" -sin t?cost? + cost?sin t? sin2t? + cos2t? "0 lj' and similarly, as the reader can verify, /?'tf/?tf = /2. This leads us more generally to the following notion. Definition Annxn matrix A is said to be orthogonal if AA' = IR = A'A. It follows from the above that to every rotation of axes in two dimensions we can associate a real orthogonal matrix ('real' in the sense that its elements are real numbers). EXERCISES 2.1 If A is an orthogonal n x n matrix prove that A' is also

orthogonal. 2.2 If A and B are orthogonal nxn matrices prove that AB is also orthogonal. x' y'

2. Some Applications of Matrices 19 2.3 Prove that a real 2x2 matrix is orthogonal if and only if it is of one of the forms a b -b a b -a where a2 + b2 - 1. Consider now the effect of one rotation followed by another. Suppose that we transform (xt y) into (x'ty') by a rotation through t?, then (*',/) into (*",/') by a rotation through (p. Then we have x" /'. cosip -sin (p cosy? -sin <p sin cos sin cos cost? sin t? | * shit? cost?J [y This suggests that the effect of one rotation followed by another can be described by the product of the corresponding rotation matrices. Now it is intuitively clear that the order in which we perform the rotations does not matter, the final frame of reference being the same whether we first rotate through t? and then through <p or whether we rotate first through <p and then through t?. Intuitively, therefore, we can assert that rotation matrices commute. That this is indeed the case follows from the identities which the reader can easily verify as an exercise using the standard identities for cos(t? + ip) and sin (t? + <p). Example 2.1 Consider the hyperbola whose equation is x2 - y2 = 1. If this is rotated through 45° anti-clockwise about the origin, what does its new equation become? To answer this, observe first that rotating the hyperbola anti-clockwise through 45° is equivalent to rotating the axes clockwise through 45°. Thus we have = /?_ n/4 Now since ^ir/4^-ir/4 ~ ^*/4-n/4 = fy) - ' we can multiply the above equation by /?w/4 to obtain X -^»/4 = i i I 1 .~vS v3. V" y. so that r — 1 *•' J. 1 v' & - I r' I ,./ y=-vx + vy

20 Basic Linear Algebra Thus the equation x2 - y2 = 1 transforms to i.e. to2x'y= 1. EXERCISES 2.4 Two similar sheets of graph paper are pinned together at the origin and the sheets are rotated. If the point (1,0) of the top sheet is directly above the point (^, ‖) of the bottom sheet, above what point of the bottom sheet does the point (2,3) of the top sheet lie? 2.5 For every point (xty) of the cartesian plane let (*',/) be its reflection in the x-axis. Prove that 10" H- 0 -1 2.6 In the cartesian plane let L be a line passing through the origin and making an angle t? with the x-axis. For every point (x,y) of the plane let (xLtyL) be its reflection in the line L. Prove that "cos2t? sin 2t?l I sin2t? - cos2t? \\y *L 2.7 In the cartesian plane let L be a line passing through the origin and making an angle t? with the jc-axis. For every point {x,y) of the plane let (x*, y*) be the projection of (x, y) onto L (i.e. the foot of the perpendicular

from (x, y) to L). Prove that cos2t? sin t?cost? sin t?cost? sin2t? 2. Systems of linear equations As we have seen above, a pair of equations of the form 011*1+ 012*2 = ^1. 021*1 + 022*2 = ^2 can be expressed in the matrix form Ax = b where A = 0ii 012 .021 022. Let us now consider the general case. x = *2 b =

2. Some Applications of Matrices 21 Definition By a system of m linear equations in the n unknowns xx,..., xn we shall mean a list of equations of the form 0ii*i + <»i2*2 + 013*3 + •■■ + alnxn = bx «21*1 + fl22*2 + ^23*3 + ' " + fl2«*« = *>2 alxxx + fl32x2 + fl33x3 + • ■ • + a3nxR = b3 flmi*i + am2x2 + flm3x3 + ■•• + fl/nMxM = bm where the a/y- and the b, are numbers. Since clearly "a,, a,2 fli3 ... aXnl pr,*| f" a,,*, + anx2 + a,3x3 + ■ ■ ■ + alnxn ' Q2x Q22 fl23 ■ • • a2n x2 fl2I*l + fl22*2 + fl23*3 + " ■ ■ + <*2nXn fl3l fl32 «33 • • • aln *3 = fl3l*l + fl32*2 + «33*3 + ■ ■ ■ + OytXn Pml aml <*m3 • • • flmJ L* J L0,nI*I + ^2*2 + flm3*3 + ' ' ' + "»7,«*«. we see that this system can be expressed succinctly in the matrix form Ax = b where A = [a;j]mxn and x, b are the column matrices r*ii *2 -*«- , b = rM h X. The m x n matrix A is called the coefficient matrix of the system. Note that it transforms a column matrix of length n into a column matrix of length m. In the case where b = 0 (i.e. where every bt = 0) we say that the system is homogeneous. If we adjoin to A the column b then we obtain an m x (n + 1) matrix which we write as A\b and call the augmented matrix of the system. Whether or not a given system of linear equations has a solution depends heavily on the augmented matrix of the system. How to determine all the solutions (when they exist) will be the objective in Chapter 3. EXERCISES 2.8 If, for given A and b, the matrix equation Ax = b has more than one solution, prove that it has infinitely many solutions. [Hint. If x, and x2 are solutions, show that so also is pxx + ^x2 where p + q=l.)

22 OCtoK, LlllCdl /"Ujjooiu 3. Equilibrium-seeking systems Consider the following situation. In normal population movement, a certain proportion of city dwellers move into the country every year and a certain proportion of country dwellers decide to become city dwellers. A similar situation occurs in national employment where a certain percentage of unemployed people find jobs and a certain percentage of employed people become unemployed. Mathematically, these situations are essentially the same. The problem of how to describe them in a concrete mathematical way, and to answer the obvious

question of whether or not such changes can ever reach a 'steady state* is one that we shall now illustrate. To be more specific, let us suppose that 75% of the unemployed at the beginning of a year find jobs during that year, and that 5% of people with jobs become unemployed during the year. These proportions are of course somewhat optimistic, and might lead one to conjecture that 'sooner or later* everyone will have a job. But these figures are chosen to illustrate the point that we want to make, namely that the system 'settles down* to fixed proportions. The system can be described compactly by the following matrix and its obvious interpretation: unemployed employed into unemployment J ^ into employment J g Suppose now that the fraction of the population that is originally unemployed is L0 and that the fraction that is originally employed is A/0 = 1 - L0. We represent this state of affairs by the matrix V More generally, we let the matrix V signify the proportions of the unemployed/employed population at the end of the /-th year. At the end of the first year we therefore have M.-jLo + gAfo and we can express these equations in the matrix form V which involves the 2 x 2 matrix introduced above. = •I I • 4 20 3 19 .4 70. MQ

2. Some Applications of Matrices 23 Similarly, at the end of the second year we have L2 = \LX + ±M, M2 = \Lx + gM, and consequently ^2 M2 -Bflfe]- J."-2 20 19 20 J M« Using induction, we can say that at the end of the it-th year the relationship between LktMk and L0,M0 is given by Mt ri 4 3 U 20 12 20 J MQ Now it can be shown (and we shall be able to do so much later) that, for all positive integers it, we have rl 4 3 L4 I I 20 19 20 J k _ I ~ 16 1 + ^ 1 + 5* 1- 15(1 -f) 15 + £ This is rather like pulling a rabbit out of a hat, for we are far from having the machinery at our disposal to obtain this result; but the reader will at least be able to verify it by induction. From this formula we can see that, the larger k becomes, the closer is the approximation ri 4 3 U i 1 20 19 20 J * rv r i 16 IS L 16 I 1 16 IS 16 J Since LQ + M0 = 1, we therefore have Lk Mk 16 16 II ^0 _ Put another way, irrespective of the initial values of L0 and M0, we see that the system is 'equilibrium-seeking* in the sense that 'eventually* one sixteenth of the population remains unemployed. Of course, the lack of any notion of a limit for a sequence of matrices precludes any rigorous description of what is meant by an 'equilibrium-seeking* system. However, only the reader's intuition is called on to appreciate this particular application. 4. Difference equations The system of pairs of equations *n+i = axn + byn yn+i = cxn+dyn

24 Basic Linear Algebra is called a system of linear difference equations. Associated with such a system are the sequences (jt,,)w>, and (>w)w^i, and the problem is to determine the general values of xR,yR given the initial values of xllyl. The above system can be written in the matrix form zw+, = Azn where r *» w , A = a c b] d\ Clearly, we have z^ = Azv z3 = Ai^ = A2zv and inductively we see that zn+| = A Z|. Thus a solution can be found if an expression for A" is known. The problem of determining the high powers of a matrix is one that will also be dealt with later. 5. A definition of complex numbers Complex numbers are usually introduced at an elementary level by saying that a complex number is 'a number of the form $x + iy$ where x, y are real numbers and $i2 = -1$ \ Complex numbers add and multiply as follows: $(x + iy) + (x' + iy') = (x + x') + i(y + /)$; $(x + iy)(x' + iy') = (xx1 -yy') + i(xy' + x'y)$. Also, for every real number X we have \(x + iy)= \x + i\y. This will be familiar to the reader, even though (s)he may have little idea as to what this number system is! For example, i - \fA is not a real number, so what does the product iy mean? Is /0 = 0? If so then every real number x can be written as $x = x + i0$, which is familiar. This heuristic approach to complex numbers can be confusing. However, there is a simple approach that uses 2x2 matrices which is more illuminating and which we shall now describe. Of course, at this level we have to contend with the fact that the reader will be equally unsure about what a real number is, but let us proceed on the understanding that the real number system is that to which the reader has been accustomed throughout her/his schooldays. The essential idea behind complex numbers is to develop an algebraic system of objects (called complex numbers) that is 'larger' than the real number system, in the sense that it contains a replica of this system, and in which the equation $x2 + I = 0$ has a solution. This equation is, of course, insoluble in the real number system. There are several ways of 'extending* the real number system in this way, and the one we shall describe uses 2x2 matrices. For this purpose, consider the collection C2 of all 2 x 2 matrices of the form a b' M(a,b) = -b

2. Some Applications of Matrices 25 where a and b are real numbers. Invoking Theorem 1.10, we can write M{a,b) uniquely as the sum of a symmetric matrix and a skew-symmetric matrix: a b^ -b a a 0 0 a 0 b -b 0 Thus, if we define A = 0 1 -1 0 we see that every such matrix M(a, b) can be written uniquely in the form M(a,b) = al2 + bJ2. Observe now that the collection R2 of all 2x2 matrices in C2 that are of the form M(a, 0) = al2 is a replica of the real number

system; for the matrices of this type add and multiply as follows: xl2 + yl2 = x + y 0 0 x + y = (x + yV2'> xl2yl2 = xy 0 0 xy = to)/2, and the replication is given by associating with every real number x the matrix xl2 = Af(x,0). Moreover, the identity matrix /2 = 1 • /2 belongs to /?2, and we have -1 0 A = o nro r -i oJ[—1 o o -1 = -i 2. so that J2 + /2 = 0. In other words, in the system C2 the equation x2 + 1 = 0 has a solution (namely J2). The usual notation for complex numbers can be derived from C2 by writing each al2 as simply a, writing J2 as i, and then writing aI2+bJ2 asa+bi. Since clearly, for each scalar b, we can define J2b to be the same as bJ2 we have that a + bi = a + ib. Observe that in the system C2 we have a bVia' br -b a\[-b' a' = M{a + a\b + b')\ Under the association = M(aa'-bb',ab' + ba'). M(a,b) *-> a + ib, the above equalities reflect the usual definitions of addition and multiplication in the system C of complex numbers. This is far from being the entire story about <C, the most remarkable feature of which is that every equation of the form a,,Xn + an_xXn~* + • • • + atX + a0 = 0, where each a, G <C, has a solution.

26 Basic Linear Algebra EXERCISES 2.9 Let A be a complex 2x2 matrix, i.e. the entries of A are complex numbers. Prove that if A2 = 0 then A is necessarily of the form zw z2 1 —w2 -zw for some z, w G C. By considering the matrix 0 0* 1 0 show that the same is not true for real 2x2 matrices. 2.10 The conjugate of a complex number z = x + iy is the complex number z = x-iy. The conjugate of a complex matrix A = [z/yLx,, is the matrix A = [Iij]mxn. Prove that A = A' and that, when the sums and products are defined, A + B = A+Band7iB = AB. 2 11 A square complex matrix A is hermitian if A' - A, and skew-hermitian if A' = -A. Prove that A + A' is hermitian and that A - A' is skew- hermitian. Prove also that every square complex matrix can be written as the sum of a hermitian matrix and a skew-hermitian matrix. ASSIGNMENT EXERCISES (1) A quaternion is a matrix of complex numbers of the form H = a + ib c + id -c + id a-ib 1 0" [o ij , i = "i 0" L° _/'J , j = ' 0 I" [-1 oj , k = o r [i oj Prove that the set of quaternions is closed under the operations of addition and multiplication. Writing 1 = show that H = al + bi + cj + dk and that i2=j2 = k2 = -l, y=-ji = k, jk = —kj = i, ki=-ik = j. (2) If H is a quaternion, prove that HJT = WH= {a2 +b2 + c2 + d2)l. Un{H) = HIT prove that n{HK) = n{H)n{K). (3) Show that the system of quaternions can be represented as a system of real matrices of size 4x4.

3 Systems of Linear Equations We shall now consider in some detail a

systematic method of solving systems of linear equations. In working with such systems, there are three basic operations involved: (1) interchanging two equations (usually for convenience); (2) multiplying an equation by a non-zero scalar; (3) forming a new equation by adding one equation to another. The operation of adding a multiple of one equation to another can be achieved by a combination of (2) and (3). We begin by considering the following three examples. Example 3.1 To solve the system $y + 2z = 1$ (1) $x - 2y + z = 0$ (2) $3y - 4z = 23$ (3) we multiply equation (1) by 3 and subtract the new equation from equation (3) to obtain $-\backslash 0z = 20$, whence we see that $z = -2$. It then follows from equation (1) that y - 5, and then by equation (2) that $x = 2y - z = 12$. Example 3.2 Consider the system $x - 2y - 4z = 0$ (1) $-2x + Ay + 3z = 1$ (2) $-x + 2y - z = 1$ (3) If we add together equations (1) and (2), we obtain equation (3), which is therefore superfluous. Thus we have only two equations in three unknowns. What do we mean by a solution in this case?

28 DdSIl Ull ICOI rv\jm» " Example 3.3 Consider the system $x + y + z + t = 1$ (1) $x - y - z + t = 3$ (2) $-x - y + z - t = I$ (3) $-3x + y - 3z - 3/ = 4$ (4) Adding equations (1) and (2), we obtain $x + 1 = 2$, whence it follows that $>> + z = -1$. Adding equations (1) and (3), we obtain $z = 1$ and consequently $y = -2$. Substituting in equation (4), we obtain $-3x - 3/ = 9$ so that $x + / = -3$, which is not consistent with $x + / = 2$. This system therefore does not have a solution. Expressed in another way, given the 4 x 4 matrix A = . I 1 I 1 1 -1 -I 1 -1 -1 1 -1 -3 1 -3 -3 > there are no numbers xtytz,t that satisfy the matrix equation A x~ y z t = "I" 3 1 4 • The above three examples were chosen to provoke the question: is there a systematic method of tackling systems of linear equations that (a) avoids a haphazard manipulation of the equations; (b) yields all the solutions when they exist; (c) makes it clear when no solution exists? In what follows our objective will be to obtain a complete answer to this question. We note first that in dealing with systems of linear equations the 'unknowns1 play a secondary role. It is in fact the coefficients (which are usually integers) that are important. Indeed, each such system is completely determined by its augmented matrix. In order to work solely with this, we consider the following elementary row operations on this matrix: (1) interchange two rows; (2) multiply a row by a non-zero scalar; (3) add one row to another. These elementary row operations clearly correspond to the basic operations on equations listed above.

3. Systems of Linear Equations 29 It is important to observe that these elementary row operations do not affect the solutions (if any) of the system. In fact, if the original system of equations has a solution then this solution is also a solution of the system obtained by applying any of the operations (1), (2), (3); and since we can in each case perform the 'inverse* operation and thereby obtain the original system, the converse is also true. We begin by showing that the above elementary row operations have a fundamental interpretation in terms of matrix products. Theorem 3.1 Let P be the mx m matrix that is obtained from the identity matrix Im by permuting its rows in some way. Then for any m x n matrix A the matrix PA is the matrix obtained from A by permuting its rows in precisely the same way. Proof Suppose that the i-th row of P is thej'-th row of Im. Then we have (*= 1 m) Pik = &ik- Consequently, for every value of it, m m [PAlk = Z>,7fl* = £ Sjiatk = fly*. whence we see that the i-th row of PA is they-th row of A. O P = Example 3.4 Consider the matrix '10 0 0" 0 0 10 0 10 0 0 0 0 1 obtained from /4 by permuting the second and third rows. If we consider any 4x2 matrix A = "fll fl2 <*1 fl4 *'l h h *J and we compute the product PA = 10 0 0" 0 0 10 0 10 0 0 0 0 1 "a, bx~ a2 bi a3 b3 a4 bA "a, bx~ a, b3 a2 b2 a4 /?4 we see that the effect of multiplying A on the left by P is to permute the second and third rows of A.

30 Basic Linear Algebra EXERCISES 3.1 Explain the effect of left multiplication by the matrix "0 0 10" P = 0 0 0 1 10 0 0 0 10 0 3.2 Explain the effect of left multiplication by the matrix 0 0 10' P = 10 0 0 0 10 0 0 0 0 1 Theorem 3.2 Let Abeanmx m matrix and let Dbethemxm diagonal matrix X D = m. Then DA is obtained from A by multiplying the i-th row of A by \,for i= 1,..., m. Proof Clearly, we have dti = yfiy. Consequently, m m [DAlj = £dikakJ = £\i6ikakj = X,fl/;, and so the i-th row of DA is simply \ times the i-th row of A. D Example 3.5 Consider the matrix D = 10 0 0 0 a 0 0 0 0 p 0 0 0 0 1 obtained from /4 by multiplying the second row by ct and the third row by /?. If fl2 bj A = o3 b3 _a4 fc4_

3. Systems of Linear Equations 31 and we compute the product DA = 10 0 0" 0 a 0 0 0 o p 0 0 0 0 1 "fl, b{ a2 b2 a3 b3 a4 b4 ' a\ *i~ aa2 otbi 0a3 /3b3 . °4 ^4. we see that the effect of multiplying A on the left by D is to multiply the second row of A by a and the third row by /?. EXERCISES 3.3 Explain the effect of left multiplication by the matrix P = 0 0 7 0 0 0 0 6 a 0 0 0 0 0 0 0 Theorem 3.3 Let Pbethemxm matrix that is obtained from Im by adding \ times

the s-th row to the r-th row (where r, s are fixed with r ^ s). Then for any mxn matrix A the matrix PA is the matrix that is obtained from A by adding \ times the s-th row of A to the r-th row of A. Proof Let E)s denote the m x m matrix that has A in the (r, j)-th position and 0 elsewhere. Then we have f A if i = rj = s; 0 otherwise. 1¾ = ■ Since, by definition, P= Im + /¾. it follows that [PA)ij = [A + E"rsA)iJ [A];; ifl^r. [A)rj + \[A)sj if/ = r. Thus we see that PA is obtained from A by adding A times the $-th row to the r-th row. D

32 Basic Linear Algebra Example 3.6 Consider the matrix P = "1 0 0 \ 1 0 0" 0 1 which is obtained from /3 by adding \ times the second row of /3 to the first row. If A = "fll fl2 .fl3 *l h hi and we compute the product PA = "1 \ 0" 0 1 0 0 0 1 'a, b{ a-i bj _fl3 h. = fl| + Afl2 ^1 "*" ^^2 fl2 /?2 fl3 ^3 . we see that the effect of multiplying A on the left by P is to add \ times the second row of A to the first row. EXERCISES 3.4 Explain the effect of left multiplication by the matrix "0 0 1 Of 13 1 1" 0 0 Definition By an elementary matrix of size nxnwe shall mean a matrix that is obtained from the identity matrix /,, by applying to it a single elementary row operation. In what follows we use the 'punning notation' p, to mean 'row i\ Example 3.7 The following are examples of 3 x 3 elementary matrices: "1 0 0 0 0 1 0" 1 0 (P2 <"> P3^ "1 0 0 0 0" 2 0 0 1 (2p2); "1 0 0 0 1 0 f 0 1 (Pl+P3)

3. Systems of Linear Equations 33 Definition In a product AB we shall say that B is pre-multiplied by A or, equivalently, that A is post-multiplied by B. The following result is now an immediate consequence of Theorems 3.1,3.2 and 3.3: Theorem 3.4 An elementary row operation onanmxn matrix A is achieved on pre-multiplying A by a suitable elementary matrix. The elementary matrix in question is precisely that obtained by applying the same elementary row operation to Im. D Having observed this important point, let us return to the system of equations described in matrix form by Ax = b. It is clear that when we perform a basic operation on the equations all we do is to perform an elementary row operation on the augmented matrix A\b. It therefore follows from Theorem 3.4 that performing a basic operation on the equations is the same as changing the system Ax = b to the system EAx = Eb where E is some elementary matrix. Moreover, the system of equations that corresponds to the matrix equation EAx = Eb is equivalent to the original system in the sense that it has the same set of solutions (if any). Proceeding in this way, we see that to every string of k basic operations there corresponds a string of elementary

matrices Elt...tEk such that the the resulting system is represented by Ek - • • E2ExAx - Ek - - • E2Elb, which is of the form Bx = c and is equivalent to the original system. Now the whole idea of applying matrices to solve systems of linear equations is to obtain a simple systematic method of determining a convenient final matrix B so that the solutions (if any) of the system Bx = c can be found easily, such solutions being precisely the solutions of the original system Ax = b. Our objective is to develop such a method. We shall insist that the method (1) will avoid having to write down explicitly the elementary matrices involved at each stage; (2) will determine automatically whether or not a solution exists; (3) will provide all the solutions. In this connection, there are two main problems that we have to deal with, namely (a) what form should the matrix B have?; (b) can our method be designed to remove all equations that may be superfluous? These requirements add up to a tall order perhaps, but we shall see in due course that the method we shall describe meets all of them. We begin by considering the following type of matrix.

34 Basic Linear Algebra Definition By a row-echelon (or stairstep) matrix we shall mean a matrix of the general form "0 ... 0[* in which all the entries 'under the stairstep* are zero, all the 'corner entries' (those marked *) are non-zero, and all other entries are arbitrary. • Note that the 'stairstep* descends one row at a time and that a 'step* may traverse several columns. Example 3.8 The 5x8 matrix is a row-echelon matrix. Example 3.9 The 3x3 matrix is a row-echelon matrix. Example 3.10 Every diagonal matrix in which the diagonal entries are non-zero is a row-echelon matrix. Theorem 3.5 By means of elementary row operations, a non-zero matrix can be transformed to a row-echelon matrix. Proof Suppose that A = [aij]mxn is a given non-zero matrix. Reading from the left, the first non-zero column of A contains at least one non-zero element. By a suitable change '0 0 0 0 0 1 0 0 0 0 3 0 0 0 0 0 4 0 0 0 2 9 0 1" 0 3 5 1 0"1 1 1 1 0 0 0 0 0 0 0 0

3. Systems of Linear Equations 35 of rows if necessary, we can move the row containing this non-zero entry so that it becomes the first row, thereby obtaining a matrix of the form B = b2X bn b22 b2k 0 b. 'm\ uml — umk. in which bu ^0. Now for i = 2,3,... ,#i subtract from the i-th row biXlbu times the first row. This is a combination of elementary row operations and transforms B to the matrix 0 ... 0^^ ... blk 0 ... 0 0 c22 ... c2k c= 0 0 0 cml ••■ Cmk_ in which we see the beginning of the stairstep. Now leave the first row alone and

concentrate on the (m - 1) x (* - 1) matrix [Cy]. Applying the above argument to this submatrix, we can extend the stairstep by another row. Clearly, after at most m applications of this process we arrive at a row-echelon matrix. Q The above proof yields a practical method of reducing a given matrix to a row echelon matrix. This process is often known as Gaussian elimination. Example 3.11 1 1 3 0 0 1 1 1 1 0 1 2 0 0 1 1 f 2 1 2 *\A '1 0 ■oil 0 1 0 1 "1 0 T\\ 0 0^ 0 0 1 -1 -2 2 1 -1 -1 3 0 0 1 1 0 0 1 1 1 1 -2 2 r i -3 1 P2~P\ P3"3pi Pi~ Pi P*-Pi 10 10 1 0 1-10 1 0 0 \-\ 1 -3 0 0 0|4 -8J p4 + 3p3 It should be noted carefully that the stairstep need not in general reach the bottom row, as the following examples show.

36 Basic Linear Algebra Example 3.12 0 1 -1 0 -1 1 2" 1 -4 ■** 1-1 1 p, +- +p2 0 0 2 -1 1 -4 1 -1 1 0 0] 2 0 0 -3J p3+pi . P3 + fP2 "1 0 0 -1 f 0 2 0 0 Example 3.13 1 -2 1 2-4 2 -1 2 -1 '1 0 0 -2 1" 0 0 0 0 P2"2Pl P3+P1 EXERCISES 3.5 Reduce to row-echelon form the following matrices: "1 2 3" 3 1 2 5 5 8 * "1 2 3" 3 1 2 2 3 1 3.6 Reduce to row-echelon form the following matrix: 1 1 1 1 2 2 0 1 0 3 1 1 3 3 1 2 1 3 3 1 3.7 Given the matrix A = 110 14 2 0 0 4 7 1110 5 I -3 -1 -10 a determine the value of a such that, in a row-echelon form of A, the stairstep reaches the bottom.

3. Systems of Linear Equations 37 Definition By an Hermite (or reduced row-echelon) matrix we mean a row-echelon matrix in which every comer entry is 1 and every entry lying above a comer entry is 0. An Hermite matrix therefore has the general form 0 .. 0 .. 0 .. 0 .. 0 .. 0 .. 0|1 0 0 . 0 0.. 0 0.. 0 0.. 0 0 . 0 . 0 1 .00.. .00.. .00.. .00.. . 0 . 0 . 0 . 0 0 0 1 0 .. 0 .. 0 .. 0 0 0 . 0 1 .00.. .00.. 0 0 0 0 . 0[1_ . 0 0 in which the unmarked entries lying above the stairstep are arbitrary Example 3.14 The 4x9 matrix "1 0 0 0 0 1 0 0 10 2 2 0 0 0 0 2 0 0 0 0 0 1 0 0 0 0 0 0 1 2 0 f 0 1 0 is an Hermite matrix. Example 3.15 The identity matrix /,, is an Hermite matrix. Theorem 3.6 Every non-zero matrix A can be transformed to an Hermite matrix by means of elementary row operations. Proof Let Z be a row-echelon matrix obtained from A by the process described in Theorem 3.5. Divide each non-zero row of Z by the (non-zero) comer entry in that row. This makes each of the comer entries 1. Now subtract suitable multiples of every such non-zero row from every row above it to obtain an Hermite matrix. □ The systematic procedure that is described in the proof of Theorem 3.6 is best illustrated by an example.

Examp Ie3.16 "1 2 1 2 4 4 3 6 5 2 8 7 1" 4 7 '1 0 0 "1 0 0 1 0 0 ~1 0 0 1 0 0 2 12 1" 0 12 4 2 0 2 14 2 12 1" T|2 4 2 0 0~|-3 2 2 12 1* ~5"|l 2 1 0 oji -I 2 1 0 J" Th o i o o]i 4 2 0 0 0" Th o i 0 o]i -|_ p2-2p, P3"3pi row-echelon P3"P2 Pi"2p3 P2 _ 2P3 Pl-P2 Hermite EXERCISES 3.8 Reduce to Hermite form the matrix 1 -1 0 -1 -5 -1 2 1-1-4 1-1 111-4-63 14 2-8-58 3.9 Reduce to Hermite form the matrices '0 0 0 1" 0 0 10 0 10 0 10 0 0 * "1110 110 1 10 11 0 111 .10 Reduce to Hermite form the matrices "10 0 1" 0 110 0 110 10 0 1 • '0110 10 0 1 10 0 1 0 110

3. Systems of Linear Equations 39 The astute reader will have observed that we have refrained from talking about the row-echelon form of a matrix. In fact, there is no unique row-echelon form. To see this, it suffices to observe that we can begin the process of reduction to row- echelon form by moving any non-zero row to become the first row. However, we can talk of the Hermite form since, as we shall see, this is unique. In fact, it is precisely because of this that such matrices are the focus of our attention. As far as the problem in hand is concerned, namely the solution of Ax - b, we can reveal that the Hermite form of A is precisely the matrix that will satisfy the requirements we have listed above. In order to establish these facts, however, we must develop some new ideas. For this purpose, we introduce the following notation. Given anmxn matrix A = [a/y] we shall use the notation and quite often we shall not distinguish this from the i-th row of A. Similarly, the i-th column of A will often be confused with the column matrix fli, a.= Definition By a linear combination of the rows (columns) of A we shall mean an expression of the form \lxl + \2x2 + • • • + \pxp where each xk is a row (column) of A. Example 3.17 The row matrix [2 - 3 1] can be written in the form 2[1 0 0]-3[0 1 0]+1[0 0 1] and so is a linear combination of the rows of /3. Example 3.18 The column matrix can be written 0" 4 -2 *0" 1 0 -2 "0" 0 1 and so is a linear combination of the columns of /3.

40 Basic Linear Algebra Example 3.19 The column matrix 'I" 0 0 is not a linear combination of the columns of the matrix "0 1 0 0 0 1 0' 1 0 Definition If jc, ,..., xp are rows (columns) of A then we shall say that x,,..., xp are linearly independent if the only scalars \,,..., \p which are such that \, jc, + \2X2 + • • • + ^pXp = 0 are\, = \2 = '"- \» = 0- Expressed in an equivalent way, the rows (columns) x,,..., xp are linearly independent if the only way that the zero row (column) 0 can be expressed as a linear combination of x,,..., xp is the trivial

way, namely $0 = 0jc, + 0*2 + \cdots + 0xp$. If x,,..., xp are not linearly independent then they are linearly dependent. Example 3.20 The columns of the identity matrix /,, are linearly independent. In fact, we have *i 1 0 0 0 + \2 0 1 0 0 +\3 0 0 1 0 + \. roi 0 0 n = M >2 >3 .K. and this is the zero column if and only if X, = X2 = $\cdots$ = Xn = 0. Similarly, the rows of/,, are linearly independent. Example 3.21 In the matrix A = 12 0 2 0 111 10 10

3. Systems of Linear Equations 41 the columns are not linearly independent. Indeed, the second and fourth columns are the same, so that a2 = a4 which we can write in the form la2 - la4 = 0. However, the first three columns of A are linearly independent. To see this, suppose that we have X,a, + \2a2 + \3a3 = 0. Then 1 0 1 + \2 2 1 0 + X, 0 1 1 = 0 0 0 and consequently from which it is easily seen that \, = \2 = \3 = 0. EXERCISES *I *l + 2\2 >2 + + >3 >3 = = = 0 0 0 3.11 Prove that in the matrix "1 2 0 1 1 0 0] 1 1 the rows are linearly independent. 3.12 For the matrix "1 2 1 2 1 0 1 1 0 3 1 1 3 3 1 2 determine the maximum number of linearly independent rows and the maximum number of linearly independent columns. Theorem 3.7 If the rows/columns xu...txpare linearly independent then none can be zero. Proof If we had xt = 0 then we could write Ox, + ■ • ■ + Ox,., + \xj + 0xi+l + $\cdots$ + 0xp = 0, so that x,,..., xp would not be independent. □ Theorem 3.8 The following statements are equivalent: (1) xu...txp$(p^2)$are linearly dependent;

42 Basic Linear Algebra (2) one of the xt can be expressed as a linear combination of the others. Proof (1) => (2): Suppose that xXt..., xp are dependent, where $p \geq 2$. Then there exist \,,..., \p such that \xxx + --- + \pxp=0 with at least one of the \t not zero. Suppose that \k f 0. Then the above equation can be rewritten in the form JC = -^-x. ^x ■** a* i A* P' i.e. x* is a linear combination of xx ,..., xk_x, xk+x,..., xp. (2) =}▶ (1): Conversely, suppose that *k = /*i*i + "' +/U-i**-i + /i*+i**+i + $\cdots$ + /iPV Then this can be written in the form /!,*, + ••• + /i*-i**-l + (-1)^ + /**+, **+, + $\cdots$ + npxp = 0 where the left-hand side is a non-trivial linear combination of xx,...txp. Thus xx,..., xp are linearly dependent. D Corollary The rows of a matrix are linearly dependent if and only if one can be obtained from the others by means of elementary row operations. Proof It suffices to observe that every linear combination of the rows is, by its very definition, obtained by a sequence of elementary row operations. D Example 3.22 The rows of the matrix A = 12 0 0 2 1-11 5 4-22 are linearly dependent. This follows from the fact that A3 = A, +

2A2 and from the Corollary to Theorem 3.8. We now consider the following important notion. Definition By the row rank of a matrix we mean the maximum number of linearly independent rows in the matrix.

3. Systems of Linear Equations 43 Example 3.23 The matrix A of the previous example is of row rank 2. To see this, recall that the three rows A,, A2, A3 are dependent. But the rows A, and A2 are independent since A | A j + A 2 «2 — IA | t £^2 ZA j ■ A2 — A2 "21 and this is zero if and only if X, = \2 - 0- Hence the maximum number of independent rows is 2. Example 3.24 The identity matrix /,, has row rank n. Example 3.25 By Theorem 3.7, a zero matrix has row rank 0. EXERCISES 3.13 Determine the row rank of the matrix "1 2 5" 2 1 4 0-1-2 * 0 1 2 It turns out that the row rank of the augmented matrix v4|b of the system Ax = b determines precisely how many of the equations in the system are not superfluous, so it is important to have a simple method of determining the row rank of a matrix. The next result provides the key to obtaining such a method. Theorem 3.9 Elementary row operations do not affect row rank. Proof It is clear that the interchange of two rows has no effect on the maximum number of independent rows, i.e. the row rank. If now the Jt-th row A* is a linear combination of p other rows, which by the above we may assume to be the rows A,,..., Ap, then clearly so is A Ak for every non-zero scalar A. It follows by Theorem 3.8 that multiplying a row by a non-zero scalar has no effect on the row rank. Finally, suppose that we add the i-th row to they-th row to obtain a newy-th row, say AJ = A,- + Ay. Since then X, A, + • • • + A.-A,- + • • • + AyAJ + • • • + ApAp = A,A, + • • • + (X,- + Ay)Ay + • • • + AyAy + • • • + ApAp,

44 Basic Linear Algebra it is clear that if A,,..., A,,..., Ay Ap are linearly independent then so also are A,,..., \h..., Ay,..., Ap. Thus the addition of one row to another has no effect on row rank. □ Definition A matrix B is said to be row-equivalent to a matrix A if B can be obtained from A by means of a finite sequence of elementary row operations. By Theorem 3.4, we can equally assert that B is row-equivalent to A if there is a product F of elementary matrices such that B = FA. Since row operations are reversible, we have that if an m x n matrix B is row- equivalent to the m x n matrix A then A is row-equivalent to B. The relation of being row-equivalent is therefore a symmetric relation on the set ofmxn matrices. It is trivially reflexive; and it is transitive since if F and G are each products of elementary matrices then clearly so is

FG. Thus the relation of being row equivalent is an equivalence relation on the set of mxn matrices. The following result is an immediate consequence of Theorem 3.9. Theorem 3.10 Row-equivalent matrices have the same row rank. D The above concepts allow us now to establish: Theorem 3.11 Every non-zero matrix can be reduced by means of elementary row operations to a unique Hermite matrix. Proof By Theorem 3.6, every non-zero matrix M can be transformed by row operations to an Hermite matrix. Any two Hermite matrices obtained from M in this way are clearly row-equivalent. It suffices, therefore, to prove that if A and B are each Hermite matrices and if A and B are row-equivalent then $A = B$. This we do by induction on the number of columns. We begin by observing that the only m x 1 Hermite matrix is the column matrix 1 0 » 0 so the result is trivial in this case. Suppose, by way of induction, that all row- equivalent Hermite matrices of size m x [n - 1) are identical and let A and B be

3. Systems of Linear Equations £5 row-equivalent Hermite matrices of size m x n. Then by Theorem 3.4 there is an m x n matrix F, which is a product of elementary matrices, such that $B = FA$. Let A and B be the m x (n -1) matrices that consist of the first /i - 1 columns of A, B respectively. Then we have B- FA and so A and B are row-equivalent. By the induction hypothesis, therefore, we have A - B. The result will now follow if we can show that the n-th columns of A and B are the same. For this purpose, we observe that in an Hermite matrix every non-zero row contains a corner entry 1, and these corner entries are the only non-zero entries in their respective columns. The non-zero rows of such a matrix are therefore linearly independent, and the number of such rows (equally, the number of corner entries) is therefore the row rank of the matrix. Now since the Hermite matrices A and B are row-equivalent, they have the same row rank and therefore the same number of corner entries. If this is r then the row rank of $A = B$ must be either r or r -1. In the latter case, the n-th columns of A and B consist of a corner entry 1 in the r-th row and 0 elsewhere, so these columns arc equal and hence in this case $A = B$. In the former case, we deduce from $B = FA$ that, for $1 < i < r$, r (1) [bn ... />,J = £>*K ... akn]. *=i In particular, for the matrix $A = B$ we have r [*il "• fl«>-l]=[*«l ••• *Vll =$>*[*« ••• ak,n-\Y *=1 But since the first r rows of A are independent we deduce from this that X, = 1 and \k = 0 for k f i. It now follows from (1) that [bn ... bin] = [an ... ain] and hence that bin = ain. Thus the n-th columns of A and B coincide and so $A = B$ also in this case. D Corollary The row rank of a matrix is the

number of non-zero rows in any row-echelon form of the matrix. Proof Let B be a row-echelon form of A and let H be the Hermite form obtained from B. Since H is unique, the number of non-zero rows of B is precisely the number of non-zero rows of //, which is the row rank of A. D The uniqueness of the Hermite form means that two given matrices are row- equivalent if and only if they have the same Hermite form. The Hermite form of A is therefore a particularly important 'representative* in the equivalence class of A relative to the relation of being row-equivalent.

46 basic Linear Miyeum EXERCISES 3.14 Consider the matrix 1 1 1 2 1 \ 1 2 1 1 \ 2 1 4 1 4 3-X 6 \ 6 v4 = Show that if \ f 1,2 then the row rank of /\ is 4. What is the row rank of A when \ = 1, and when X = 2? 3.15 Determine whether or not the matrices "2 2 f 1 3 1 1 2 2 • 2 1-1 0 2-1 -3 -2 3 are row-equivalent. Similar to the concept of an elementary row operation is that of an elementary column operation. To obtain this we simply replace 'row* by 'column* in the definition. It should be noted immediately that such column operations cannot be used in the same way as row operations to solve systems of linear equations since they do not produce an equivalent system. However, there are results concerning column operations that are 'dual* to those concerning row operations. This is because column operations on a matrix can be regarded as row operations on the transpose of the matrix. For example, from the column analogues of Theorems 3.1, 3.2 and 3.3 (proved by transposition) we have the analogue of Theorem 3.4: Theorem 3.12 An elementary column operation on an m x n matrix can be achieved by post-multiplication by a suitable elementary matrix, namely that obtained from ln by applying to ln precisely the same column operation. D The notion of column-equivalence is dual to that of row-equivalence. Definition The column rank of a matrix is defined to be the maximum number of linearly independent columns in the matrix. The dual of Theorem 3.9 holds, namely: Theorem 3.13 Column operations do not affect column rank D

3. Systems of Linear Equations 47 Since it is clear that column operations can have no effect on the independence of rows, it follows that column operations have no effect on row rank. We can therefore assert: Theorem 3.14 Row and column rank are invariant under both row and column operations. D EXERCISES 3.16 Determine the row and column ranks of the matrix "0 2 3-4 1" 0 0 2 3 4 2 2-5 2 4' 2 0-6 9 7 We now ask if there is any connection between

the row rank and the column rank of a matrix; i.e. if the maximum number of linearly independent rows is connected in any way with the maximum number of linearly independent columns. The answer is perhaps surprising, and bears out what the reader should have observed in the previous exercise. Theorem 3.15 Row rank and column rank are the same. Proof Given a non-zero m x n matrix A, let H(A) be its Hermite form. Since H(A) is obtained from A by row operations it has the same row rank, p say, as A. Also, we can apply column operations to H(A) without changing this row rank. Also, both A and H(A) have the same column rank, since row operations do not affect column rank. Now by suitable rearrangement of its columns H(A) can be transformed into the the general form I i ' 'p 0 0 ' in which the submatrix marked ? is unknown but can be reduced to 0 by further column operations using the first p columns. Thus H(A) can be transformed by column operations into the matrix \ o" 0 0 ' Now by its construction this matrix has the same row rank and the same column rank as A. But clearly the row rank and the column rank of this matrix are each p. It therefore follows that the row rank and the column rank of A are the same. D

48 Basic Linear Algebra Because of Theorem 3. IS we shall talk simply of the rank of a matrix, meaning by this the row rank or the column rank, whichever is appropriate. The following is now immediate: Corollary rank A = rank A'. D If we reflect on the proof of Theorem 3.15 we see that every non-zero m x n matrix A can be reduced by means of elementary row and column operations to a matrix of the form riP 0 L0 0 the integer p being the rank of A. Invoking Theorems 3.4 and 3.12, we can therefore assert that there is an m x m matrix P and annxn matrix Q, each of which is a product of elementary matrices, such that PAQ = Ip 0 0 0 The matrix on the right-hand side of this equation is often called the normal form of A. How can we find matrices P, Q such that PAQ is in normal form without having to write down at each stage the elementary row and column matrices involved? If A is of size m x n, an expedient way is to work with the array as follows. In reducing A, we apply each row operation to the bottom m rows of this array, and each column operation to the first n columns of this array. The general stage will be an array of the form FXF2-.FS XAY E, ■••E2ElAFlF2 •••Fs tLf ■ ■ tL'jtL 2M where E,, E2,..., E, are elementary matrices corresponding to the row operations and F,, F2,..., Fs are elementary matrices corresponding to the column operations. If N is the normal form of A, the final configuration is Q_ N P

Example 3.26 Consider the matrix A = 1 2 -1 -2 -1 -1 1 1 0 1 2 1 Applying row and column operations as described above, we obtain 1 0 0 0 1 -1 0 0 1 0 0 2 -1 1 0 0 1 0 -i 1 2 0 0 0 1 -2 1 1 1 0 0 0 1 0 0 0 1 1 0 0 1 0 0 ^ 0 0 1 2 0 1 0 1 0 0 1 0 -1 0 2 0 0 0 I -2 -1 1 1 0 1 1 0 0 0 0 1 1 0 0 0 0 1 0 0 0 0 1 0 ^0 0 0 1 1 2 -1 -2 0 1 0 0 0 -I 2 2 1 -1 1 0 -1 1 1 0 0 1 0 0 v 0 0 1 2 0 1 0 0 0 0 1 0 -1 0 1 0 0 0 1 -2 -1 1 1 i 1 i 2 0 1 ~2 0 0 1 2 1 0 0 0 i 0 0 -2 1 1 0 0 1 0 0 0 0 1 0 0 1 2 0 0 1 0 -1 1 i 1 ~2 0 0 1 0 ~2 2 1 0 0 0 1 0 0 -2 1 0 1 0 1 0 1 0 0 0 0 1 0 0 0 1 0 0 0 1 1 1 1 -i 0 0 1 0 1 1 2 2 1 0 0 0 i 0 0 -2 1 1 0 0 1 0 0 0 0 1 0 0 1 -1 1 -1 1 0 0 0 i 1 ~2 0 0 I 0 1 1 2 2

Thus we see that so that A is of rank 3; and N = P = 1 1 L 2 1 0 0 0" 0 1 0 0 0 0 1 0 , Q = "1 0 0 0 -2 1 1 0 0 1 0 0 -1 1 -1 1 A simple computation on the reader's part will verify that PAQ = N. • It should be noted that the matrices P and Q which render PAQ = N are not in general uniquely determined. The reader should do the above example in a different way (e.g. reduce A to Hermite form before doing column operations) to obtain different P and Q. EXERCISES 3.17 Show that the matrix A = 2 -3 -2 1 -2 -3 is of rank 2 and find matrices P, Q such that /2 0 0 0 PAQ = 3.18 Write down all the normal forms that are possible for non-zero 4x5 matrices. 3.19 Suppose that A is an n x n matrix and that m rows of A are selected to form an m x n submatrix B. By considering the number of zero rows in the normal form, prove that rank B ^ m - n + rank A. Definition We say that two mxn matrices are equivalent if they have the same normal form. Since the rank of matrix is the number of non-zero rows in its normal form, it is clear from the above that two mxn matrices are equivalent if and only if they have the same rank. The reader can easily check that the relation of being equivalent is an equivalence relation on the set of mxn matrices. The normal form of A is a particularly simple representative of the equivalence class of A

EXERCISES 3.20 Prove that if A and B are row-equivalent then they are equivalent. 3.21 Prove that every square matrix is equivalent to its transpose. 3.22 Show that the following matrices are equivalent: 1 2 1 5 3 -1 0 2 -7 -1 0 1 2 4 -1 -6 0 0 1 -2 -3 0 1 0 2 1 1 0 0 3 4 We now have to hand enough machinery to solve the problem in hand. This is dealt with in the next three results. Theorem 3.16 If A is an m x n matrix then the homogeneous system of equations Ax = 0 has a non-trivial solution if and only if rank A < n. Proof Let a, be the i-th column of A Then there is a non-zero

column matrix x = such that v4x = 0 if and only if *iai +*2a2 + ■ • • + V»n = 0; for, as is readily seen, the left-hand side of this equation is simply Ax. Hence a non-trivial (=non-zero) solution x exists if and only if the columns of A are linearly dependent. Since A has n columns in all, this is the case if and only if the (column) rank of A is less than n. D Theorem 3.17 A non-homogeneous system Ax — b has a solution if and only if rank A = rank v4|b. Proof If A is of size m x n then there is an n x 1 matrix x such that v4x = b if and only if there are scalars jc, ,..., xn such that *lal + *2a2 + •■+xn*n = *> This is the case if and only if b is linearly dependent on the columns of v4, which is the case if and only if the augmented matrix A\b is column-equivalent to A, i.e. has the same (column) rank as A D

52 Basic Linear aigeui a Definition We shall say that a system of linear equations is consistent if it has a solution (which. in the homogeneous case, is non-trivial); otherwise we shall say that it is inconsistent. Theorem 3.18 Let a consistent system of linear equations have as coefficient matrix the m x n matrix A. If rank A = p then n -p of the unknowns can be assigned arbitrarily and the equations can be solved in terms of them as parameters. Proof Working with the augmented matrix v4|b, or simply with A in the homogeneous case, perform row operations to transform A to Hermite form. We thus obtain a matrix of the form H(A)\c in which, if the rank of A is p, there are p non-zero rows. The corresponding system of equations H{A)x = c is equivalent to the original system, and its form allows us to assign n-poi the unknowns as solution parameters. D The final statement in the above proof depends on the form of H(A). The assignment of unknowns as solution parameters is best illustrated by examples. This we shall now do. It should be noted that in practice there is no need to test first for consistency using Theorem 3.17 since the method of solution will determine this automatically. Example 3.27 Let us determine for what values of a the system jc+ y+ z= 1 2x- y + 2z= 1 jc + 2y + z = a has a solution. By Theorem 3.17, a solution exists if and only if the rank of the coefficient matrix of the system is the same as the rank of the augmented matrix, these ranks being determined by the number of non-zero rows in any row-echelon form. So we begin by reducing the augmented matrix to row-echelon form: 1 1 2 -1 1 2 1 1 0 -3 0 1 "1 1 1 0 1 0 1 1 0 -1 0 a-1 1 1 3 0 1 0 a-1 A* 1 1 1 0 1 0 0 0 0 a-f

3. Systems of Linear Equations 53 It is clear that the ranks are the same (and

hence a solution exists) if and only if In this case the rank is 2 (the number of non-zero rows), and the Hermite form is "1 0 1 f 0 10}. 0 0 0 0 Using Theorem 3.18, we can assign 3 -2 = 1 of the unknowns as a solution parameter. Since the corresponding system of equations is x + z = j we may take as the general solution y=\ and jc = | - z where z is arbitrary. Example 3.28 Consider now the system jc+ y+ z + x+fiy+ z + jc+ > + /?z + (3 2jc + 2> + 2z The augmented matrix of the system is "111 1 4" 1 P 1 14 1 1 0 3-0 6 2 2 2 0 6 which can be reduced to row-echelon form by the operationsP2 ~Pi» P3 ~Pi» ar,d p4-2p,. We obtain "11 1 14" 0/9-10 0 0 K) 0 0 /3-1 2-0 2 0 0 0 P-2 -2 Now if p f 1,2 then the rank of the coefficient matrix is clearly 4, as is that of the augmented matrix. By Theorem 3.18, therefore, a unique solution exists (there being no unknowns that we can assign as solution parameters). / = 4 / = 4 -/?)/ = 6 /?/ = 6.

54 Basic Linear Algebra To find the solution, we reduce the above row-echelon matrix to Hermite form: 1 1 1 0 1 0 0 0 1^-^- u u * /3-1 0-1 POO 1 ^ 1111 0 10 0 0 0 10 0 0 0 1 4 0 0 /3-2J n o o o 4 + ^2 0 10 0 0 0 10 pooi ^ The system of equations that corresponds to this is x = 4 + -2- H T /3-2 y = 0 z = 0 ' /3-2 which gives the solution immediately. Consider now the exceptional values. First, let /? = 2. Then the matrix (2) becomes '1111 4" 0 10 0 0 0 0 10 2 0 0 0 0 -2 and in the system of equations that corresponds to this augmented matrix the final equation is Ojc + Oy + 0z + 0/ = -2, which is impossible. Thus when 0 = 2 the system is inconsistent. If now p - 1 then the matrix (2) becomes 1111 4 0 0 0 0 0 0 0 0 0 1 2 0 0 0-1-2

3. Systems of Linear Equations 55 which reduces to the Hermite form 1110 2 0 0 0 12 0 0 0 0 0 0 0 0 0 0 The corresponding system of equations is x + y + z - 2, r = 2. Here the coefficient matrix is of rank 2 and so we can assign 4-2 = 2 of the unknowns as solution parameters. We may therefore take as the solution ' = 2, jc = 2-y-z where v and z are arbitrary. EXERCISES 3.23 Show that the system of equations 2y 2y + 3z 3r = 3 3/ = 1 / = 3 2/ = 1 has no solution. 3.24 For what value of a does the system of equations jc - 3v - z - 10/ = a x+y+z = 5 2jc -4/=7 jc + y + / = 4 have a solution? Find the general solution when a takes this value. 3.25 Show that the equations 2x + y + z = -6a 2jc + y + {0 + \)z = 4 /?jc + 3> + 2z = lot has a unique solution except when /3=0 and when 0=6. If /3 = 0 prove that there is only one value of a for which a solution exists, and find the general solution in this case. Discuss the situation when 0 =

6.

SUPPLEMENTARY EXERCISES 3.26 Let A be a real 4x3 matrix. Determine the elementary matrices which, as post-multipliers, represent each of the following operations (1) multiplication of the third column by -2; (2) interchange of the second and third columns; (3) addition of -2 times the first column to the second column. 3.27 If the real matrix 'a 1 a 0 0 0" 0 6 16 0 0 0 0 c 1 c 0 0 0 0 d \ d has rank r, prove that (1) r>2; (2) r = 3 if and only if a = d = 0 and be = 1; (3) r = 4 in all other cases. 3.28 Given the real matrices "3 2 -1 5' 1-12 2 0 5 7a , B = ro 3" 0 -1 0 6 prove that the matrix equation AX = B has a solution if and only if a = -1. 3.29 Show that the equations x - y - u - 5t - a 2jc+ y - z-4w+ t = P jc+ y + z - 4u -6/ = 7 jc + Ay + 2z - 8w - 5/ = 6 have a solution if and only if 8a-/?-ll7 + 56 = 0. Find the general solution when a = P = -1,7 = 3,6 = 8. 3.30 Discuss the system of equations -2jc + (n + 3)y - \z = -3 jc + \z = 1 2jc + 4y + 3Xz = -X. 3.31 How many different normal forms are there in the set ofmxn matrices?

3. Systems of Linear Equations 57 3.32 Give an example of a matrix that cannot be reduced to normal form by means of row operations alone. 3.33 Give an example of two matrices that are equivalent but are not row-equivalent. 3.34 Prove that two matrices are equivalent if and only if one can be obtained from the other by a finite sequence of row and column operations. 3.35 Determine the values of a, b,c for which the parabola $y = ox^2 + bx + c$ passes through the three points (-1,1), (0,2), (1,3). 3.36 Determine the set of integer solutions of the system 3jc, + jc2 + 2jc3 - jc4 = 2jc, + jc2 + 4jc3 + 3jc4 + 2jc5 = jc, - 2jc2 + 3*4 + *5 = *i — *3 — *4 + 3*e zE 2*, + 3*2 + 3*3 - *4 - JC5 = in which =5 denotes congruence modulo 5. ASSIGNMENT EXERCISES (1) Determine the rank of the matrix 0 2 4 4 8 2 6 3 (2) If r is the rank of the matrix 1 -/? 0 0 4 4 0 2 a 1 -1 -6 2 8 10 9 0 0 1 1 2" 0 2 1 0" 0 1 6 show that {a) r>l; (b) r = 2 if and only if ap - -1 and ^ = 6 = 0; (c) r = 3 if and only if either*) = 6 or aft = -1 and ^,6 are not both zero. (3) If A, B, C are (rectangular) matrices such that A = BC prove that rank A^ min{rank B, rank C}.

(4) Let A be an m x n matrix. Prove that the rank of A is the smallest integer p such that A can be written as a product BC where B'ismx p and C'ispx n. (5) Show that if X f 0 then the system of equations \x{+

atxn+x = bt (/ = 1,..., n) n £ fli*i + **n+l = bn+l n has a unique solution if and only if X2 - £ af f 0. i=l (6) Consider on the one hand the system 5, of n linear equations in n unknowns: a\\x\ + fll2*2 + 0|3*3 + •■• + fl|nJfn = bx\ a2lxx + a22*2 + a23jr3 + ••• + fl2n^ = fc2; fl3l*l + fl32*2 + fl33*3 + ■" + a3nXn = fc3; fl«i*i + fl«2*2 + fl,,3^3 + ••■ + amxn = bn, and on the other hand the homogeneous system S2 of n linear equations in n + 1 unknowns: fln>i + fl|2>2 + fli3>3 + ■•■ + fl|n>n - *>i)Vi = 0; 021?! + fl22>2 + 023?3 + " ' + flfcJn " ^n+l = °I «31>1 + fl32>2 + a33>3 + ■■■ + a3«>n " Mn+I = °i flni>i + an2y2 + a,,3>3 + ••• + anny,, - b,,yn+l = 0. Show that the homogeneous system associated with S, has only the trivial solution if and only if every non-trivial solution of 52 is such that yn+l f 0.

4 Invertible Matrices In Theorem 1 3 we showed that every mxn matrix A has an additive inverse, denoted by -A, which is the unique mxn matrix X that satisfies the equation A + X = 0. We shall now consider the multiplicative analogue of this. Definition Let A be an m x n matrix. Then an n x m matrix X is said to be a left inverse of A if it satisfies the equation XA = /,,; and a right inverse of A if it satisfies the equation AX = Im. Example 4.1 Consider the matrices A = •1 4 3 .0 r 3 4 0. »*-[-: -3 3 1 0 a 0 1 b A simple computation shows that Xaj,A = /2, and so A has infinitely many left inverses. In contrast, A has no right inverse. To see this, it suffices to observe that if Y were a 2 x 4 matrix such that AY = /4 then we would require [AY]4 4 = 1 which is not possible since all the entries in the fourth row of A are 0. Example 4.2 The matrix "1 0 0 2 0 0 o" 0 3 has a common unique left inverse and unique right inverse, namely "1 0 0 0 1 2 0 0" 0 1 3.

60 Basic Linear Algebra Theorem 4.1 Let Abe an m x n matrix. Then {1) A has a right inverse if and only if rank A = m\ (2) A has a left inverse if and only if rank A - n. Proof (1) Suppose that the n x m matrix X is a right inverse of A, so that we have AX = /,,,. If X; denotes the i-th column of X then this equation can be expanded to the m equations (i= l,...,»t) Ax, = A(where A, denotes the i-th column of /,,,. Now each of the matrix equations Ax, = A- represents a consistent system of m equations in n unknowns and so, by Theorem 3.17, for each i we have rank A - rank A\AS. Since A,,..., A,,, are linearly independent, it follows by considering column ranks that rank A = rank v4|A, = rank j4|A, |A2 = rank A\AXlA^... |A,, = rank A\Im = m. Conversely, suppose that the rank of A is m. Then necessarily we have that n^m. Consider the Hermite form of

A'. Since H(A') is an n x m matrix and rank H(A') = rank A' = rank A- m, we see that H(A') is of the form 1- H(A') = 0,,-m, As this is row-equivalent to A\ there exists annxn matrix Y such that yA' = ^n-m,m Taking transposes, we obtain Now let Z be the n x m matrix consisting of the first m columns of Y'. Then from the form of the immediately preceding equation we see that AZ = /m, whence Z is a right inverse of A. (2) It is an immediate consequence of Theorem 1.9 that A has a left inverse if and only if its transpose has a right inverse. The result therefore follows by applying (1) to the transpose of A. D

4. Invertible Matrices 61 EXERCISES 4.1 Show that the matrix '13 4 7" 2 3 5 8 14 5 9 has neither a left inverse nor a right inverse. Theorem 4.2 If a matrix A has both a left inverse X and a right inverse Y then necessarily [1] A is square; (2) X = Y. Proof (1) Suppose that A is of size m x n. Then by Theorem 4.1 the existence of a right inverse forces rank A = m, and the existence of a left inverse forces rank A = n. Hence m = n and so A is square. (2) If A is of size p x p then XA = Ip = AY gives, by the associativity of matrix multiplication, X = XIp = X{AY) = (XA)Y=IpY=Y. D For square matrices we have the following stronger situation. Theorem 4.3 If A is an n x n matrix then the following statements are equivalent: (1) A has a left inverse; (2) A has a right inverse; (3) A is of rank n\ (4) the Hermite form of A is /,,; (5) A is a product of elementary matrices. Proof We first establish the equivalence of (1), (2), (3), (4). That (1), (2), (3) are equivalent is immediate from Theorem 4.1. (3) =>• (4): If A is of rank n then the Hermite form of A must have n non-zero rows, hence n corner entries 1. The only possibility is /,,. (4) =>{3): This is clear from the fact that rank /,, = n. We complete the proof by showing that (3) =>■ (5) and (5) =>■ (3) (3) =>• (5): If A is of rank n then, since (3) => (1), A has a left inverse X. Since XA = /,, we see that X has a right inverse A so, since (2) => (4), there is a finite string of elementary matrices Fl,F2,...}Fq such that Fq-F2FlX = ln.

62 Basic Linear Algebra Consequently, we have A = InA = {Fq • • • F2FXX)A = Fq • • - F2FX (XA) = Fq ■ ■ ■ F2FX and so A is a product of elementary matrices. (5) =» (3): Suppose now that A = EXE2-Ep where each £, is an elementary matrix. Observe that Ep is of rank n since it is obtained from /,, by a single elementary operation which has no effect on rank. Also, pre-multiplication by an elementary matrix is equivalent to an elementary row operation, which has no effect on rank. It follows that the rank of the product

ExE2>Ep is the same as the rank of £p, which is n. Thus the rank of Aisn. D It is immediate from the above important result that if a square matrix A has a one-sided inverse then this is a two-sided inverse (i.e. both a left inverse and a right inverse). In what follows we shall always use the word 'inverse' to mean two-sided inverse. By Theorem 4.2, inverses that exist are unique. When it exists, we denote the unique inverse of the square matrix A by A~l. Definition UA has an inverse then we say that A is invertible. If A is an invertible nx n matrix then so is A~l. In fact, since AA~l = In = A~lA and inverses are unique, we have that A is an inverse of A-1 and so {A~l)~l = A. We note at this juncture that since, by Theorem 4.3, every product of elementary matrices is invertible, we can assert that B is row-equivalent to A if and only if there is an invertible matrix E such that B = EA. Another useful feature of Theorem 4.3 is that it provides a relatively simple method of determining whether or not A has an inverse, and of computing A~* when it does exist. The method consists of reducing A to Hermite form: if this turns out to be /,, then A is invertible; and if the Hermite form is not /,, then A has no inverse. In practice, just as we have seen in dealing with normal forms, there is no need to compute the elementary matrices required at each stage. We simply begin with the array A\In and apply the elementary row operations to the entire array. In this way the process can be described by A\I,, ^ ElA\El ** E2^i^l^2^i *^ ■ • ■ . At each stage we have an array of the form S\Q = Er-E2ExA\Er-E2Ex in which QA - S. UA has an inverse then the Hermite form of A will be /,, and the final configuration will be h\Ep-E2Ex so that Ep • • • E2EXA = I,, and consequently A~l = Ep • • • E2EX.

4. Invertible Matrices 63 Example 4.3 Consider the matrix Applying the above procedure, we obtain "l 2 1 3 1 4 3] 4 4 1 2 3 1 3 4 1 4 4 1 0 0 namely 0 1 0 0 0 1 " 4 0 -1 ■** ■** ■** ■** ■** -4 -1 2 1 0 0 1 0 0 1 0 0 1 0 0 1 0 0 2 1 2 2 1 0 2 1 0 2 1 0 0 1 0 I" 1 -1 3 1 1 3 1 -1 3 0 -1 0 0 -1 0 0 1 1 -1 -1 1 -1 1 1 0 1 4 0 1 4 0 -1 0 0 1 0 0 1 0 0 1 0 -2 1 0 0 -1 1 -2 1 -6 3 -1 1 -2 1 -4 1 -1 1 2 -1 EXERCISES 4.2 Determine which of the following matrices are invertible and find the inverses: "l 1 f 1 2 3 0 1 1 > "1 2 f 1 3 2 1 0 1 > "I 2 2" 1 3 1 1 1 3 1111 12-12 1-12 1 13 3 2 11 11 13 12 12-11 5 9 16 4.3 If A is an n x n matrix prove that the homogeneous system Ax - 0 has only the trivial solution x = 0 if and only if A is invertible.

64 Basic Linear Algebra 4.4 Prove that the real 2x2 matrix A = a b c d "l 1 1 1

0 2 0" 0 a is invertible if and only if ad-bcf 0, in which case find its inverse. 4.5 Determine for which values of a the matrix A = is invertible and describe A"1. We shall see later other methods of finding inverses of square matrices. For the present, we consider some further results concerning inverses. We first note that if A and B are invertible nxn matrices then in general A + Bis not invertible. This is easily illustrated by taking A = /,, and B = -In and observing that the zero nxn matrix is not invertible. However, as the following result shows, products of invertible matrices are invertible. Theorem 4.4 Let A and B be nxn matrices. If A and B are invertible then so is AB\ moreover, (AB)'" = B~lA~l. Proof It suffices to observe that ABB'1 A'1 = AI,,A~l = AA~l = /,, whence B~XA~X is a right inverse of AB. By Theorem 4.3, AB therefore has an inverse, and {AB)~l = B~lA~l. D Corollary If A is invertible then so is Am for every positive integer m\ moreover, (Am)'" = (A-')m. Proof The proof is by induction. The result is trivial for m = 1. As for the inductive step, suppose that it holds for m. Then, using Theorem 4.4, we have {A~l)m+l = A-l(A~l)m = A-l(Amyl = (AmA)-1 = (A"*1)"1, whence it holds for m + 1. D

4. Invertible Matrices 65 Theorem 4.5 If A is invertible then so is its transpose A'\ moreover, (AT1 = (A'1)'. Proof By Theorem 4.4 we have /,, = /; = (AA-ly = (A-lyA* and so (A~l)' is a left inverse, hence the inverse, of A'. D EXERCISES 4.6 If A,, A2,..., Ap are invertible n x n matrices prove that so also is the product AXA2 • ■ ■ Ap> and that 4.7 Let i? be a fixed real number and let 0 1 -sintf A = -1 0 costf — sin i9 cost? 0 Show that A3 = 0. For each real number jc define the 3 x 3 matrix Ax by Ax = I3 + xA + ±x2A2. Prove that AxAy = Ax+y and deduce that each Ax is invertible with A~l = 4.8 For each integer n let 1-« -n n 1 +n Prove that AnAm = An+m. Deduce that An is invertible with /l^1 = A_,,. Do the same for the matrices " [-An 1+2« What is the inverse of AmBml A = Recall that annxn matrix is orthogonal if it is such that AA' = /,, = A'A. By Theorem 4.3, we see that in fact only one of these equalities is necessary. An orthogonal matrix is therefore an invertible matrix whose inverse is its transpose

66 Basic Linear Algebra Example 4.4 A 2 x 2 matrix is orthogonal if and only if it is of one of the forms ' a b' -b a * a b b -a in which a2 + ^2 = 1. In fact, suppose that '-til is orthogonal. Then from the equation [: :i t;] — t \A'-- we obtain a2 + b2=\, ac + bd = 0} c2 + d2=\. The first two equations show that if a = 0 then d = 0, and the second and third equations show that ifd = 0 then a = 0.

Now when a = d = 0 the equations give b = ±1 and c = ±1. On the other hand, when a and J are not zero the middle equation gives c/d - ~{b/a), so either c = -b, d = a or c = b, d = -a. It follows that in all cases the matrix is of one of the stated forms. Note that the one on the left is a rotation matrix (take a = cos i? and b = sin i9). Example 4.5 The matrix l/v/3 1/vf -√y/T l/v/3 -2/v^ 0 l/v/3 √y/S 1/vT is orthogonal. If P and Aarcnxn matrices with P invertible then for all positive integers m we have {P-XAP)M = p-lAmP. The proof of this is by induction. In fact the result is trivial for m = 1; and the inductive step follows from {P'lAP)m+} = p-xAP(p-xAP)m = P~lAPP-lAmP = p-lAI,,AmP = p-'Am+,P.

4. Invertible Matrices 67 In certain applications it is important to be able to find an invertible matrix P such that P~XAP is of a particularly simple form. Consider, for example, the case where P can be found such that P~lAP is a diagonal matrix D. Then from the above formula we have ZT = {p-lAP)m = p-lAmp. Consequently, by multiplying the above equation on the left by P and on the right by P~l, we see that Am = pirp-1. Since VF is easy to compute for a diagonal matrix D (simply take the m-th power of the diagonal entries), it is then an easy matter to compute Am. The problem of computing high powers of a matrix is one that we have seen before, in the 'equilibrium-seeking* example in Chapter 2, and this is precisely the method that is used to compute Ak in that example. Of course, how to determine precisely when we can find an invertible matrix P such that P~lAP is a diagonal matrix (or some other 'nice* matrix) is quite another problem. A similar problem is that of finding under what conditions there exists an orthogonal matrix P such that PAP is a diagonal matrix. Why we should want to be able to do this, and how to do it, are two of the most important questions in the whole of linear algebra. A full answer is very deep and has remarkable implications as far as both the theoretical and practical sides are concerned. EXERCISES 4.9 If A is an n x n matrix such that /,, + A is invertible prove that (In-A)(In + A)-l = (In+A)-l(In-A). Deduce that if P=(!n-A)(ln + A)-1 then P is orthogonal when A is skew-symmetric. Given that 0 cosi? 0 A = -cos i? 0 sin i9 0 -sin i? 0 prove that P = sin219 -cost? sintfcost? costf 0 —sin i? sin i? cos t? sin i? cos21?

68 Basic Linear wyema SUPPLEMENTARY EXERCISES 4.10 Find the inverse of H 1 2 r 0-200 12 1-2 0 3 2 1 4.11 Let A and B be row-equivalent nxn matrices. If A is invertible, prove that so also is B. 4.12 Let A and B be

nxn matrices. If the product AB is invertible, prove that both A and B are invertible. 4.13 If A and B are n x n matrices with A invertible, prove that {A + B)A~\A -B) = (A -B)A~\A + B). 4.14 If A is an n x n matrix establish the identity ln-Ak+l = (ln-A)(ln+A+A2 + -- + Ak). Deduce that if some power of A is the zero matrix then /,, -A is invertible. Suppose now that '2 2-1 -r -10 0 0 -1-110 0 1-11 A = Compute the powers (74 -A)' for i = 1,2,3,4 and, by considering A = I4 - (/4 -A), prove that A is invertible and determine A~l. 4.15 Let A and B be n x n matrices such that AB - /,, is invertible. Show that (BA-In)[B(AB-InylA-In] = ln and deduce that BA - /,, is also invertible. ASSIGNMENT EXERCISES (1) If A e MatnxlI R is such that al} = 0 for i = j and a%i = 1 for i ^,, show that A satisfies a polynomial equation of degree 2. Hence determine A'1. (2) Determine the inverse of the complex matrix 1 l + i - i 0 i l-2i 1 1 i

5 Vector Spaces In order to proceed further with matrices we have to take a wider view of matters. This we do through the following important notion. Definition By a vector space we shall mean a set V on which there are defined two operations, one called 'addition' and the other called 'multiplication by scalars', such that the following properties hold: (V,) jc + v = v + jc for all jc, y G V; (V2) (jc + y) + z = jc + {y + z) for all x,y,z G V; (V3) there exists an element 0 € V such that jc + 0 = jc for every jc € V; (V4) for every jc e V there exists -jc G V such that jc + (-jc) = 0; (V5) X(jc +v)= \x + \y for all jc,> G V and all scalars X; (V6) (^ + n)x = ^x + n* for all jc G V and all scalars X, ^; (V7) (^n)x - Mm*) *°r all jc € V and all scalars X,^; (V8) ljc = jtforalljcG V. When the scalars are all real numbers we shall often talk of a real vector space; and when the scalars are all complex numbers we shall talk of a complex vector space. • It should be noted that in the definition of a vector space the scalars need not be restricted to be real or complex numbers. They can in fact belong to any 'field' F (which may be regarded informally as a number system in which every nonzero element has a multiplicative inverse). Although in what follows we shall find it convenient to say that 'V is a vector space over a field F* to indicate that the scalars come from a field F, we shall in fact normally assume (i.e. unless explicitly mentioned otherwise) that F is either the field IR of real numbers or the field C of complex numbers. • Axioms (V,) to (V4) above can be summarised by saying that the algebraic structure (V; +) is an abelian group. If we denote by F the field of scalars

70 Basic Linear Algebra (usually IR or C) then multiplication by scalars can be considered as an action by F on V, described by (X, jc) »-► Xjc, which relates the operations in F (addition and multiplication) to that of V (addition) in the way described by the axioms (V5) to (V8). Example 5.1 Let MatmxlI IR be the set of all m x n matrices with real entries. Then Theorems 1.1 to 1.4 collectively show that Matmxn IR is a real vector space under the usual operations of addition of matrices and multiplication by scalars Example 5.2 The set IR" of n-tuples (jc, ,..., jc,,) of real numbers is a real vector space under the following component-wise definitions of addition and mutiplication by scalars: (*i xn) + (yu...tyn) = (*i +?i *,,+?,,), M*i,•••,*,,) = (^xu...t\xn). Geometrically, IR2 represents the cartesian plane, whereas IR3 represents three- dimensional space. Similarly, the set C" of n-tuples of complex numbers can be made into both a real vector space (with the scalars real numbers) or a complex vector space (with the scalars complex numbers). Example 5.3 Let Map(IR, IR) be the set of all mappings/: IR -• R. For two such mappings/, g define/ + g : IR -»IR to be the mapping given by the prescription (f+*)M=/M + *M, and for every scalar X € IR define X/ : R -» R to be the mapping given by the prescription W)M = x/W. Then it is readily verified that (V,) to (V8) are satisfied, with the role of the vector 0 taken by the zero mapping (i.e. the mapping t? such that t?(x) = 0 for every jc € IR) and -/ the mapping given by (-/)(*) = -/(jc) for every jc € R These operations therefore make Map(IR, IR) into a real vector space. Example 5.4 Let IR,,[X] be the set of polynomials of degree at most n with real coefficients The reader will recognise this as the set of objects of the form a0 + axX + a2X2 + ■ • • + anXn

5. Vector Spaces 71 where each a, G F and X is an 'indeterminate', the largest suffix i for which at f 0 being the degree of the polynomial. We can define an addition on IR,,[X] by setting {aQ + axX + .. + anXn) + {b0 + bxX + • • • + ^X") = («o + *o) + («« +*«)* + - + (0. + 6,,^ and a multiplication by scalars by \(a0 + axX + • ■ • + a.X") = Xa0 + ^i* + • ■ • + \anXn. In this way IR,,[X] has the structure of a real vector space. We now list some basic properties of the multiplication by scalars. For clarity, we shall denote (at least for a while) the additive identity element of V by Ov and that of F (i.e. IR or C) by Of. We also use the symbol V as an abbreviation of 'for all*. Theorem 5.1 If V is a vector space over afield F then (1) (VXGF) X0V = 0V; (2) (Vjc G V) Of jc = 0/, (3) if \ x = Ov then either \ = 0Forx = Ov; (4) (Vjc G V)(VX G F) (-X)jc = -(X*) = X(-*). Proof (0 By (V3) and (V5) we have X0V = X(0V + 0V) = \0V + \0V.

Now add ~{\0V) to each side. (2) By (V6) we have 0f x = (Of + Of)jc = Of jc + Of jc. Now add -(0f jc) to each side. (3) Suppose that \x = 0V and that X f 0f. Then X has a multiplicative inverse X"1 and so, by (V7) and (1), jc = lf jc = (X_,X)jc = X-'(Xjc) = X-'Ov = 0V. (4) By (2) and (V6) we have Ov=(X + (-X)]jc=Xjc + (-X)jc. Now add -(Xjc) to each side. Also, by (1) and (V5) we have 0V = X(jc + (-jc)] = Xjc + X(-jc). Now add -(Xjc) to each side. D EXERCISES 5.1 Verify the various items of Theorem 5.1 in the particular case of the vector space MatOTXn IR.

72 Basic Linear Algebra In order to study vector spaces we begin by concentrating on the substructures, i.e. those subsets that are themselves vector spaces. Definition Let V be a vector space over a field F. By a subspace of V we shall mean a non-empty subset W of V that is closed under the operations of V, in the sense that (1) ifjc.y G Wthenjc + y £ W; (2) if jc G W and X G Flhcn \x G W. Note that (1) says that sums of elements of W belong also to W, and (2) says that scalar multiples of elements of W also belong to W. When these properties hold, W inherits from the parent space V all the other properties required in the definition of a vector space. For example, by taking X = -lf in (2) we see that if jc G W then -jc G W; and then, by taking y = -jc in (1), we obtain 0V G W. Example 5.5 Every vector space V is (trivially) a subspace of itself. V itself is therefore the biggest subspace of V. Example 5.6 By Theorem 5.1(1), the singleton subset {0V} is a subspace of V. This is then the smallest subspace of V since, as observed above, we have that 0V G W for every subspace W of V. Example 5.7 R is a subspace of the real vector space C. In fact, it is clear that properties (1) and (2) above hold with W = IR and F = IR. Example 5.8 In the real vector space IR2 the set X = {(jc,0) ; jc G IR} is a subspace; for we have (*i,0)+(*2,0)=(*,+*2,0); \(x,0) = (Xjc.O), and so X is closed under addition and multiplication by scalars. Thus (1) and (2) are satisfied. This subspace is simply the 'jc-axis' in the cartesian plane IR2. Similarly, the 'y-axis* >'={(0,y);>GiR} is a subspace of IR2.

5. Vector Spaces 73 Example 5.9 We can expand on the previous example. In the cartesian plane IR2 every line L through the origin has a description in the form of a particular subset of IR2, namely L = {{x,y);ax + Py = 0}. If L makes an angle t? with the jc-axis then the gradient of L is given by tan t? = -a/0. Nowif(jc,,y,) G Land(jc2,y2) G £ then we have ckjc, = -/?>, and a a 2 = -fiyi whence a(jc, + x2) = -fi[y\ + yi) and therefore (^1,^1) + (^2,^2) = (^1 +^2,^1

+>2)G^ and if (jc,, >,) G L then ajc, = -/3y, gives aXjc, = -fi\yx for every \ G IR, so that >(^i,>i) =(^1,^1) G^. Thus we see that every line L that passes through the origin is a subspace of IR2. As we shall see later, apart from {(0,0)} and IR2 itself, these are the only sub- spaces of IR2. Example 5.10 In the cartesian space IR3 every plane through the origin has a description in the form P={(x,y,z)\ ax + 0y + iz = O}. To see the geometry of this, observe that if we fix z, say z = kt then 'the plane z = ky (i.e. the set {(jc, y, k); jc, y G IR}) slices through P in the line {{x,y,k); ax + 0y = -7k}. Nowif (jc,,y,,z,) G ^and(jc2,y2,Z2) G P then it is readily seen that (*i.yi.Zi) + (*2,?2,Z2) = (-^1 +- ^2»>i + >2.Zi + Z2) G P\ and if (jc, , y,, z,) G P then, for every X G IR, HxuyuZi) = (\xlt\yu\zi)^P- Thus we see that every plane through the origin is a subspace of IR3. We shall see later that, apart from {(0,0,0)} and IR3 itself, the only subspaces of IR3 are lines through the origin and planes through the origin. Example 5.11 Annxn matrix over a field F is said to be lower triangular if it is of the form [flll fl2I «31 -flnl 0 0 . a22 0 . fl32 fl33 • atil fln3 • ' 01 . 0 . 0 • 0iw-

74 Basic Linear Algebra i.e. if ay - 0 whenever i < j. The set of lower triangular n x n matrices is a subspace of the vector space Mat,,XII F. In fact, if A and B are lower triangular then clearly so is A + B and so is \A. EXERCISES 5.2 In the vector space IR4 of 4-tuples of real numbers, determine which of the following subsets are subspaces: 0) {(x,y,z,t)> x = y, z = t}\ (2) {(*,y,z,/); x + y + z + t = 0}; (3) {{x,y,z,t)\ *=1}; (4) {(*,y,z,0; xt = yz). 5.3 Consider the set Con(IR.IR) of all continuous functions/ : IR -> IR. What well-known theorems from analysis ensure that Con(IR, R) is a subspace of the vector space Map(IR, IR)? 5.4 Let Diff (IR, R) be the subset of Con(lR, IR) consisting of all differen- tiable functions/: IR —»IR. What well-known theorems from analysis ensure that Diff (IR, IR) is a subspace of the vector space Con (IR, IR)? 5.5 Determine which of the following are subspaces of the vector space MatnxlIIR: (1) the set of symmetric nxn matrices; (2) the set of invertible nxn matrices; (3) the set of non-invertible nxn matrices. 5.6 If A is a real mx n matrix prove that the solutions of the homogeneous system Ax = 0 form a subspace of the vector space Matnx, IR. 5.7 Show that a line in R3 that does not pass through the origin cannot be a subspace of IR3. As the above examples illustrate, in order to show that a given set with operations is a vector space it is often very easy to do so by proving that it is a subspace of some well-known and considerably larger vector space. Suppose now that A

and B are subspaces of a vector space V over a field F and consider their intersection A OB. Since 0V must belong to every subspace we have that 0V G A and 0V G B, and therefore 0V G A D B so that A D B f 0. Now if xty G ADBthen jc,y G A gives jc + y G A, andjc.y G B gives Jc + y G B, whence we have that jc + y G A n B. Likewise, if jc G A f) B then jc G A gives \x G A, and x G B gives \x G B, whence we have that \x G A n B. Thus we see that A D B is also a subspace of V.

5. Vector Spaces 75 We can in fact prove a much more general statement: Theorem 5.2 The intersection of any set of subspaces of a vector space V is a subspace of V. Proof Let C be a set of subspaces of V and let T be their intersection. Then 7^0 since every subspace of V (and therefore every subspace in C) contains Or/, whence so also doesf. Suppose now that x,y € T. Since jc and y belong to every subspace W in the set C, so does jc + y and hence jc + y G T. Also, if jc € T then jc belongs to every subspace W in the set C, whence so does Xjc and so Xjc € T. Thus we see that T is a subspace of V. D In contrast with the above situation, we note that the union of a set of subspaces of a vector space V need not be a subspace of V: Example 5.12 In IR2 the jc-axis X and the y-axis Y are subspaces, but X U Y is not. For example, we have (1,0) G X and (0,1) G K, but (1,0)+(0,1) = (1, \)£XuY so the subset X U Y is not closed under addition and therefore cannot be a subspace. Suppose now that we are given a subset S of a vector space V (with no restrictions, so that S may be empty if we wish). The collection C of all the subspaces of V that contain S is not empty, for clearly V itself belongs to C. By Theorem 5.2, the intersection of all the subspaces in C is also a subspace of V, and clearly this intersection also contains S. This intersection is therefore the smallest subspace of V that contains S (and is, of course, S itself whenever 5 is a subspace). We shall denote this subspace by (5). Example 5.13 In IR2 consider the singleton subset S = {(jc,>)}. Then (5) is the line joining (jc,y) to the origin. Our immediate objective is to characterise the subspace (5) in a useful alternative way. For this purpose, we consider first the case where S is not empty and introduce the following notion. Definition Let V be a vector space over a field F and let 5 be a non-empty subset of V. Then we say that v € V is a linear combination of elements of 5 if there exist jct,..., jc,, € S

76 Basic Linear Algebra and X,,..., X,, € F such that n i=I n m It is clear that if v = £ Kxi and m> = £ M;?.are'>near combinations of elements i=l i=I of S then

so is v + *v; moreover, so is Xv for every X G F. Thus the set of linear combinations of elements of 5 is a subspace of V. We call this the subspace spanned by S and denote it by Span 5. The above notions come together in the following result. Theorem 5.3 (5) = Span 5. Proof For every jc € S we have jc = lf jc € Span S and therefore we see that S C Span S. Since, by definition, (5) is the smallest subspace that contains 5, and since Span S is a subspace, we see that (5) C Span S. For the reverse inclusion, let jc, ,..., jc,, € S and X,,..., X,, € F. If W is any subspace of V that contains S we clearly have jc, ,..., jc,, € W and Consequently we see that Span S CW. Taking W in particular to be (S), we obtain the result. D An important special case of the above arises when Span S is the whole of V. In this case we often say that 5 is a spanning set of V. Example 5.14 Consider the subset S = {(1,0), (0,1)} of the cartesian plane R2. For every (.*, y) € IR2 we have (x.y) = (jt.O) + (0,y) - jt(I,0) + y(0,1), so that every element of IR2 is a linear combination of elements of S. Thus 5 is a spanning set of IR2. Example 5.15 More generally, if the n-tuple 6, = (0,...,0,1,0,...,0) has the 1 in the i-th position then for every (jc, ,..., jcJ G IR" we have (*i ^) = ^1 + --- + -½. Consequently, {elt...,en} spans IR".

5. Vector Spaces 77 Example 5.16 In IR3 we have Span{(l,0,0)} = {X(l,0,0); X € IR} = {(X.0,0); X G IR}; i.e. the subspace of IR3 spanned by the singleton {(1,0,0)} is the jc-axis. Example 5.17 In IR3 we have Span{(l,0,0),(0,0,l)} = {jc(l,0,0) + z(0,0,l);jc,zGlR} = {(x.O.z); jc.zgIR}; i.e. the subspace of IR3 spanned by the subset {(1,0,0), (0,0,1)} is the 'jc, z-plane*. EXERCISES 5.8 Show that for all a, btc G IR the system of equations X 4 y + z = a x + 2y + 3z = b x + 3v + 2z = c is consistent. Deduce that the column matrices 1 1 1 *1 2 3 • 1 3 2 span the vector space Mat3x, IR. 5.9 Let IR2[X] be the vector space of all real polynomials of degree at most 2. Consider the following elements of IR2[*]: p(X) = 1 + 2X + X\ q(X) = 2 + X2. Does {p(X),q(X)} span IR: iW .10 Does the set {[i :]• span the vector space Mat2 "0 0' 1 1 x2R? * "1 0' 0 1 ■c :]} We now formalise, for an arbitrary vector space, a notion that we have seen before in dealing with matrices.

78 Basic Linear Algebra Definition Let 5 be a non-empty subset of a vector space V over a field F. Then S is said to be linearly independent if the only way of expressing 0V as a linear combination of elements of S is the trivial way (in which all scalars are 0F). Equivalently, S is linearly independent if, for any given xu..., xn e S, we have Mi + - + UB0v => >, = -•- = >kB = oF.

Example 5.18 The subset {(1,0), (0,1)} of IR2 is linearly independent. For, if \,(1,0) + \2(0,1) = (0,0) then (X,, \2) = (0.0) an<*hence *i = *2 = °f- Example 5.19 More generally, if et = (0,..., 0,1,0,...,0) with the 1 in the i-th position then {el,..., en} is a linearly independent subset of the vector space IR". Example 5.20 Every singleton subset {x} of a vector space V with x f 0 is linearly independent. This is immediate from Theorem 5.1(3). The following result is proved exactly as in Theorem 3.7: Theorem 5.4 No linearly independent subset of a vector space V can contain 0V. D A subset that is not linearly independent is said to be linearly dependent. Note that, by the last example above, every dependent subset other than {0V} must contain at least two elements. Linearly dependent subsets can be characterised in the following useful way, the proof of which is exactly as in that of Theorem 3.8: Theorem 5.5 Let V bea vector space over afield F. IfS is a subset ofV that contains at least two elements then the following statements are equivalent: (1) S is linearly dependent; (2) at least one element ofS can be expressed as a linear combination of the other elements ofS. D Example 5.21 The subset {(1,1,0), (2,5,3), (0,1,1)} of IR3 is linearly dependent. In fact, we have (2,5,3) = 2(1,1,0) + 3(0,1,1).

5. Vector Spaces 79 Example 5.22 In the vector space VR2[X] consider p{X} = 2 + X + X2t q{X} = X + 2X2, r{X} = 2 + 2X + 3X2. A general linear combination of these vectors is 2X, +2>3 + (X, + >2 + 2X3)X + (\, + 2\2 + 3X3)X2. This is the zero polynomial if and only if each of the coefficients is 0; i.e. if and only if 2 0 2 1 1 2 1 2 3 *i >2 w = 0 0 0 The reader can easily verify that the above coefficient matrix is of rank 2 so that, by Theorem 3.16, a non-trivial solution exists. Hence there are scalars ^1,^2. ^3 which are not all zero such that \iP{X) + \2q{X) + \3r(X) = 0 and so the given set is linearly dependent. EXERCISES 5.11 Let 5, and S2 be non-empty subsets of a vector space such that 5, C S2. Prove that (1) if S2 is linearly independent then so is 5,; (2) if 5, is linearly dependent then so is S2. 5.12 Determine which of the following subsets of Mat 3x, IR are linearly dependent. For those that are, express one vector as a linear combination of the others: f (1) - (2) - (3) « T 0 0 1 0 0 1 1 -1 * * > "-f 2 1 1 2 -1 0 1 1 t t • "2 V 1 1 0 ' 1 1 . J J 1 ' 2 0

80 Basic Linear Algebra 5.13 Determine which of the following subsets of Mat 2x2,R are linearly dependent. For those that are, express one matrix as a linear combination of the others: (i) (2) (3) { {[■ "1 I" 1 1 "1 f 1 1)) "1 0' 0 2 "2 3" 1 2 * * "0 2' 0 2 "2 2" 1 1 }■■ 1 0 0 2 1 1 2 1 0 3 2 1 3 1 2 1 2 3 4 3 }}■

]} 5.14 Determine which of the following subsets of IR2M are linearly dependent. For those that are, express one vector as a linear combination of the others: (1) {X,3 + X2,X + 2X2}\ (2) {-2 + X.3 + X.1+X2}; (3) {-5+X + 3X2,13 + X,l+X + 2X2}. We now combine the notions of linearly independent set and spanning set to obtain the following important concept. Definition A basis of a vector space V is a linearly independent subset of V that spans V. Example 5.23 The subset {(1,0), (0,1)} is a basis of the cartesian plane IR2. Likewise, the subset {(1,0,0),(0,1,0),(0,0,1)} is a basis of IR3. More generally, {ex ej is a basis for IR" where ', = (0 0,1,0 0), the 1 being in the /-th position. These bases are called the natural (or canonical) bases. Example 5.24 In IR2 the subset {(1,1),(1,-1)} is a basis. In fact, for every (jc, y) (= IR2 we have (*,y)-x,(i,i) + Mi,-0 where X, = \{x +y) and \2 = i(x -y). Thus {(1,1), (1,-1)} spans IR2; and if <*(!, 0 + /3(1,-1) = (0,0)

5. Vector Spaces 81^ thena + /?= Oanda-/?= 0 whencea = 0 = 0,so{(l, 1), (1,-1)} is also linearly independent. EXERCISES 5.15 Prove that the monomials 1,X,... ,X" form a basis (the natural basis) forlR,,[X]. 5.16 Prove that the m x n matrices EM described by pw,-{! if;=p:;=*; [0 otherwise form a basis (the natural basis) for Matmxn IR. 5.17 Prove that the diagonal nxn matrices form a subspace of Mat nxn R and determine a basis of it. 5.18 Annxn matrix all of whose diagonal entries are the same is called a Toeplitz matrix. Prove that the set of Toeplitz matrices is a subspace of Matnxn IR and exhibit a basis for this subspace. 5.19 Let/, g, h : IR -»IR be the mappings given by f{x) = cos2 xt g{x) = sin2 x, h(x) = cos 2x. Consider the subspace of Diff(IR, IR) given by W = Span {/, g, h}. Find a basis for W. A fundamental characterisation of bases is the following. Theorem 5.6 A non-empty subset S of a vector space V is a basis of V if and only if every element of V can be expressed in a unique way as a linear combination of elements of S. Proof =» : Suppose first that 5 is a basis of V. Then V = Span S and so, by Theorem 5.3, every x € V is a linear combination of elements of 5. Now since S is linearly independent, only one such linear combination is possible for each Jt £ V; for if 51 ^ixi * 51 Mi*/ where xf G S then 5^(X, -/*i)*i * ®v whence each \ -/*,• = 0V and therefore \ = ^, for each i. «j= : Conversely, suppose that every element of V can be expressed in a unique way as a linear combination of elements of S. Then, by Theorem 5.3, Span S is the whole of V. Moreover, by the hypothesis, 0V can be expressed in only one way as a linear combination of elements of S. This can only be the linear combination in which all the scalars

are 0F. It follows, therefore, that 5 is also linearly independent. Hence 5 is a basis of V. D

82 Basic Linear Algebra Example 5.25 For i = 11... t n let Then the set {flj,..., a,,} is a basis of R" if and only if the matrix A = [fl/y]nxn is invertible (or, equivalently, has maximum rank ri). To see this, let x = (x,,..., jcJ G IR" and consider the equation x-\lai + \2a2 + ■ + \nan. By equating corresponding components, we see that this is equivalent to the system X2 = >ifl,2 + >2fl22 + • + Ka>a Xn = ^l<*ln + *2fl2n + • + *nam i.e. to the system whereA = [fl0l,,XII. From these observations we see that every x G IR" can be written uniquely as a linear combination of ax,..., an if and only if the above matrix equation has a unique solution. This is so if and only if A' is invertible, which is the case if and only if A is invertible (equivalently, A has maximum rank ri). Example 5.26 Consider the set Seq IR of all real sequences (fl«tai = oua2la3l...tan,a,,^u... of real numbers. We can make Seq IR into a real vector space in an obvious way, namely by defining an addition and a multiplication by scalars as follows: Mn>i + {bn)n>l = {an + b,,),,zi\ M<On>i'M^ntai- Define a sequence to be finite if there is some element am of the sequence such that ap = 0 for all p> m\ put another way, if there are only finitely many non-zero elements in the sequence. A'

5. Vector Spaces 83 The set Seq^- IR of all finite sequences is clearly closed under the above addition and multiplication by scalars and so is a subspace of Seq IR. Consider the (finite) sequences that are represented as follows: e, = 1,0,0,0,... e2 = 0,1,0,0,... e3 = 0,0,1,0,... *,= 0,0,0,0 0,1,0,0,... * v ' i-1 Clearly, every finite sequence can be expressed in a unique way as a linear combination of ex, e2, £3,.... Consequently, {ex, e2, eit...} forms a basis for the subspace Seq^ IR of finite sequences. Note that this basis is infinite. EXERCISES 5.20 Determine which of the following are bases of IR3: (1) {(1,1,1),(1,2,3),(2,-1,1)}; (2)((1,1,2),(1,2,5),(5,3,4)}. 5.21 Show that {(1,1,0,0),(-1,-1,1,2),(1,-1,1,3),(0,1,-1,-3)} is a basis of IR4 and express a general vector (a, b, c, d) as a linear combination of these basis elements. 5.22 Show that 1. ■f 2 1 * Y 2 3 >>, r-' 'o" 0 1) T 2 5 are bases for the same subspace of Mat 3xl IR. Our objective now is to prove that if a vector space V has a finite basis B then every basis of V is finite and has the same number of elements as B. This is a consequence of the following result. Theorem 5.7 Let V be a vector space that is spanned by the finite set G = {v,,..., v,,}. If I =

{w,,..., wm} is a linearly independent subset of V then necessarily m^n.

84 Basic Linear Algebra Proof Consider w, G /• Since G is a spanning set of V, there exist scalars \,,..., \n such that w, = V,+ •+Vn and at least one of the \; is non-zero (otherwise every \,- = 0 whence w, = 0V and this contradicts Theorem 5.4). By a suitable change of indices if necessary, we may assume without loss that X, f 0. We then have v, = \\xwx -\\x\2v2-...-\\l\nvnt which shows that V= Span G = Span {v,,v2 vj C Span{wI,v2,v3,...,vII}. It follows that Vr=Span{w1,v2,V3,...,vII}. Now u>2 can be written as a linear combination of w,, v2,..., v,, in which at least one of the coefficients of the vj is non-zero (otherwise w2 is a linear combination of w,, a contradiction). Repeating the above argument we therefore obtain V=Span{wI,M>2,V3,...,vJ. Continuing in this way, we see that if p - min{m, n} then V=Span{n>, *W,,...,vJ. Now we see that m > n is impossible; for in this case p - n and we would have V = Span{wl,...twn} whence the elements wn+lt...twm would be linear combinations of w,,..., w,, and this would contradict the fact that / is independent. Thus we conclude that m ¢. n. D Corollary 1 If V has a finite basis B then every basis of V is finite and has the same number of elements as B. Proof Suppose that B* were an infinite basis of V. Since, clearly, every subset of a linearly independent set is also linearly independent, every subset of B* is linearly independent. Now B*, being infinite, contains finite subsets that have more elements than B. There would therefore exist a finite independent subset having more elements than B. Since this contradicts Theorem 5.7, we conclude that all bases of V must be finite. Suppose now that the basis B has n elements and let B* be a basis with n* elements. By Theorem 5.7, we have n* < n. But, inverting the roles of B and B*, we deduce also that n ¢. n*. Thus n* = n and so all bases have the same number of elements. D

5. Vector Spaces 85 Corollary 2 If V has a finite basis then all linearly independent subsets of V are finite. Proof If V has a finite basis of n elements and if there existed an infinite independent subset then this would contain an independent subset of n + 1 elements, and by the above this is not possible. D Definition By a finite-dimensional vector space we shall mean a vector space V that has a finite basis. The number of elements in any basis of V is called the dimension of V and will be denoted by dim V. Example 5.27 The vector space IR" has dimension n. In fact, as we have seen before, {ex,..., en} is a basis. Example 5.28 The vector space Mat,,,XII IR is of dimension mn. To see this,

observe that if E^ is the m x n matrix that has 1 in the (/,;)-th position and 0 elsewhere then {Etj\ i = 1,...,m, ;= 1,...,n} is a basis for Matmxn IR. Example 5.29 The vector space IR,,[X] of real polynomials of degree at most n is of dimension n+1. In fact, {1, X, X2,..., X"} is a basis for this space. Example 5.30 The set V of complex matrices of the form a 0' _7 -a forms a real vector space of dimension 6. In fact, V is a subspace of the real vector space Mat2x2 C. Morover, the matrix a + ib c + id e + if -a-ib can be written as 1 0' 0 -1 + b i 0" 0 -i + c 0 1" 0 0 + d "0 i" 0 0 + e "0 0" 1 0 + / "0 0" and as the six matrices involved in this belong to V and are clearly linearly independent over F, they form a basis of the subspace that they span, which is V. a 0 7 -Of

Basic Linear Algebra Example 5.31 The set W of complex matrices of the form ' a 0 -P -a is a real vector space of dimension 4. In fact, W is a subspace of the real vector space Mat2x2 C. Morover, the matrix can be written as a 1 (0 - a P] -0 -a a + ib c + id -c + id -a - ib 3' 1 + b i 0 0" -i + c ' o r -1 0 +d o r i 0 and as the four matrices involved in this linear combination belong to W and are clearly linearly independent over IR, they form a basis of the subspace that they span, which is W. EXERCISES 5.23 Let V be a vector space of dimension 2. If {v,, v2} is a basis of V and if wltw2€;V prove that the following statements are equivalent: (1) Span{w,,M*} = V; (2) there is an invertible matrix A such that w2 = A v2 5.24 In the vector space IR4 let A = Span{(l, 2,0,1),(-1,1,1,1)}, B = Span{(0,0,1,1)}, (2,2,2,2)}. Determine A n B and compute its dimension. 5.25 If V is a vector space over C of dimension n, prove that V can be regarded as a vector space over IR of dimension In. [Hint. Consider {vj,..., vB, iv,,..., ivn} where {v,,..., v,,} is a basis over C] The reader will recall that the notion of linear independence was defined for a non-empty subset of a vector space. Now it is convenient to extend to the empty set 0 the courtesy of being linearly independent, the justification for this being that the condition for a set of elements to be linearly independent can be viewed as being satisfied 'vacuously' by 0.

5. Vector Spaces 87 Since we know that the smallest subspace of V is the singleton {0V}, and since this is clearly the smallest subspace to contain 0, we can also regard the zero subspace as being spanned by 0. These courtesies concerning 0 mean that we can usefully regard 0 as a basis for the zero subspace, which we can then say has dimension 0. This we shall do henceforth. We shall now establish some important facts concerning bases.

Theorem 5.8 Let V be a finite-dimensional vector space. If G is a finite spanning set of V and if I is a linearly independent subset of V such that I⊆G then there is a basis B of V such that I ⊂ B ⊂ G. Proof Observe first that if I also spans V then / is a basis of V and there is nothing to prove. Suppose then that V f Span /. Then we must have I ⊂ G (for otherwise I = G and is a spanning set of V). We note first that there exists g, ∈ G\I such that g, ∉ Span /; for otherwise every element of G\ belongs to Span I whence V = Span G ⊂ Span I and we have the contradiction V = Span I. We then observe that I∪ {g\} is linearly independent; otherwise we have the contradiction g, ∈ Span /. Now if / ∪ {g|} spans V then it is a basis, in which case no more proof is required since we can take B = I∪ {gt}. If / ∪ {gt} does not span V then we can repeat the above argument to produce an element g2 ∈ G\[I ∪ {gt}) with I∪ {gt, &} linearly independent. Proceeding in this way we see, since G is finite by hypothesis, that for some m the set B = I∪ {g,,g2» • • >8m} >s a basis of V with I ⊂ B ⊂ G. ☐ Corollary 1 Every linearly independent subset I of a finite-dimensional vector space V can be extended to form a basis. Proof By Corollary 2 of Theorem 5.7,1 is finite. Take G = I∪ B where B is any basis of V. Then by the above there is a basis B* with I ⊂ B* ⊂ I∪ B. ☐ Corollary 2 If V is of dimension n then every linearly independent set consisting of n elements is a basis of V. Proof This is immediate by Corollary 1 and Corollary 1 to Theorem 5.7. ☐

88 Basic Linear Algebra Corollary 3 If S is a subset of V then the following statements are equivalent: (1) S is a basis; (2) S is a maximal independent subset (in the sense that if I is an independent subset with S⊆I then S = /); (3) S is a minimal spanning set (in the sense that if G spans V and G⊆S then G = S). Proof (1) => (2): If / is independent with S⊆I then by Corollary 1 there is a basis B such that I ⊂ B. Since S is a basis, and since all bases have the same number of elements, we deduce that S - B- I. (2) => (1): By Corollary 1 there is a basis B with S ⊂ B. But B is also independent so, by (2), we have S = B and therefore S is a basis. (1) => (3): If G spans V then (recalling that 0 is independent) there is a basis B with 0 ⊂ B ⊂ C. If G⊆S then B ⊂ S and both are bases. Again since bases have the same number of elements, we deduce that B = G = S. (3) => (1): There is a basis B with 0 ⊂ B ⊂ S. But B also spans V so, by (3), we have B = S and so S is a basis. ☐ Corollary 4 If V is of dimension n then every subset containing more than n elements is linearly dependent. No subset containing fewer than n elements can span V. Proof This is immediate

from Corollary 3. D As for subspaces of finite-dimensional vector spaces, we have the following consequence. Theorem 5.9 Let V be a finite-dimensional vector space. If W is a subspace of V then W is also of finite dimension, and dimW< dirnV. Moreover, we have dimW = dimV <=> W=V. Proof Suppose that V is of dimension n. If / is a linearly independent subset of W then, by Theorem 5.7,1 has at most n elements. A maximal such subset B is then, by Corollary 3 of Theorem 5.8, a basis of W. Hence W is also of finite dimension, and dim W < dim V.

5. Vector Spaces 89 Finally, if dim W = dim V = n then B is a linearly independent subset of V having n elements whence, by Corollary 2 of Theorem 5.8, B is a basis of V. Hence W = SpanB=V. U Example 5.32 Consider the real vector space IR2. This is of dimension 2 and so if W is a subspace of IR2 then by Theorem 5.9 the dimension of W is either 0,1, or 2. If dim W = 0 then we have W = {(0,0)}. If dim W = 2 then, by Theorem 5.9, we have W = IR2. If dim W = 1 then W has a basis of a single non-zero element (jc, y), so that W= {\(x,y); \£lR} = {(\x,\y)\ X G IR}, which is none other than the line passing through the origin (0,0) and the point (a ,)). Example 5.33 Arguing in a similar way to the above, we can show that the subspaces of IR3, corresponding to the dimensions 0,1,2,3, are: the zero subspace {(0,0,0)}; any line through the origin; any plane through the origin; IR3 itself. Example 5.34 If V is a vector space with dim V = 10 and X, Y are subspaces of V with dim X - 8 and dim Y = 9 then the smallest possible value of dim (X n Y) is 7. To see this, begin with a basis {v,,..., vp} of X n Y and, using Corollary 1 of Theorem 5.8, extend this on the one hand to a basis {vi,--,VVi»--»v8} of X, and on the other hand to a basis (vi vptw^lt...tw9} of Y. Observe that none of w^,,..., n>9 belongs to X (for otherwise it belongs to X n K, a contradiction), and so {^1,...,^^,,...,^,^,,...,^9} is linearly independent (otherwise one of the vectors w^,,..., w9 would belong to Span {v,,..., v8} = X, a contradiction). Since dim V = 10, this set contains at most 10 elements. For this we must have p ^ 7. To see that this lower bound of 7 is attainable, consider V = IR10 and take for X the subspace consisting of those 10-tuples whose first and third components are 0, and for Y the subspace consisting of those 10-tuples whose second component is 0.

90 Basic Linear Algebra EXERCISES 5.26 Consider the subset of IR4 given by X = {(2,2,1,3),(7,5,5,5),(3,2,2,1),(2,1,2,1)}. Find a basis of Span X and

extend this to a basis of IR4. 5.27 Find a basis for Mat 3x, IR that contains both the vectors "l" 0 2 • "0" 2 4 5.28 Find a basis of IR3[XJ containing the polynomials 1+X+X2 andX -X3. SUPPLEMENTARY EXERCISES 5.29 For each of the following statements give a proof if it is true and a counter-example if it is false: (1) If V is a vector space over a field F then a non-empty subset W of V is a subspace of V if and only if (Vjc.y G W)(VX,/j € f) ** + py G W. (2) The subspace {(jc, jc, jc) ; jc € IR} of IR3 is of dimension 3. (3) Every spanning set contains a basis. (4) The subspace of IR3 spanned by {(1,2,1), (2,2,1)} is (a) {(a + 2b, 2a+2bta + b); a,b G IR}. (b) {{x + y,2y,y)\ Jf.yGlR}; (5) If P, Q are subspaces of a finite dimensional vector space then (a) P C Q implies dim P < dim Q\ (b) dim P ^ dim Q implies P C Q. (6) If {xty, z} is a basis of IR3 and w is a non-zero vector in IR3 then {w + x,y, z} is also a basis of IR3. 5.30 Determine whether or not the following subsets of IR4 are subspaces: (1) {{atbtctd)\ a + b = c + d}-t (2) {(a,b,c,d)- a + b=\}; (3) {(fl,M, <0;fl2+ ^= 0)-. (4) {(atbtctd)-a2 + b2=l}. 5 31 Determine whether or not the following subsets of IR4 are subspaces: (1) {(* +2yt0t2x -y,y); J.yGlR}; (2) {{x + 2ytxt2x-y,y)'t *,yGlR}.

5. Vector Spaces 91 5.32 Prove that the subset {(3-,, 2 + 2/,4),(2,2 + 41, 3),(1 -,', -2i, -1)} is a basis of the complex vector space <C3. Express each of the vectors (1,0,0), (0,1,0), (0,0,1) as a linear combination of these basis vectors. 5.33 Find a basis for the solution space of the homogeneous system [12 2 1 -11 0 2 2-1-2 2 6 2 \ -4 .14 0 0 -3. "jc* y z t = r°i 0 0 LoJ 5.34 Let V be a finite-dimensional vector space. If AtB are subspaces of V, prove that so also is the set A + B - {a + b ; a e A, b G B). Prove further that if C is any subspace of V such that ACC and BCC then A + B C C (in other words, A + B is the smallest subspace of V that contains both A and B). If L, A/, N are subspaces of V prove that Ln[M + (Lr)N)} = (LnM) + (LnN). Give an example to show that in general Ln(M + N)?(LnM) + {Lr\N). 5.35 Let n be a positive integer and let En be the set of mappings/: IR —»IR that are given by a prescription of the form n /(jc) = a0 + Y^akcos kx + bk sin kx) jt= i where ak,bk£ IR for each k. Prove that En is a subspace of Map(IR, IR). If/ € £,, is the zero mapping, prove that all the coefficients ak, bk must beO. [Hint. Proceed by induction. For this, find a prescription for D2/+ n2f] Deduce that the In + 1 functions jr i—► 1, x»->cosA:jc, x i-* sin kx (k - 1,...,n) form a basis for E,,.

92 Basic Linear Algebra 5.36 Let a,@ G IR with a f @ and let r,s be fixed

positive integers. Show that the set of rational functions/ : IR —»IR given by a prescription of the form where each a{ G IR, is a subspace of Map(IR, IR) of dimension r + s. [Hint. Show that the functions X *->f;(x) = -. 7-. —— (x-aY(x-0Y for i = 0,..., r + s -1 constitute a basis.] Show also that if g- and hj are given by ftM = (jr-a)\ hj(x) = (x-py then is also a basis. [Hint. It suffices to prove that B is independent.] 5.37 For each positive integer k \elfk: IR —»IR be given by fk{x} = exp rkx where each rk e IR. Prove that (/,)1^,-^,, is linearly independent if and only if r,,...., rn are distinct. 5.38 Let P0(X),P,(X),...,Pn(X) be polynomials in IR,,[X] such that, for each 1, the degree of P,(X) is 1. Prove that {P0(X)tPl(X)t...tPn(X)} is a basis of IR,,[X]. 5.39 A net over the closed interval [0,1] of IR is a finite sequence (fl,)o^, ^n+i such that 0 = fl0<fli < --■ <fl,, < fl,,*i = 1. A step function on the semi-open interval [0,1[is a map/: [0,1[—»IR for which there is a net (0,)0^,,+1 over (0» 1) ar,d a finite sequence (bi)o<i<n such that (VJtGk,flI+1[) f(x) = bt. Sketch a picture of a step function. Show that the set E of step functions on [0,1 [is a vector space.

5. Vector Spaces 93 Show also that a basis of £ is the set of functions ek : [0, l[-» IR given by eM {0 if0O<Jt; 1 if*^jc<l. A piecewise linear function on [0,1[is a map/: [0,1[-* IR for which there is a net (0,)0^,,+1 and finite sequences (fc()0<^M, (0<k,$m sucn that (V*ek,fl,+1[) f{x} = bix + ci. Sketch a picture of a piecewise linear function. Show that the set F of piecewise linear functions on [0,1[is a vector space. If G is the subset of F consisting of those piecewise linear functions g that are continuous with g(0) = 0, show that G is a subspace of F. Show also that a basis of G is the set of functions gk : [0,1[—▶ IR given by {0 if0<jc<Jt; x-k lfit^jc < 1. Show finally that every/ G F can be written uniquely in the form g + e where g € G and e £E. ASSIGNMENT EXERCISES (1) By a magic matrix we mean a real square matrix in which all row sums, all column sums, and both diagonal sums are equal to some value a. If M = [wjfyk^ is magic, prove that 0 = 3m22. Deduce that, given a,b,c € IR there is a unique 3x3 magic matrix M(a,b, c) such that m22 - fl, mn = a + bt m3l = a + c. Show that {M(a,b, c); a, b, c G IR} is a subspace of Mat 3x3 IR and that 5={M(1,0,0),M(0,1,0),M(0,0,1)} is a basis of this subspace. By a pseudomagic matrix we mean a real square matrix in which all row sums and all column sums are equal to some value 0 (no requirement being made about the diagonal sums). Prove that a matrix A is pseudomagic if and only if AJ = JA = oJ where J is the matrix all of whose elements are 1.

94 Basic Linear Algebra Deduce that if A, B are pseudomagic matrices of the same size then so also are xA + yB and AB. In contrast, provide an example to show that the product of two magic matrices need not be magic. (2) If A is a non-empty subset of IR show that Map (A, IR) is a subspace of the vector space Map(IR, IR). If, for each a G A, the mapping/,: A -»IR is given by , v f 1 if* = a; /aW"\0 ifjc^fl, prove that {fg ; a G A} is a basis of Map (A, IR). (3) Let V be the real vector space of all polynomial functions / : IR —» IR of degree at most 2, i.e. functions of the form f(x) = a0 + axx + a2x2 where a0l alta2 € IR- If r is a fixed real number define/,,/2,/3 : IR —► IR by the prescriptions /iW=l. /2(*) = * + rt /sM-fr + r)2. Prove that B = {/, /2,/3} is a basis for V. If/: IR -► IR is given by f{x) = a0 + axx + a2x2, express/ as a linear combination of the elements in B. (4) Let F be the vector space of infinitely differentiable functions / : C —» C and let Pn be the subspace consisting of the complex polynomial functions of degree at most n. For each a € C define Pnja = {eazp(z)\p(z)ePHY Show that Pna is a subspace of F and that ^e1*^;*-0,...,11} is a basis of Pnor. If D denotes the differentiation mapping, prove that £^-^7) = ^(0- <* >d)n/.

6 Linear Mappings In the study of any algebraic structure there are two concepts that are of paramount importance. The first is that of a substructure (i.e. a subset with the same type of structure), and the second is that of a morphism (i.e. a mapping from one structure to another of the same kind that is 'structure-preserving'). So far, we have encountered the notion of a substructure for a vector space; this is called a sub space. In this chapter we shall consider the notion of a morphism between vector spaces, i.e. a mapping from one vector space to another that is 'structure- preserving* in the following sense. Definition If V and W are vector spaces over the same field F then by a linear mapping (or linear transformation) from V to W we shall mean a mapping /: V —» W such that (1)(Vjc,vGV) f(x + y)=f(x)+f(y); (2)(VjcGV)(VXgF) f(\x) = \f(x). • If/: V —» W is linear then V is sometimes called the departure space and W the arrival space of/. Example 6.1 The mapping/: IR2 -»IR3 given by f(a,b) = (a + bta-b,b) is linear. In fact, for all (a,b) and (a'tb') in IR2 we have f((a,b) + W,b'))=f(a + a'tb + b') = {a + a' + b + b'ta + a'-b-b'tb + b') = (fl + 6,fl-6,6) + (fl' + 6',fl'-6',6') = /(*,*)+/(*',*')

96 Basic Linear Algebra and, for all (a, b) G IR2 and all X G IR, f(\(atb)) = f(^,\b) = (\a + \bt\a-M>,\b) = \(a + b,a-b,b) = >/(*,*>)• Example 6.2 The

mapping pr, : IR" -»IR described by pr((jr,,..., *,,) = */ (i.e. the mapping that picks out the i-th coordinate) is called the i-th projection of IR" onto IR. It is readily seen that (1) and (2) above are satisfied, so that pr, is linear. Example 6.3 Consider the differentiation map D: IR,,[X] -»IR,,[X] given by D(a0 + axX + • • • + anXn) = ax + 2a2X + • • • + nanXn~l. This mapping is linear; for if p(X) and q(X) are polynomials then we know from analysis that D(j>(X) + q(X)) = Dp(X) + Dq(X) and that, for every scalar X, D(\p(X)) = \ Dp{X}. EXERCISES 6.1 Decide which of the following mappings/: IR3 -»IR3 are linear: (l)f(x,y,z)=(y,z,0); (2) f{xtytz) = (zt-ytx)\ (3)/(*,>,*)= (M,-z,0); (4) f(x,y,z) = {x-\txty)\ (5)/(*,y,z) = (* +)\z,0); (6)/(jc,y,z) = (2jc,v-2,4>). 6.2 Let B G MatnxlI IR be fixed and non-zero. Which of the following mappings TB : Matnxn IR -» MatnxlI IR are linear? (1) TB(X) = XB-BX- (2) TB{X} = XB2 + BX; (3) TB(X) = XB1-BX2. 6.3 Which of the following mappings are linear? (D/ (2)/ (3)/ R,,[X] -> IR3[X] given by /fc>(X)) = p(0)X2 + Dp(0)X3; IR,,M - IRn+iM given by f{p(X)) = p(0) + Xp(X); «,,[*] - K,,+iM given by/fc>(X)) = 1 + Xp(X).

6. Linear Mappings 97 6.4 Let A be a given real m x n matrix. Prove that the mapping /A:MatnxlIR->MatmxlIR described by fA{\) - Ax is linear. 6.5 Let I: IR,,[X] -»IR be the integration map defined by i(pM)= fpW- Jo Prove that I is linear. 6.6 For anmx n matrix A let/(/1) be the Hermite form of A if A f 0, and let/(0) = 0. Is/linear? The following result contains two important properties of linear mappings that will be used constantly in what follows. Theorem 6.1 If the mapping/: V —♦ W is linear then (1)/(0,) = 0^ (2) (VjcGV) /(-*)=-/(*). Proof (1) We have/(0v) =/(0f0v) = 0f/(0v) = 0W. (2) Using (1) we have, for every jc € Vt f(x)+f(-x)=f[x + (-x))=f(Qv) = Ow from which the result follows on adding -/(jc) to each side. □ EXERCISES 6.7 Let B G Mat,,XII IR be fixed and non-zero. Prove that the mapping TB : MatnxBIR-MatBXIIIR given by TB{A) = (A + B)2 -(A + 2B)(A - 3B) is linear if and only if B2 = 0. We shall now consider some important subsets that are associated with linear mappings. For this purpose we introduce the following notation. If/ : V ~* W is linear then for every subset X of V we define /~*(X) to be the subset of W given by r(X) = {f(x);xeX}; and for every subset Y of W we define f*~{Y) to be the subset of V given by r(Y) = {x£V-f(x)eY}. We often call/~*(X) the direct image of X under/, and/*~(K) the inverse image of Y under/.

98 Basic Linear Algebra • The reader should be warned that this is not

'standard* notation, in the sense that most authors write/(X) for/-(X), and fl(Y) for f*~{Y). We introduce this notation in order to reserve the notation f~l as the standard notation for the inverse of a bijection. One advantage that this non-standard notation has to offer is that it gives a visually appealing reminder that/"" sends subsets of V to subsets of W, and /*" lifts back subsets of W to subsets of V. EXERCISES 6.8 Consider the differentiation map D: IR,,[X] -»IR,,[X]. Describe the sets D"*(IR,,[Xj)andD^(0). 6.9 Prove that/- = /-* o/- o/- and that/- =f~of-+of-. The mappings/-* and/4- are each inclusion-preserving in the sense that (a)X,CX2 =*/-(X,)C/-(X2). For, if v G/^(X,) then y =/(*,) where jc, G X, C X2. (b) YXCY2 * r{Yx)cr(Y2). For, if x $r(Yx) then/(jc) eK.C Y2. Moreover, each of these mappings carries subspaces to subspaces: Theorem 6.2 Letf : V —»IV te /meor If X is a subspace of V thenf~*(X) is a subspace of W\ and if Y is a subspace of W thenf*~(Y) is a subspace of V Proof Observe first that if X is a subspace of V then we have 0V G X and therefore 0W = /(Ov)er(X).1Tius/-(X)^0. If now yXty2 G /-(X) then yx = f(xx) and y2 = /(jc2) for some jc,,jc2 G X. Consequently, since X is a subspace of V, >. + >2 =/(*,) +f{x2) =f(xx + *2) G/^(X); and, for every scalar \t av, = v(*,)=/(xjc.) enn Thus/~*(X) is a subspace of W. Suppose now that Y is a subspace of W. Observe that f(0v) = 0W g Y gives Ov G /^(0. and therefore/^(K) ^ 0. If now jc, , jc2 G f~(Y) then/(jc,),/(jc2) G Y and therefore

6. Linear Mappings 99 whence jc, + jc2 e f*~(Y)\ and, for every scalar X, f(\xl) = \f(xy)eY whence X*, G /-(^). Thus f^(Y) is a subspace of V. □ EXERCISES 6.10 Show that the subset X of polynomials in IF^nM a" of whose odd coefficients fl2i+i are zero forms a subspace of IF^nM- Describe D~*{X) and D^(X). 6.11 Show that the mapping/: IR2 -»IR2 given by f(x,y) = (x + ytx-y) is linear. For each subspace X of IR2 describe f~*(X) and/*~(X). 6.12 Let/: V —* W be linear. If X is a subspace of V and Y is a subspace of W, prove that r[xnr(Y))=r(x)nY. Deduce that Of particular importance relative to any linear mapping /: V —» IV are the biggest possible direct image and the smallest possible inverse image. The former is/~*(V); it is called the image (or range) of/ and is denoted by Im/. The latter is/*~({0w}); it is called the kernel (or null-space) of/ and is denoted byKer/. Pictorially, these sets can be depicted as follows: Ker/ W

100 Basic Linear Algebra Example 6.4 Consider the i-th projection pr, : IR" -» R. Recall that pr,(jc,,..., jcJ = jc,. The image of pr, is therefore the whole of IR;

and the kernel of pr, is the set of n-tuples whose i-th component is 0. Example 6.5 Consider the differentiation map D : IR,,[X] -» IR,,[X]. Its image is the set of all polynomials of degree at most n -1; in other words, it is IR,,_|[X]. The kernel of D is the set of polynomials whose derivative is zero; in other words it is IR. Example 6.6 If A is a given real n x n matrix, consider the linear mapping /A:MatnxlIR-»MatnxlIR described by fA[x] = Ax. The image o(fA consists of all n x 1 column matrices y = y\ for which there exists x = *i such that Ax = y; i.e. the set of all y such that there exist jc, ,..., jc,, with y=*,a,+ • +xnan. In other words, Im fA is the subspace of Mat,,yl IR that is spanned by the columns of A. As for the kernel of/A, this is the subspace of Mat nx, IR consisting of the column matrices x such that Ax =■ 0; i.e. the solution space of the system Ax = 0. Example 6.7 Consider the subspace of Map(IR, IR) that is given by V= Span {sin, cos}, i.e. the set of all real functions/ given by a prescription of the form /(jc) = a sin jc + 6 cos jc. Let I: V —»IR be given by i(/)= T/. Jo Using basic properties of integrals, we see that I is linear. Now if/(jc) = a sin jc + bcos x then/ G Ker I if and only if I (a sin jc + bcos x)dx = 0. Jo This is the case if and only if a = 0. Consequently, we see that Ker I = Span {cos }

6. Linear Mappings 101 Example 6.8 Consider the mapping/: IR4 -> F3 given by f(a,btctd) = (a + b,b -cya + d). Since (a + btb -cta + d) = a(l,0,1) + b{\, 1,0) + c(0, -1,0) + </(0,0,l) we see that Im/=Span{(l,0,1),(1,1,0),(0,-1,0),(0,0,1)} To find a basis for Im/, proceed as follows. Observe that Im/ is the subspace spanned by the rows of the matrix "I 0 r 110 0-10' 0 0 1. The Hermite form of A is "1 0 0" 0 1 0 0 0 1' 0 0 0. Since the rows of this matrix span the same subspace, and since they are linearly independent, we deduce that a basis for Im / is {(1,0,1),(1,1,0),(0,-1,0)}. Thuslm/=IR3. EXERCISES 6.13 Find Im/ and Ker/ when/: IR3 -»IR3 is given by f{atbtc)= {a + btb + c,a + c). 6.14 If/: F2 -> IR2 is given by /(a, b) = (6,0), prove that Im /= Ker/. 6.15 Give an example of a linear mapping for which Im/ C Ker/; and an example where Ker/ C Im/. 6.16 Let/: IR5 — IR4 be given by f(a,b,ctdte) = (a-c + 3d-e,a + 2d-et2a-c + 5d-et-c + d). Find bases for Im / and Ker/. 6.17 Let/: IR2[XJ -> IR3[XJ be given by f(p(X)) = X2Dp(X). Prove that/ is linear and determine bases for Im / and Ker/.

102 Basic Linear Algebra 6.18 Consider the subspace A of IR4 given by /1 = {(jc,0,z,0); jc.zG R}. Determine linear mappings/, g : IR4 -»IR4 such that (1)

Im/=/l; (2) Kerg = A. Definition A linear mapping/ : V —» W is said to be surjective if Im / = W (in other words, if every element of W is the image under/ of some element of V)\ and injective if /(jc) f f(y) whenever jc f y (in other words, if/ carries distinct elements to distinct elements). We say that/ is bijective if it is both injective and surjective. Example 6.9 The i-th projection pr,: IR" —»IR is surjective but not injective. Example 6.10 The linear mapping/: IR2 -»IR3 given by /(*,>) = (y,0,jr) is injective but not surjective. Example 6.11 The differentiation map D : IR,,[X] -»IR,,[X] is neither injective nor surjective From the above definition, a linear mapping/ : V —» W is surjective if Im/ is as large as it can be, namely W. Dually, we can show that/ : V —» W is injective if Ker/ is as small as it can be, namely {0W}: Theorem 6.3 If f: V —» W is linear then the following statements are equivalent: (1) f is injective; (2)Ker/={0}. Proof (1) =» (2): Suppose that/ is injective. Then/ is such that xfy^Mtfb) or, equivalent^, f(x)=f(y) => * = y- Suppose now that jc € Ker/. Then we have f(x) = 0w=f(0v)

ungs whence we see that x = Ov and consequently Ker/ = {0}. (2) => (1): Suppose that Ker/= {0} and let/(jc) =/(>). Then /(jc -y)=/[x + (-y)] = f(x) +/(-y) = /(jc) -/(>) = 0,, so that jc - y G Ker / = {0V} and hence jc = y, i.e. / is injective. D Example 6.12 The linear mapping/: F3 -» Fl3 given by f(x>y>z) = (x + Z* x + y + 2zt2x + y + 3z) is neither surjective nor injective. In fact, we have that (a, bt c) € Im / if and only if the system of equations jc + z = a x + y + 2z = b 2x + y + 3z = c is consistent. The augmented matrix of the system is 'l 0 1 a' 112 6 2 13c and this has Hermite form 'l 0 1 a 0 1 1 b-a 0 0 0 c-b-a

104 Basic Linear Algebra 6.20 Prove that if the linear mapping/: V -» W is injective and {v,,..., vn} is a linearly independent subset of V then {/(v,),...,/(v,,)} is a linearly independent subset of W. In the case of finite-dimensional vector spaces there is an important connection between the dimensions of the subspaces Im / and Ker/. Theorem 6.4 [Dimension Theorem] Let V and W be vector spaces of finite dimension over afield F.Iff:V->Wis linear then dim V = dim Im / + dim Ker/. Proof Let {w,,..., wm} be a basis of Im /, and let {v,,..., vn} be a basis of Ker/. Since each w, G Im /, we can choose v\,..., v* G V such that/(v*) = w, for / = 1,..., m. We shall show that M Oi »>J is a basis of V, whence the result follows. Suppose that x G V. Since/(jc) G Im / there exist X,,..., \n G F such that /w=t *.". = E vw)=£/(vr)=/(e vr). i=l i=l i=l »=1 It follows that *-f>fv?GKer/ i=i and so there exist n,,...,

mn G F such that m n *-EVf = EM;Vr i=i ;=i Thus every x G V is a linear combination of vf,..., v*n, v,,..., vn and so V=Span{W v;,v, vj. Suppose now that (1) EV: + Ew = o. i=i y=i Then we have EVf=-X>;VyGKer/ i=l y=l and consequently EVi = EVW)=/(EV?) = o /=i i=i /=i

6. Linear Mappings 105 whence X, = . ■. = Xm = 0 since {w,,..., wm} is a basis of Im /. It now follows n from (1) that £ pjvj - 0 whence /z, = ■ • • = (jl,, = 0 since {v,,..., vn} is a basis of Ker/. Thus we see that the spanning set W Oi VJ is also linearly independent and is therefore a basis of V. D Definition If/ is a linear mapping then dim Im / is called the rank of/, and dim Ker / is called the nullity of/. With this terminology, the dimension theorem above can be stated in the form: rank + nullity - dimension of departure space. Example 6.13 Consider pr, :R3-»fi given by pr,(x,y,z) = x. We have Im pr, = IR which is of dimension 1 since {1} is a basis of the real vector space IR; so pr, is of rank 1. Also, Ker pr, is the y, z-plane which is of dimension 2. Thus pr, is of nullity 2. EXERCISES 6.21 Let V be a vector space of dimension n ^ 1. If/: V —» V is linear, prove that the following statements are equivalent: (1) Im/=Ker/; (2) / f 0, fl = 0, n is even, and the rank of/ is jn. 6.22 Give an example of a vector space V and a linear mapping / : V —» V with the property that not every element of V can be written as the sum of an element of Im / and an element of Ker/. 6.23 In the vector space of real continuous functions let W= Span {sin, cos}. Determine the nullity of i?: IV -» F when t? is given by 0M(/)= /"/; Jo (2)m= /"/; Jo (3) iH/) = D/(0).

106 Basic Linear Algebra 6.24 In the vector space of real continuous functions let W= Span {/, gth} where M = e\ g(x) = e-\ h(x) = x. Let T: W -» W be the linear mapping given by f(t?)=D2i9-i9. Determine the rank and nullity of T. As an application of the dimension theorem, we now establish another result that is somewhat surprising. Theorem 6.5 Let V and W be vector spaces each of dimension n over a field F. If f : V —» W is linear then the following statements are equivalent. (1) f is injective; (2) / is surjective; (3) / is bijective; (4) / carries bases to bases, in the sense that if{vlt...tvn}isa basis of V then {/M f(vn)}isabasisofW. Proof (1) =» (3): Suppose that/ is injective. Then Ker/ = {0} and so dim Ker/ = 0. By Theorem 6.4, it follows that dim Im / = n = dim V - dim W. It now follows by Theorem 5.9 that Im / = W and so/ is also surjective, and hence is bijective. (2) =}► (3): Suppose that / is surjective. Then Im / = W and so, by Theorem 6.4, dim Im / = dim W= n - dim

$V = \dim \operatorname{Im} f + \dim \operatorname{Ker} f$ whence $\dim \operatorname{Ker} f = 0$. Thus $\operatorname{Ker} f = \{0\}$ and so, by Theorem 6.3, f is also injective, and hence is bijective. (3) => (1) and (3) => (2) are clear. (1) =>. (4): Suppose that f is injective. If $\{v_1,..., v_n\}$ is a basis of V then the elements $f(v_1),...,f(v_j)$ are distinct. If now $\pounds X_i f(v_i) = 0$ then $f(\pounds X_i v_j) = 0$ i=i i=i n and so, since $\operatorname{Ker} f = \{0\}$, we have $\pounds^i v_i = 0$ and hence $X_i = \bullet \bullet \bullet = \backslash n = 0$. Thus $\{f(v_1),...,f(v_n)\}$ is linearly independent. That it is now a basis follows from Corollary 2 of Theorem 5.8.

6. Linear Mappings 107 (4) => (2): Since every linear combination of $f(v_1),...,f(v_n)$ belongs to $\operatorname{Im} f$, it is clear from (4) that $\operatorname{Im} f = W$ and so f is surjective. D Definition A bijective linear mapping is called a linear isomorphism, or simply an isomorphism. We say that vector spaces V, W are isomorphic, and write $V \sim W$, if there is an isomorphism $f: V \longrightarrow W$. Example 6.14 Le$M = \{(x.y.O); x,y \in IR\}$ be the x,y-plane in F^3, and let $B = \{\{x,0,z\}; x,z \in IR\}$ be the x, z-plane. Consider the mapping $f: A \longrightarrow B$ given by $f(x,y,0)=(x,0,y)$. Clearly, f is linear and bijective. Thus f is an isomorphism and so $A \sim B$. This example is a particular case of the following general situation. Theorem 6.6 Let V be a vector space of dimension $n^{\wedge}$ over a field F. Then V is isomorphic to the vector space F^n. Proof Let $\{v_1,..., v_n\}$ be a basis of V. Consider the mapping $f : V \longrightarrow F^n$ given by the prescription $f(\pounds V_i) = (x_1,...f$ U n Since for every $x \in V$ there are unique scalars $X_1,..., \backslash n$ such that $x = \pounds X_i v_i$, it is clear that f is a bijection. It is clear that f is linear. Hence f is an isomorphism. D Corollary If V and W are vector spaces of the same dimension n over F then V and W are isomorphic. Proof There are isomorphisms $f_V : V \longrightarrow F^n$ and $f_W : W \longrightarrow F^n$. Since the inverse of an isomorphism is clearly also an isomorphism, so then is the composite mapping $f \longrightarrow 0 f_V : V \longrightarrow W$. D EXERCISES 6.25 Exhibit an isomorphism from $f_n[X]$ to F^{n+1}.

108 Basic Linear Algebra 6.26 If $f: F \to F$ is linear and $i?: F^2 \longrightarrow F^2$ is defined by $*(*,•>)= (*,y-/M)$, prove that $i9$ is an isomorphism. Our next objective is to prove that a linear mapping enjoys the property of being completely and uniquely determined by its action on a basis. This is a consequence of the following result. Theorem 6.7 Let V and W be vector spaces over a field F. If $\{v_1,..., v_n\}$ is a basis of V and $w_1,..., w_n$ are elements of W (not necessarily distinct) then there is a unique linear mapping $f: V \longrightarrow W$ such that $(i =1,...,n)$ $f(v_i) = w_i$. Proof n Since every element of V can be expressed uniquely in the form $J^{\wedge} X_i \cdot v_i$, we can define a mapping $f: V \longrightarrow W$ by the prescription $f(\pounds V_i) =$

£>.".. 1=1 /=1 i.e. taking x as a linear combination of the basis elements, define/(x) to be the same linear combination of the elements w,,..., wn. It is readily verified that/ is linear. Moreover, for each i, we have /(vt) = /(£ Sift) = E SijWj = w,. >=i ;=i As for the uniqueness, suppose that g : V -» W is also linear and such that n g(vt) = w, for each i. Given x G Vt say x = £ \tvh we have i=l 8M = g(t^) = tKgM = t^wi=M i=i i=i i=i whence g = /. □ Corollary 1 >4 //near mapping is completely and uniquely determined by its action on a basis. Proof If / : V -> W is linear and B = {v,,..., v,,} is a basis of V let w, = /(v() for each i. Then by the above / is the only linear mapping that sends v^ to wt. Moreover, knowing the action of/ on the basis B, we can compute f(x) for every x\ for jc = EV, gives/(*) = £x/(vf). □ i=I i=I

6. Linear Mappings 109 Corollary 2 Two linear mappings f,g : V —* W are equal if and only if they agree on any basis ofV. Proof If/(vi) = #(vi) f°r every basis element v, then by the above uniqueness we have that /=* a Example 6.15 Consider the basis {(1,1,0), (1,0,1), (0,1,1)} of IR3. If /: IR3 -»IR2 is linear and such that /(1,1,0)=(1,2), /(1,0,1) = (0,0), /(0,1,1)=(2,1), then we can determine/completely. In fact, we have (1,0,0)=1(1,1,0)+1(1,0,1)-1(0,1,1) and therefore /(1,0,0) =1/(1,1,0) +1/(1,0,1)-1/(0,1,1) . = 1(1,2) + 1(0,0)-1(2,1) = (-1 1) Likewise, give (0,1,0)= 1(1,1,0)-1(1,0,1) + 1(0,1,1), (0,0,1) = -1(1,1,0)+1(1,0,1) + 1(0,1,1) /(0,1,0)= 1(1,2)+1(2,1)=(|,|) /(0,0,1)=-1(1,2) + 1(2,1)=(1,-1). Consequently,/ is given by f(x,y,z)= /Ml,0,0) + y(0,1,0) + 2(0,0,1)] = */(!, 0,0) + y/(0,1,0) + 2/(0,0,1) = *(-!,i) + y(M) + z(i,-l) = (!(-* +3y + z), i(x + 3y-z)). Note that, alternatively, we could first have expressed (x, y, z) as a linear combination of the given basis elements by solving an appropriate system of equations, then using the given data. Finally, let us note that Theorem 6.5 is not true for vector spaces of infinite dimension:

110 Basic Linear Algebra Example 6.16 Let V = Seq^ IR be the infinite-dimensional vector space of finite sequences of real numbers described in Example 5.26. Since every element of V is a (finite) linear combination of basis elements, we can define a linear mapping/: V —» V by speci- fying/(e,) for the basis elements elte2,e3t... and extending to all of V by linearity. Consider then the definition {0 if/ is odd; .... e\i if i is even. Since/(e,) = 0 = /(e3) we see that/ is not injective. But, given any basis element en we have en = f{e2,,) G Im /, so the subspace spanned by these elements (namely, the whole of V) is contained in Im /. Hence Im / = V and so/ is surjective. If we define g :

$V \longrightarrow V$ by specifying $g(ei) - e2i$ for every i then we obtain an injective linear mapping that is not surjective. EXERCISES 6.27 Show that {(1,1,1), (1,2,3), (1,1,2)} is a basis of IR3. If/: IR3 -> IR3 is linear and such that /(1,1,1)= (1,1,1),/(1,2,3) = (-1,-2,-3),/(1,1,2)=(2,2,4), determine f(x.y.z) for all (x.y.z) € IR3- 6.28 If/: IR2[XJ - R3[XJ is linear and such that/(1) = 1,/(X) = X2 and /(X2) = X + X3, determine f(a + bX + cX2). SUPPLEMENTARY EXERCISES 6.29 Let/: C -» Cbegivenby/(x+iy) = jt-iy. Prove that ifC is considered as a real vector space then / is linear, whereas if C is considered as a complex vector space/ is not linear. 6.30 Let V be a vector space of dimension 3 with {v,, v2, v3} a basis. Let W be a vector space of dimension 2 with {w,, w2) a basis. Let/: V ->W be defined by /(Vi + *2V2 + V3) = (*i + MK + (*2 + ^3)^2- Determine the values of /j for which/ is linear. For these values of /j, determine a basis of Ker/. 6.31 For the linear mapping/: IR3 -»IR3 given by f{x,y,z}={x + y,0, y-z) determine Im /, Ker/, and a basis of each. If A is the subspace {(x.y.z) G IR3 ; * = >-}, determine f~(A) and a basis of it.

6. Linear Mappings 111 6.32 Let/: IR3 -> IR3[X] be the linear mapping such that /(1,0,0) = 2X + X3; /(0,1,0) = -2X + X2; /(0,0,1) = X2+X\ Determine (1)/(x,y,z)forall(x,y,z)GlR3; (2) Im / and a basis of it; (3) Ker/ and a basis of it. Extend the basis of (3) to a basis of IR3. 6.33 Show that {a, 6, c] is a basis of IR3 where fl = (-1,1,1), 6 = (1,-1,1), ¢ = (1,1,-1). Let/: IR3 —» R4 be the linear mapping such that /W = (l.O.l.X), /(6) = (0,1,-1,0), f(c) = (1,-1,X.-l). (1) Determine f{x,y,z) for all (jt.y.z) € R3. (2) For which values of X is/ injective? (3) Consider the subspace W of IR4 given by W = Span {/(0),/(6)}. Determine dim W when X = -1. (4) With X = 2 determine/(1,1,0) and/-{(1,1,0,0)}. 6.34 A non-empty subset 5 of a vector space is convex if tx + (1 - t)y € S forallx.y eSandallr €[0,1]. Prove that if 5 is a convex subset of IR" and/: IR" -♦ IR" is linear then /~*(5) is also convex. 6.35 A diagram of finite-dimensional vector spaces and linear mappings of the form vx > v2 * v3 ► * vn+l is called an exact sequence if (1)/, is injective; (2) /,, is surjective; (3) (/=1 n-I) Im/. = Ker/;+1. Prove that, for such an exact sequence, £(-l)'diniV;. = 0. i=l 6.36 Determine the rank and nullity of the linear mapping /:Mat3xlIR->Mat3XlIR

112 Basic Linear Algebra given by/(x) = Ax where "l 0 2 1 0 1 2" 2 -2 A = 6.37 If /: V -> W and g : W -> X are linear, prove that (1) Im(go/)CImg; (2) Ker/CKer(goy); (3) rank / + rank g - dim W ^ rank gof^ min {rank /, rank g}.

6.38 Determine the rank and nullity of f: IR3[X] -> IR3[X] given by $f(p(X)) = (X-1)D^3 p(X)$. 6.39 Given $x = \{x_1, x_2, x_3\}$ and $y = (y_1, y_2, y_3)$ 'n ^{3}» define the wedge product of x,y by $x \wedge y = (x_2 y_3 - x_3 y_2, *_3 y, -*, y_3, x, y_2 - j^\wedge i)$- Define f,: IR3 —► IR3 by $f_y\{x\} = x \wedge y$. Show that f, is linear. If $>• \wedge 0$, prove that Kerf_y is the subspace spanned by $\{y\}$. 6.40 Let V be the real vector space of 2 x 2 hermitian matrices. Prove that the mapping f: IR4 -» V given by "$w + x$ $y + iz$ $f(x,y,z,w) = y-iz$ $w-x$ is an isomorphism. ASSIGNMENT EXERCISES (1) Let A, B be subspaces of a finite-dimensional vector space V. Recall from Exercise S.34 that the smallest subspace of V that contains $A \cup B$ is given by $A + B = \{a + b ; a \pounds A, b \pounds B)$. Consider the mapping f: A x B —» V that is given by $f\{a,b\} = a + b$. Prove that Ker$f$= $\{(jc, -jc) ; jc \in A \cap B\}$ and use the Dimension Theorem to show that $\dim(A + B) = \dim A + \dim B - d\backslash m(A \cap B)$. (2) Show that the set E of real numbers of the form $a + b\sqrt{2} + c\sqrt[3]{}i$ (fl.fr.cGO) is a vector space over $\mathbb{Q}$. Prove that for a, b, c $\in \mathbb{Q}$ the polynomials $a + bX + cX^2$ and $X^3 - 2$ have no common factor. Hence establish a vector space isomorphism f: E -»$\mathbb{Q}^3$.

7 The Matrix Connection We shall now proceed to show how a linear mapping from one finite-dimensional vector space to another can be represented by a matrix. For this purpose, we require the following notion. Definition Let V be a finite-dimensional vector space over a field F. By an ordered basis of V we shall mean a finite sequence (v_i), $1 \le i \le n$ of elements of V such that $\{v_1,..., v_n\}$ is a basis of V. Note that every basis of n elements gives rise to n! distinct ordered bases, for there are n! permutations on a set of n elements, and therefore n! distinct ways of ordering the elements $v_1,..., v_n$. In what follows we shall find it convenient to abbreviate $(v_i)_{1 \le i \le n}$ to simply $(v_i)_n$. Suppose now that V and W are vector spaces of dimensions m and n respectively over a field F. Let $(v_i)_m$, $(w_j)_n$ be given ordered bases of V, W and let f: V -» W be linear. We know from Corollary 1 of Theorem 6.7 that f is completely and uniquely determined by its action on the basis $(v_i)_m$. This action is described by expressing each $f(v_i)$ as a linear combination of elements from the basis $(w_j)_n$: $f(v_1) = x_{11} w_1 + x_{12} w_2 + \cdots + x_{1n} w_n$, $f(v_2) = x_{21} w_1 + x_{22} w_2 + \cdots + x_{2n} w_n$, $f(v_m) = x_{m1} w_1 + x_{m2} w_2 + \cdots + x_{mn} w_n$. The action of f on $(v_i)_m$ is therefore determined by the mn scalars $x^\wedge$ appearing in the above equations. Put another way, the action of f is completely determined by a knowledge of the m x n matrix $X = [x_{ij}]$. For technical reasons that will be explained later, the transpose of this matrix X

114 Basic Linear Algebra is called the matrix of/ relative to the fixed ordered bases $(v,)_m$, $(w()_{,,}$. When it is clear what these fixed ordered bases are, we denote the matrix in question by Mat /. • The reader should note carefully that it is an n x m matrix that represents a linear mapping from an m-dimensional vector space to an n-dimensional vector space. Example 7.1 Consider the linear mapping/: IR3 —»IR2 given by f{x,y,z)={2x-3y + z, 3x-2y). The action of/ on the natural basis of IR3 is described in terms of the natural basis of IR2 as follows: /(1,0,0)= (2,3)= 2(1,0) + 3(0,1) /(0,1,0) = (-3,-2)=-3(1,0)-2(0,1) /(0,0,1)= (1,0)= 1(1,0) + 0(0,1) and so we see that the matrix of/ relative to the natural ordered bases of IR3 and IR2 is the transpose of the above coefficient matrix, namely the 2 x 3 matrix '2 -3 f 3 -2 0 ' • Note how the rows of this matrix relate to the definition of/. Example 7.2 The vector space IR,,[X] is of dimension n + 1 and has the natural ordered basis $\{i,x,x2,...,r\}$. The differentiation mapping D : IR,,[X] -» Bn[X] is linear, and Dl = 0 \ +0 >X + > • + 0 Xn~l +0 ■ Xn DX =\\+0X+-+0Xn-l+0Xn dx2 = o-i+2-jf + -+or' + or DXn = 0 1 +0 X + • • + n Xn~l +0 Xn so the matrix of D relative to the natural ordered basis of IR,,[X] is the (n +1) x («+1) matrix ■0 1 0 ... o- 0 0 2 ... 0 0 0 0 ... n 0 0 0 ... 0

7. The Matrix Connection 115 EXERCISES 7.1 Consider the linear mapping/: F2 -»IR3 given by f{x,y)={x + 2yt2x-yt-x). Determine the matrix of/ (1) relative to the natural ordered bases; (2) relative to the ordered bases {(0,1), (1,1)} and {(0,0,1),(0,1,1),(1,1,1)}. 7.2 Consider the linear mapping/: IR3 -»IR2 given by f(x,y,z) = {2x-y,2y-z). Determine the matrix of/ (1) relative to the natural ordered bases; (2) relative to the ordered bases {(1,1,1),(0,1,1), (0,0,1)} and {(0,1),(1,1)}. 7.3 Consider the linear mapping/: IR3 -»IR3 given by fix,y,z)={2x + z,y-x + z, 3z). Determine the matrix of/ with respect to the ordered basis {(1,-1,0),(1,0,-1),(1,0,0)}. 7.4 Suppose that the mapping/: R3 -»IR3 is linear and such that /(1,0,0)=(2,3,-2)-, /(1,1,0)=(4,1,4)-, /(1,1,1)= (5,1,-7). Find the matrix of/ relative to the natural ordered basis of IR3. It is natural to ask what are the matrices that represent sums and scalar multiples of linear mappings. The answer is as follows. Theorem 7.1 If VtW are of dimensions m,n respectively and if ftg : V —> W are linear then, relative to fixed ordered bases, Mat(/+ g) - Mat/+ Mat g and, for every scalar \, MatX/=XMat/.

116 Basic Linear Algebra Proof Let Mat / = $[jc,y]_{nxin}$ and Mat g = $[yiy]_{nxm}$

relative to fixed ordered bases (v,),,, of V and (wf)n of W. Then for i = 1,..., m we have (recalling the transposition involved) n n y=i y=i It follows that n and therefore Mat (f + g) = [xy + yij]nxm = Mat / + Mat g. Similarly, for every scalar X we have (V)(vJ = X>yiwy /-1 and so Mat X/ = X Mat /. D We can express Theorem 7.1 in a neater way as follows. Consider the set Linm n{V, W) of linear mappings from a vector space V of dimension m to a vector space W of dimension n (each over the same field F). It is clear that, under the usual addition and multiplication by scalars, Linmn(V, IV) is a vector space. Consider now the mapping ^L<n(V\W)-+MatnxmF given by !?(/)= Mat/ where V and W are referred to fixed ordered bases (v,),,, and (wt)n throughout. This mapping t? is surjective. To see this, observe that given any n x m matrix M = [ffijy], we can define n (1= 1,...,w) /(V|)=^m7lwy. By Theorem 6.7, this produces a linear mapping / : V -» W; and clearly we have Mat /= A/. Moreover, 1? is injective. This follows immediately from Corollary 2 of Theorem 6.7 and the definition of the matrix of a linear mapping. Thus t? is a bijection and, by Theorem 7.1, is such that ti(f + g) = t?(/") + t?(g) and t?(X/) = \ti(f). In other words, t? is a vector space isomorphism and we have: Theorem 7.2 IfV,W are of dimensions m, n respectively over F then Lin^HO-Mat^F. D

7. The Matrix Connection 117 It is reasonable to ask if, for given ordered bases (v,)m of V and (w/)n of W, there is a 'natural' basis for the vector space Lin mn(V, W). Indeed there is, and this can be obtained from the natural basis of Matnxm Ft namely {EM\ p= 1,....n and q=\,...tm) where EM is the matrix that has 1 in the (p, q)-lh position and 0 elsewhere. To see this, consider the linear mapping/w : V —» W given by wp if i = q\ (v,)={ W1" ■ n u • 0 otherwise. Then we have /w(v,) = 0w,+ .. + 0wp+-+0wn /w(vJ = 0w, + - . + 0wp+. .+0wB from which we see that Mat /w = £w, i.e. that t?(/^) = E^. Now since the inverse of an isomorphism is also an isomorphism, it follows by Theorem 6.5 that a (natural) basis for Linmn(Vr, W) is {fpq\ P= 1,...,1 and 0=l,...,/n}. We now turn our attention to the matrix that represents the composite of two linear mappings. It is precisely in investigating this that we shall see how the definition of a matrix product arises in a natural way, and why we have chosen to use the transpose in the definition of the matrix of a linear mapping. Consider the following situation: U; Mm -^- V; (v,),, -^- W't (W<), in which the notation U\ (u()m for example denotes a vector space U with a fixed ordered basis («,),,, and f\A denotes a linear mapping / represented, relative to the fixed bases, by the matrix A. The composite mapping is described by v-M)m—

&■—>wi(Wl)r What is the matrix of this composite linear mapping? It is natural to expect that this will depend on A and B. That this is so is the substance of the following result. Theorem 7.3 Mat(go/) = Matg.Mat/.

118 Basic Linear Algebra Proof To find Mat (g of) we must express each image (g o /)(«,) in terms of the ordered basis (wt)p of W. Now since Mat / = A we have n 0=1 "0 /(«,)= 5>/ivy. /= i and since Mat g = fl we have p (/= 1,....n) g(vy)=Efr*yM'*- Thus, for each i, #,)1 = ^^) = ^^(^.) ;=l /= i " / p \ /=1 *=l *=i /= i n Consequently the (A, i)-th element of Mat (g of) is £ fc^flyi, which is precisely the (*, i)-th element of BA = Mat g Mat /. D Corollary A square matrix is invertible if and only if it represents an isomorphism. Proof Suppose that A is an n x n matrix that is invertible. Then there is an n x n matrix B such that BA = /,,. Let V be a vector space of dimension n and let (v,),, be a fixed ordered basis of V. If/, g : V —► V are linear mappings that are represented by A, B respectively then by Theorem 7.3 we have that g o/ is represented by BA = /,,. It follows that g o/ = idv whence, by Theorem 6.5,/ is an isomorphism. Conversely, if / : V —» V is an isomorphism that is represented by the matrix A then the existence of/"1 such that/"1 o/ = idv implies the existence of a matrix B (namely that representing/"1) such that BA = /,,, whence A is invertible. D Example 7.3 Consider IR3 referred to the natural ordered basis. If we change reference to the ordered basis {(1,1,0), (1,0,1),(0,1,1)}

7. The Matrix Connection 119 then the matrix of the identity mapping is obtained from the equations id(l,0,0) =(1,0,0)= 1(1,1,0)+1(1,0,1)-5(0,1,1) id(0,1,0) =(0,1,0)= 1(1,1,0)-1(1,0,1)+1(0,1,1) id(0,0,l)= (0,0,1)= 4(1,1,0)+1(1,0,1)+1(0,1,1) i.e. it is 1 1 -1 1 -1 1 -1 1 1 The identity mapping being an isomorphism, this matrix is invertible. • The reader should note that it is to maintain the same order in which g,f appear in Theorem 7.3 that we choose to call the transpose of the coefficient matrix the matrix of the linear mapping. If, as some authors do, we were to write mappings on the right (i.e. write xf instead of/(jc)) then this convention is unnecessary. EXERCISES 7.5 A linear mapping/: IR3 -»IR3 is such that /(1,0,0)=(0,0,1),/(1,1,0)= (0,1,1),/(1,1,1)=(1,1,1). Determine /(x,y,z) for all (x,ytz) G IR3 and compute the matrix of/ relative to the ordered basis B= {(1,2,0),(2,1,0),(0,2,1)}. If g : IR3 —»IR3 is the linear mapping given by g(*,;y,z)= {2x,y + z,-x), compute the matrix of/ ogof relative to the ordered basis B. 7.6 Show that the matrix A = 2

0 4 -1 1 -2 2 3 3 is invertible. If/ : IR3 —» IR3 is a linear mapping whose matrix relative to the natural ordered basis of IR3 is A, determine the matrix of/"1 relative to the same ordered basis. 7.7 If a linear mapping/: V —» V is represented by the matrix A prove that /" is represented by A".

120 Basic Linear Algebra We now consider the following important question. Suppose that we have the situation v;(v,L—-—>w-(wt)m. If we refer V to a new ordered basis (vj)m and W to a new ordered basis (tvj)n then clearly the matrix of/ will change. How does it change? To see how we can proceed, consider the particular case where W = V and / is the identity mapping on V. We then have the situation V;(v,L — >V;(v\)m T T old basis new basis • This is precisely the situation described in the previous example. We call A the transition matrix from the basis (v()m to the basis (v))m. The following result is clear from the Corollary of Theorem 7.3: Theorem 7.4 Transition matrices are invertible. D We can now describe how a change of bases is governed by the transition matrices that are involved. Theorem 7.5 [Change of bases] If a linear mapping f:V—*Wis represented relative to ordered bases (v,)m, (w,),, by thenxm matrix A then relative to new ordered bases (vj)m, (wj),, the matrix representing f is thenxm matrix Q~lAP where Q is the transition matrix from (w))n to {w{)n and P is the transition matrix from (vj)m to (v,)m. Proof Using the notation introduced above, consider the diagram V;(v,), >W;(ua i<V./» idw.Q v;«L >w; (*;.),, /** We have to determine the matrix X. Now this diagram is 'commutative* in the sense that travelling from the southwest corner to the north-east corner is independent of whichever route we choose; for, clearly, f o\dv = f = idwo f. It therefore follows by Theorem 7.3 that the matrices representing these routes are equal, i.e. that AP = QX. But Q, being a transition matrix, is invertible by Theorem 7.4 and so X = Q~lAP. □

7. The Matrix Connection 121 Example 7.4 Suppose that/: IR3 -* IR3 is the linear mapping whose matrix relative to the natural ordered basis is " 1 -3 f A = 3-2 0 -4 1 2 and let us compute the matrix of/ when IR3 is referred to the ordered basis B= {(1,-1,1),(1,-2,2),(1,-2,1)}. We apply Theorem 7.5 with W = V = IR3, {wt) = (v,) =the natural ordered basis, and (wj) = (vj) =the new ordered basis B. The transition matrix from the new ordered basis to the old is 1 1 1 P = -1 -2 -2 1 2 1 • Note that this is obtained by taking the elements of B and turning them into the columns of P. This becomes clear on observing that we have id(l,-1,1)= (1,-1,1)= 1(1,0,0)-1(0,1,0)+ 1(0,0,1) id(l,-2,2) = (1,-2,2)=

1(1,0,0)-2(0,1,0) + 2(0,0,1) id(l,-2,1) = (1,-2,1)= 1(1,0,0)-2(0,1,0)+ 1(0,0,1) and the transition matrix is the transpose of the coefficient matrix. Now P is invertible (by Theorem 7.4) and the reader can easily verify that P~' = 2 1 0 -1 0 1 0 -1 -1 The matrix of/ relative to the new ordered basis B is then, by Theorem 7.5, P~lAP = 15 25 23 -8 -11 -12 -2 -5 -3 Example 7.5 Suppose that the linear mapping/: IR3 —> IR2 is represented, relative to the ordered bases{(l,0,-l),(0,2,0),(l,2,3)}oflR3and{(-l,l),(2,0)}oflR2,bythematrix A = 2-13 3 1 0 To determine the matrices that represent/ relative to the natural ordered bases, we first determine the transition matrices P, Q from the natural ordered bases to the

122 Basic Linear Algebra given ordered bases. Making use of the observation in the previous example, we can say immediately that ^10 1 >-i <r = r-i 2i 1 0 The reader can easily verify that P=\ P~l = 3 0-1 -1 2 -1 1 0 1 0 2 2 -10 3 and hence that the required matrix is <2-Mp=i 3 3-5 5 -1 1 We now establish the converse of Theorem 7.5. Theorem 7.6 Let {vj)mi {\Vj)n be ordered bases of vector spaces V, W respectively. Suppose that A, B are n x m matrices such that there are invertible matrices P, Q such that B = Q~XAP. Then there are ordered bases (v|)m, (u>-)„ of V, W and a linear mapping f : V —*W such that A is the matrix off relative to (v,)m, {w,)n and B is the matrix off relative to (vJL.K),,. Proof If P = \PijLxm &"& Q = fajlnxn* define m n ('=1 «) vi = X>;.v/' (i =!,...,«) w\\ = Y,aawi- Since P is invertible there is, by the Corollary of Theorem 7.3, an isomorphism fP : V —* V that is represented by P relative to the ordered basis (r,-)m- Since by definition v\ = fP(v,) for each i, it follows that (vj)m is an ordered basis of V and that P is the transition matrix from (vj)^ to (v,)m. Similarly, (w[)n is an ordered basis of W and Q is the transition matrix from (uj),, to (w,-),,. Now let/: V —» W be the linear mapping whose matrix, relative to the ordered bases (v,-)m and (h'-)m is A. Then by Theorem 7.5 the matrix of/ relative to the ordered bases (v;)m and (w'X is Q~lAP = B. □ EXERCISES 7.8 Determine the transition matrix from the ordered basis {(1,0,0,1),(0,0,0,1),(1,1,0,0),(0,1,1,0)} of IR4 to the natural ordered basis of IR4.

7. The Matrix Connection 123 "0 1 1 1 f 0 1 1 0 7.9 Consider the linear mapping/: IR3 -»IR3 given by f{x,y,z) = (y,-x,z)- Compute the matrix A of/ relative to the natural ordered basis and the B matrix of/ relative to the ordered basis {(1,1,0),(0,1,1),(1,0,1)}. Determine an invertible matrix X such that A =

X~lBX. 7.10 Let/ : IR3 —► IR3 be a linear mapping which is represented relative to the natural ordered basis by the matrix A. If P is the invertible matrix P = determine an ordered basis of IR3 with respect to which the matrix of/ is P~lAP. We have seen in Chapter 3 that a matrix of rank p can be transformed by means of row and column operations to the normal form 7, 0 0 0 We can also deduce this as follows from the results we have established for linear mappings. The proof is of course more sophisticated. Let V and W be of dimensions m and n respectively and let/: V —» W be a linear mapping with dim Imf- p. By Theorem 6.4, we have dim Ker / = dim V - dim Im / = m -/?, so let {v,,..., vm_p} be a basis of Ker/. Using Corollary 1 of Theorem 5.8, extend this to a basis B={Ui Mp,v, vm.p} of V. Observe now that IfM /(",)} p p is linearly independent. In fact, if j^A/faj) = 0 then^j^w,) = 0 and so i=I i=I p p m-p 53 \«/ G Ker/ whence £ ^/M/ = 51 /f/vy ^en i=I i=I ;=I P nt-P EM/- E/w = o i=i y=i and so, since B is a basis, every \{ and every nj is 0.

124 Basic Linear Algebra It follows by Corollary 2 of Theorem 5.8 that {/(«i),...,/(«p)} is a basis of the subspace Im /. Now extend this, courtesy of Corollary 1 of Theorem 5.8, to a basis C={f{ui) f{up),wlt...,wn_p} of W. Then, since/(v,) = • • = f(vm_p) = 0, we have /(«i) = 1/(«i) + 0/(«2) + • • + 0/(«p) + ..- + 0wn.p; f(u2) = 0/(«,) + 1/(«2) + ■• + 0/(«p) + ..+ o*vP; f(up) = 0/(«,) + 0f(u2) + .. + l/(«p) + .. + 0vvp; /(v.) = 0/(«,) + 0/(«2) + • • • + 0/(«p) +•+ 0w,,_p; /(^-,) = 0/(«,) + 0/(«2) + • • • + 0/(«p) + ..+ 0w,,_p. The matrix of/ relative to the ordered bases B and C is then [" °1 0 0j' where p is the rank of/. Suppose now that A is a given n x m matrix. If, relative to fixed ordered bases BV,BW this matrix represents the linear mapping/ : V -+ W then, Q and P being the appropriate transition matrices from the bases B, C to the fixed ordered bases Bv and Bw, we have *♦"-[;';;]■ Now since transition matrices are invertible they are products of elementary matrices. This means, therefore, that A can be reduced by means of row and column operations to the form 0 0 " The above discussion shows, incidentally, that the rank of a linear mapping f is the same as the rank of any matrix that represents /. Definition If A, B are n x n matrices then B is said to be similar to A if there is an invertible matrix P such that B = P~lAP. It is clear that if B is similar to A then A is similar to B; for then A = PBP-1 = {P-1)-1AP-K

7. The Matrix Connection 125 Also, if B is similar to A and C is similar to B then C is similar to A\ for B = P~lAP and C=Q~lBQ give c=0-^4/^=(per'^e-

Thus the relation of being similar is an equivalence relation on the set of n x n matrices. The importance of similarity is reflected in the following result.

Theorem 7.7 Two n x n matrices A,B are similar if and only if they represent the same linear mapping relative to possibly different ordered bases. Proof This is immediate from Theorems 7.5 and 7.6 on taking W = V and, for every /, W; = v,- and w\ = v\. D Corollary Similar matrices have the same rank. D The notion of similar matrices brings us back in a more concrete way to the discussion, at the end of Chapter 4, concerning the problem of deciding when (in our new terminology) a square matrix is similar to a diagonal matrix; or, equivalcntly, when a linear mapping can be represented by a diagonal matrix. We are not yet in a position to answer this question, but will proceed in the next chapter to develop some machinery that will help us towards this objective. EXERCISES 7.11 Show that if matrices A, B are similar then so are A', B>. 1.12 Prove that if A, B are similar then so are Ak, Bk for all positive integers it. 7.13 Prove that, for every t? G IR, the complex matrices are similar. "cost? sin t? -sin t? cost? i [V 0] 0 e~"m SUPPLEMENTARY EXERCISES 7.14 Determine the matrix of the differentiation map D on IR,,[X] relative to the ordered bases (1) {i,x,x2,...,*"}; (2) {X\Xn~x X,l}; (3) {1,1+X,1 + X2,...,1+X"}.

126 Basic Linear Algebra 7 15 Let V be a vector space of dimension n over a field F. A linear mapping /: V -» V is said to be nilpotent if f = 0 for some positive integer p. The smallest such integer p is called the index of nilpotency of/. Suppose that/ is nilpotent of index p. If x G V is such that/?'"(*) ^ 0, prove that {xj(x)j2(x),~,rl(x)} is linearly independent. Hence prove that/ is nilpotent of index n if and only if there is an ordered basis {vj)n of V such that the matrix of/ relative to (v,),, is of the form ro o o ... o o 1 0 0 ... 0 0 0 1 0 ... 0 0 0 0 1 ... 0 0 0 0 0 1 0 7.16 Let/: R,,[X] -> R,,[X] be given by f(p(X)) = p(X+\). Prove that/ is linear and find the matrix of/ relative to the natural ordered basis {1, X,..., X"}. 7.17 Consider the mapping/: IR2[X] —» Mat2x2 R given by [D C Show that/ is linear and determine the matrix of/ relative to the ordered bases {1,X, 1 + X2} of IR2[XJ and {[1 0 0 0 "i r 0 0 i "i i" 1 0 * "i r i i of Mat 2X2IR- Describe also Im / and Ker /. 7.18 Let/: R3[XJ -» R3[XJ be given by /(a + bX + cX2 + <fX3) = a + (d - c - a)X + (d - c)X\ Show that/ is linear and determine the matrix of/ relative to (1) the natural ordered basis {1,X,X2,X3}; (2) the ordered basis {1 + X3, X, X + X3, X2 + X3}. 7.19 Ifv4 = [flylnx,, then the trace of A is defined to be the sum of the diagonal elements

7. The Matrix Connection 127 Prove that (1) tr(Xv4) = Xtrv4; (2) tr(v4 + B) = trv4 + trB; (3) tr(v4B) = tr(Bv4). Deduce from (3) that if A and B are similar then they have the same trace. Give an example of two matrices that have the same trace but are not similar. ASSIGNMENT EXERCISES (1) If/ : MatnxJR —» IR is linear, prove that the following statements are equivalent: (a) f(AB) = f{BA) for all A,B e Matnxn IR; (b) / is a scalar multiple of the trace function. Deduce that X G Matnx,, IR can be written as a sum of matrices of the form AB-BA if and only if tr(X) = 0. (2) Consider the linear mapping / : IR3 -» IR3 whose matrix relative to the natural basis {¢,, e2, ¢3} is A = 0 1 -sin t? -1 0 cos t? - sin t? cos t? 0 Show that {e\, e'2, e'3} is a basis of IR3 where tf; = tf,cos0 + tf2sin0, 4=/(^1). e3=f(e2). Determine the matrix of / relative to this basis. (3) Let V be the real vector space of functions q : IR2 —»IR given by a prescription of the form 0(*. y) - axl + bxy + cy2 +dx + ey+f. Let (p : V -» V be the mapping described by setting <p{q) =^ I<i(x>y)dy + q-j i(x>y)dx- Show that <p is linear. Determine the matrix that represents <p relative to the ordered basis B of V given by B={x2txy,y21x,yt\}.

128 Basic Linear Algebra (4) Let V be a finite-dimensional vector space and let B = {v,,..., v,,} be an ordered basis of V. Let C be the ordered basis of V that is obtained from B by interchanging v, and vl+, in the ordering. If/ : V —* V is linear, explain how Matc/ is related to Mat Bf. Describe i n general the effect on Mat Bf that is caused by a change in the ordering offl. (5) Let V be a finite-dimensional vector space over IR and let/ : V -* V be a linear mapping such that/2 = -idv. Extend the action of IR on V to an action of C on V by defining, for all x € V and all a + ib£ C, (a + ib)x = ax - bf(x). Show that in this way V becomes a vector space over C. Use the identity r r r E(fl/ - ibt)vt = E aivt + E bJ(vt) i=l /=I r=l to show that if {v,,..., vr} is a linearly independent subset of the C-vector space V then {v,,..., vr,/(v,),...,/(vr)} is a linearly independent subset of the R-vector space V. Deduce that the dimension of V as a complex vector space is finite and that dimB V- 2dim€ V. Hence show that a In x In matrix A over IR is such that A2 = -I2n if and only if A is similar to the matrix ro -i, 1. 0

8 Determinants In what follows it will be convenient to write annxn matrix A in the form A = [81,82,...,8,,] where, as before, a, represents the /-th column of

A. Also, the letter F will signify either the field IR of real numbers or the field C of complex numbers. Definition A mapping D : Matnxn F -* F is determinantal if it is (a) multilinear (or a linear function of each column) in the sense that (0,)^...,^+^,...1 = ^...,^,...1 + ^...,^...1: (D2) £)[...,Xa,,...] = XZ) [...,a,,...]; (b) alternating in the sense that (D3)Z)[...,a,.,...,a;,...] = -Z) [...,ay,...,a,.,...]; (c) 1-preserving in the sense that (D4)D(/n)=lf. We first observe that, in the presence of property (D,), property (D3) can be expressed in another way. Theorem 8.1 If D satisfies property (D,) then D satisfies property (D3) if and only if it satisfies the property (1¾) D(A) = 0 whenever A has two identical columns. Proof => : Suppose that A has two identical columns, say a, = ay with i fj. Then by (D3) we have D(A) = -D(A) whence D(A) = 0.

130 Basic Linear Algebra <= : Suppose now that D satisfies (D,) and (LV3). Then we have (/ = ^...,3,+8, 8,+3,,...] l = * H~.t*i 3,+3,,...)+1)1...,3, 3,+3,,...] ^ £[...,3,- a,,...] + £)[...,3, a,,...] +D[...,SLj 3,,...1+1)1...,3, 3,,...] l=* D[...,ai 3,,...] + £)[...,a,,...,a,,...] whence (D3) follows. D Corollary D is determinantal if and only if it satisfies (Dj),(D2),(D3).(D4). D - aua22 -0|2fl2i- Example 8.1 Let D : Mat2x2 F -» F be given by D\a\\ an .fl2l fl22. Then it is an easy exercise to show that D satisfies the properties (D,).(D2),(D3),(D4) and so is determinantal. In fact, as we shall now show, this is the only determinantal mapping definable on Mat 2X2 F. For this purpose, let 6,= & = so that every A € Mat2x2 F can be written in the form A = [fl‖6| + ^21^2» fll2^l "*" fl22^2j- Suppose that/: Mat2x2 F —» F is determinantal. Then, by (D,) we have f(A)=f[au6ll ax2bx + fl22^] +/[fl2i^2. a\i^\ + ^22¾] Applying (D,) again, the first summand can be expanded to /[fl,,6,, fl|26,] +/Ifl‖6|, ^22¾] which, by (D2), is By (D'3) and (D4), this reduces to 0^22- As for the second summand, by (D,) this can be expanded to f[a2l62t al26x] +f[a2l62, a^b^

8. Determinants 131 which, by (D2), is fl2lfll2/[^2,^J + fl2lfl22/[^2.^]- By (D'3) and (D4), this reduces to -alxal2. Thus/(;4) = a\\a22~a\2a2\ an(* we conclude that there is a unique determinantal mapping on Mat2x2 F. In what follows our objective will be to extend the above observation to Mat,,XII F for every positive integer n. The case where n = 1 is of course trivial, for if A is the 1 x 1 matrix [a] then clearly the only determinantal mapping is that given by D(A) = a. For every A € Mat,,XII F we denote by Ati the (n -1) x (n -1) matrix obtained from A by deleting the i-th row and thej-th column of A (i.e. the row

and column containing atJ). The following result shows how we can construct determinantal mappings on the set of n x n matrices from a given determinantal mapping on the set of (n-1) x (n-1) matrices. Theorem 8.2 For n ^ 3 let D : Mat(n_|)x(n_|) F —» Fbe determinantal, and for i= 1,..., n define fr.MMnxnF^Fby /=i Then each ft is determinantal. Proof It is clear that D(Ai}) is independent of the j-th column of A and so avD(A jy) depends linearly on the j-th column of A. Consequently, we see that f, depends linearly on the columns of A, i.e. that the properties (D,) and (D2) hold for/. We now show that f{ satisfies (1¾). For this purpose, suppose that A has two identical columns, say thep-th and the g-th columns with p f q. Then for j f p and j f q the (n - 1) x (n - 1) matrix A^ has two identical columns and so, since D is determinantal by hypothesis, we have (/Yp.«) 0Mij) = o. It follows that the above expression for/-(A) reduces to MA) = (-\)i+"aipD(Aip) + (-ir<aiqD(Aiq). Suppose, without loss of generality, that p < q. Then it is clear that Aiq can be transformed into Aip by effecting q -1 -p interchanges of adjacent columns; so, by (D3) for D, we have D{Aiq)= (-\rl-"D(Aip).

132 Basic Linear Algebra Since aip - aiq by hypothesis, we thus have MA) = [(-1)'♦' + (-\)i+H-\rl-p)aipD(Aip) which reduces to 0 since (_!)'>/> + (_i)/+«(_i)?-i-p = (-1)^(1 + (-1)29-2^1] = (-\)i+r[\ +(-1)] = 0. Finally,/ satisfies (D4) since if A = ln then aVi = 6^ and Au = /,,_|, so that /(/n)=(-i)'%W,,-.)=i- Thus/ is determinantal for every /. D Corollary For every positive integer n there is at least one determinantal mapping on Mat nxnF. Proof We proceed by induction. By Example 8.1, the result is true for n = 2. The inductive step is Theorem 8.2 which shows that a determinantal mapping can be defined on Mat,,XII F from a given determinantal mapping on Mat(,,_i)x(»i-i) F- E D - «ll«22 ~«|2«2I» Example 8.2 If D is the determinantal mapping on the set Mat2x2 Ft i.e. if «11 «12 L«2I «22. then by Theorem 8.2 the mapping/, : Mat3x3 F —» F given by /. is determinantal. Likewise, so are/2 and/3 given by fi «11 «12 « 13 «21 «22 «23 «31 «32 «33. = «l|£ «2 .«3 2 «23 2 «33. -«I2# «21 «23 .«31 «33. + al3D «21 «22 .«31 «32. «11 «12 « 13 «21 «22 «23 «31 «32 «33. = a2lD «1 .«3 2 «13* 12 «33. -a2iD «11 «13 .«31 «33. + a23D «11 «12 .«31 «32.

8. Determinants 133 /3 «11 «12 « 13 «21 «22 «23 «31 «32 « = «3.*> 33 J «12 «13 .«22 «23. -a32D «11 «13 .«21 «23. + fl33Z) «11 «12 .«21 «22. EXERCISES 8.1 Using the formula D «11 «12 .«21 «22 - «ll«22 ~«I2«2I»

evaluate each of/j(v4),/2(v4),/3(v4) in the above example. What do you observe? 8.2 Show that the mapping/: Mat3x3 F —» F given by / «11 «12 «1 3 «21 «22 «23 «31 «32 «2 = «,,£> 13 _ 22 «23 32 «33. -a2lD «12 «13 .«32 «33. + a3lD «12 «13 .«22 «23. is determinantal. Our objective now is to establish the uniqueness of a determinantal mapping on Matnxn F for every positive integer n. For this purpose, it is necessary to digress a little and consider certain properties of permutations (= bijections) on a finite set. On the set {1,..., n} it is useful to write a permutation / in the form 1 2 3 ... n /(1) /(2) /(3)... f(n) Example 8.3 The permutation/on {0,..., 9} described by/(jc) = x+\ modulo 9 can be described by / = 0123456789 1234567890

134 Basic Linear Algebra Example 8.4 The permutation on {1,2,3,4,5} described by _/12345 I 1 245 3 'fixes' 1 and 2, and permutes cyclically 3,4,5. Given permutations /, g on {1,..., n} we can compute the composite permutation gofby simply treating them as mappings: 1 ... n M\)\-gW)\ Example 8.5 Consider the permutations (1 2 3 4 5 6\ /l 2 345 6 164352/ ^2 6 5 3 14 Working from the right, we compute the composite permutation go fas follows: (1 2 3 4 5 6\ /123456\ /l 2 3 4 5 6 265 3 14/ ^164352/ V 2 4 3 5 I 6 EXERCISES 8.3 Compute the products 12 34567 *\ /12 34 5678 8623145 7/ \5 67 184 32 1 2 34567 8\ /l 2 34567 8 8623145 7/ ^53467281 Definition By a transposition on the set {1,2,..., n} we mean a permutation that interchanges two elements and fixes the other elements. More precisely, a transposition is a permutation r such that, for some ij with i ^j\ r{i) = ;', r{j) = /, and (Vx f ij) t(x) = x.

8. Determinants 135 We shall sometimes denote the transposition r that interchanges i and j by the notation r : i *-*j. Clearly, the inverse of a transposition is also a transposition. The set of permutations on {1,2,...,«} will be denoted by Pn. Theorem 8.3 If n > 2 then every a G P,, can be expressed as a composite of transpositions. Proof We establish the result by induction on n. When n = 2 it is clear that cr is itself a transposition, in which case there is nothing to prove. Suppose, by way of induction, that n>2 and that the result holds for all permutations in P,,_|. Let cr G P,, and suppose that o(n) = k. Let r be the transposition r : n *-* k. Then rocr is such that (rocr)(n) = n\ i.e. rocr fixes n, and so induces a permutation, (rocr)* say, in P,,_,. By the induction hypothesis, there are transpositions rj,..., r* in P,,_, such that (r o cr)* = t\ o ■ ■ o r*. Clearly, r*,..., r* induce transpositions in P,,, say T|,..., rr, each of which

fixes n, and r o a = tx o • ■ ■ o rr. It follows that a = T~l O T| O ■ ■ ■ O Tr as required. D Given cr € P,, let /(cr) be the number of inversions in cr, i.e. the number of pairs (/,;') with i < ;' and o(j) < o(i). Definition For every cr G Pn the signum (or signature) of a is defined by ea = (-1)7^. Theorem 8.4 (Vp.aePj epa = epea. Proof Consider the product v. = 11(/-0- For every a G Pn define a(v;)=nb(/)-cT(/)]. Since a is a bijection, every factor of VH occurs precisely once in o(Vn), up to a possible change in sign. Consequently we have °(K) = (-l)'MV,, = ecV,,.

136 Basic Linear Algebra Given p,o ePn we have similarly pa(Vn) = £po{Vn). Consequently, SpoVn = po(Vn)=8po(Vn)=8(>80Vn whence, since Vn f 0, we obtain 8pa = 8p8a. D Corollary IfoePn then ea = ±1 and 8a-i = 8a. Proof It is clear that et = -1 for every transposition r. It follows from Theorems 8.3 and 8.4 that 8a = ±1 forevery permutation a. Moreover, since the signum of the identity permutation is clearly 1, we deduce from £a£a-i = 8aoa-i = 1 that 8a-\ = 8a. D We say that o is an even permutation if 8a = 1, and an odd permutation if ea = -1. An even permutation is therefore one that has an even number of transpositions, whereas an odd permutation is one with an odd number of transpositions. This notion of parity is therefore an invariant associated with a permutation. Example 8.6 Consider the permutation / = 12 345 6 16435 2 If we join each i on the top line with the corresponding i on the bottom line we obtain the diagram 1 1 In this the number of distinct crossings gives the number of inversions. This is 8, which is even. The permutation is therefore even. EXERCISES 8.4 Determine the parity of each of the following permutations: 1 2 345 6\ /123456 614253/' V5436I2

8. Determinants 137 Theorem 8.5 There is a unique determinantal map D : Matnxn F —» F, and it can be described by D{A)= E£afla(l),l"fla(n),n- o£Pn Proof We know by the Corollary of Theorem 8.2 that at least one determinantal mapping D exists on MatnxlI F. If we write 6j for the i-th column of /,, then we can represent annxn matrix A = [a,-] by A = [all8l +-+anl8n aXnBx + -- + 0^6,]. Using property (D,) we can write D(A) as a sum of terms of the form ^[fla(l),l^a(1)»'' •» fla(n).A(«J. where 1 ^ a(i) ^ n for every i. Using property (D2) we can then write each of these terms as aa[l),l ' ' ■ fla(n),n^o(l). • • • . &o[n)]- But, by property (D'3), each such expression is 0 except those in which we have a(i) f o{j) for i f j\ i.e. those in which a is a permutation on {1,..., n}. Thus we have that D{A) = £ fla(l)il • • ■ aa{n)jnD[6a{l) 6a{n)]. o

£P9 Now the columns 6a(l),..., 6a(nj occur in the permutation o of the standard arrangement 6|,..., 6n. If, using Theorem 8.3, we write a as a composite of transpositions, say (J=T|0"-OTi(then we have a~l = tJ1 o ■ ■ • o rj"1. Restoring the standard arrangement of the columns by applying rj"1,..., rj' we see, by property (D3) and the fact that ea - ea-i, that But by property (D4) we have Z)[6,,..., 6,] = £)(/,,) = 1. We therefore conclude that D(A)= £ eaaa{l)tr'- aa{n)jl. a£P,, The above argument also shows that D is unique. D An important consequence of the above is that the expression for fk(A) given in Theorem 8.2 is independent ofi.

138 Basic Linear Algebra Definition The unique determinantal mapping on MatnxnF will be denoted by det. By the determinant of A = [a(y]nxn we shall mean det A. By Theorem 8.5, we have that detv4= Y, £aaa(l),l • • • aa[n),n- a £P. Alternatively, by Theorem 8.2, we have that, for i - \,..., n, detv4 = tH),+Sdetv4i>, which is called the Laplace expansion along the i-th row. • Note that, as pointed out above, the Laplace expansion is independent of the row chosen. Example 8.7 Consider the matrix A = 1 1 -1 2 1 3 1 -5 1 Using a Laplace expansion along the first row, we have ' 1 3' [-5 lj -1 • det '2 3" [l lj det A - 1 • det = 16-(-1)-(-11) = 28. Expanding along the second row, we obtain = -2 • det + 1 • d = -2(-4) + 2-3(-6) = 28. Finally, expanding along the third row we obtain + (-1). det det A = -2 • det 1 [-b -1" lj + 1 • det 1 [l -1" lj -3 det 2 1 1 -5 14] det A = 1 • det 1 -1 1 3 = 4 + 5 5 + (-1) = 28 -(-5) det 1 I2 -1" 3J + 1 • det 1 r [2 lj EXERCISES 8.5 Compute, via a third row Laplace expansion, ■12-3 4- det 3 0-20 . 1 0 2-5.

6. Determinants 139 8.6 Determine det 1-X 3 2 2 -1-X 3 3 0 1-X Theorem 8.6 If A is a square matrix then det A = det A'. Proof If a G PH then whenever o(i) - j we have i = o~l(j) and therefore aa^,y = fl)i<ri(/). Consequently, fla(l).l * • • aa{n),n = fll.a-'(l) ' ' ' fln.a-'(n) = Wlr»0).l ' " Wo'1 (n),n■ Now as a ranges over P,, so does o~l; and ea = ea-\. Thus det A = £ efffloji).! • • • aa{nU = £ ea-.[v4']a-'(i),i • • • [A')0-1{n)n = det v4'. D o€P. a-'FP. Corollary Fory = 1,..., n we have Proof We have i=l det A = det A' = £(-I)'+'ay/ det Ay, = EH),'%detv4/y, i=i the second summation being obtained from the first summation by interchanging i and). D The reader should note that in the above Corollary the summation is over the first index whereas in Theorem 8.2 it is over the second index. We can thus assert that Laplace expansions via columns are also valid. EXERCISES 8.7 For the matrix A of the previous example, find

det A by Laplace expansion via each of the three columns. 8.8 Compute det 12 3-4' 0-5 6-7 0 0-89 0 0 0 10

140 Basic Linear Algebra It is useful to know how row and column operations on a square matrix effect the determinant of that matrix. It is clear from property (D3) that • if B is obtained from A by interchanging two columns of A then det B = -det A. Also, from property (D2) it follows that • ifB is obtained from A by multiplying a column of A by some scalar A then det B = A det A. Finally, ifB is obtained from A by adding A times the i-th column of A to thej-th column then, by properties (D,) and (D'3), we have detfl = det[...,a, a> + Aa,,...] = det [...,a, a;,...] + det [...,a,,...,Aa,,...] = det A + A det [..., a,,..., a,,...] = det A + A0 = det A. Thus we have that • ifB is obtained from A by adding to any column of A a multiple of another column then det B = det A. Since row operations on A are simply column operations on A', and since det A' = det v4, it is clear from the above that similar observations hold for row operations. Example 8.8 Any square matrix that has a zero row or a zero column has a zero determinant. To see this, simply perform a Laplace expansion along that zero row or column. Example 8.9 Consider again the matrix "I 1 -1 v4= 2 1 3 1 -5 1 Using row operations, we have det A = det = det "l 1 -l" 0 -1 5 0 -6 -I 51 -6 2 2 l>> P2"2pi Pi-Pi by first column Laplace = -2 + 30 = 28.

8. Determinants 141 Example 8.10 Consider the matrix A = 1 0 -1 2 0 3 1 -5 1 By a Laplace expansion using the second column, we have 1 -1 det A = -(-5) det 2 3 = 5-5 = 25. Example 8.11 For the matrix we have detv4 = A = = 1 -1 2 31 2 2 0 2 12 3 0 .5 3 2 -1. ■1 -1 2 3" 0 4-4-4 0 3 1-3 .0 8 -8 -16. = det 4-4-4 3 1 -3 8 -8 -16 = 4-8-det = 32 det = 3 (2(- ■4) "l 0 0 = P2"2pi P3-P1 P4~5p, by first column Laplace "l -1 -l" 3 1 -3 1 -1 -1 -1 4 0 0 -1 -128. -2 >2-3p, »3"Pl EXERCISES 8.9 What is the determinant of an n x n elementary matrix? 8.10 Prove that the determinant of an n x n upper triangular matrix is the product of its diagonal elements. 8.11 Determine the values of x for which det A f 0 where x 2 0 31 12 3 3 10 11 1113 A =

142 Basic Linear Algebra 8.12 Using row and column operations, show that ■0 a 0 0 0 0" det / 0 b 0 0 0 0 g 0 c 0 0 0 0 h 0 d 0 0 0 0 k 0 <? 0 0 0 0 m 0 8.13 Compute the determinant of the matrix 12 3. -10 3. -1-2 0 . = -acefhm. -1 -2 -3 ... 0 8.14 Compute the determinant of the matrix A = [fliy]nxl, given by fl iff +7 = 11 + 1; a.i = 0 otherwise. 8.15 Consider the complex matrix A = 0 1 +

i 1+2/ 1-i 0 2-3« 1-2/ 2 + 3/ 0 Show that det A = 6. We now consider some further important properties of determinants Theorem 8.7 If A, Be Mat nxnFthen det AB - det A • det B. Proof If C = AB then the Jt-th column of C can be written To see this, observe that the i-th element of c* is n n T " 1 Mi = cik = E aijbjk = E Ma;l. = E fy*a> ■ y-i y-i ly=i J«

8. Determinants 143 Thus we have that det AB = det C = det[6,,a, + • • + bnlan fclna, + • • + bnn^ = Y, detK(i).iaa(i) ba{n)naa{n}] oGfl, = £ ^(d.i ''' K[n).n det [aa(I),..., aa(n)] = £ bo(\).\ • • ba[n),nto det[» »J = det A • £ C A(l),l " ba(n),n o £P9 = det A • det B. D Corollary /fA w invertible then det A j£ 0 and det A"1 = Proof det A This follows from det A • det A"1 = det AA~l = det /,, = I. D , p= "0 1 2" 2 0 I 1 2 0 EXEflC/SES 8.16 Given the matrices ~b + &c 2c-2b 4b -Ac A- 4c -4a c + %b 2a-2c 2b-2a 4a-4b a + 8fc find P~l and compute P~lAP. Hence determine det A. 8.17 If A is a square matrix such that Ap = 0 for some positive integer p, prove that det A = 0. If A is an invertible n x n matrix then we have seen above that det A f 0. Our objective now is to show that the converse of this holds. This will not only provide a useful way of determining when a matrix is invertible but will also give a new way of computing inverses. For this purpose, we require the following notion. Definition If A € Matnxn F then the adjugate (or adjoint) of A is the n x n matrix adj A given by [adjA]iy = (-irdetAyi, • It is important to note the reversal of the suffices in the above definition.

144 Basic Linear Algebra The adjugate matrix has the following useful property. Theorem 8.8 For every n x n matrix A, A • adj A = (det A)I,, = adj A • A. Proof We have 14 • adj 41,-, = £aik[ad)A)kj *=i = £ Oikdet Ajk k=l {det A if i=j\ 0 \U+j% the last equality resulting from the fact that when i f] the expression represents the determinant of a matrix whose/-th row is the same as its i-th row. Thus A • adj A is a diagonal matrix all of whose diagonal entries is det A; in other words, A • adj A = (det A)In. The second equality is established similarly. D EXERCISES 8.18 Compute the adjugate of each of the following matrices: a b c d a h g h b f g f c * -1 0 -1 0-10 -1 0 -1 Theorem 8.9 A square matrix A is invertible if and only ifdel A ^ 0, in which case the inverse is given by i-i >T = —— adj A. detv4 Proof If det A f 0 then by Theorem 8.8 we have whence A is invertible with A~l = adj A. det A J Conversely, if A l exists then, as we have observed above, it follows from Theorem 8.7 that det A f 0. □

8. Determinants 145 Theorem 8.9 provides a new way of computing inverses. In purely numerical examples, one can become quite skilful in its use. However, the adjugate matrix has to be constructed with some care! In particular, notice should be taken of the factor {-l),+> = ±1. The sign is given according to the scheme + -+- -+- + + -+- Example 8.12 We have seen previously that the matrix A = 1 1 -1 2 1 3 1 -5 1 is such that det A = 28. By Theorem 8.9, A is therefore invertible. Now the adjugate matrix is ' 1 -1 -5 1 det -det det 1 -5 "2 1 2 1 - 3 1 3 1 1 -5 -det det 1 -1 1 3 det -det 1 -1 1 1 1 1 1 -5 -det det 1 -1 2 3 1 1 2 1 i.e. it is the following matrix (which with practice can be worked out mentally): 16 4 4" B= 12-5 -11 6 -1 It follows by Theorem 8.9 that A'1 - —Bt which can of course be verified by direct multiplication. EXERCISES 28 8.19 For each of the following matrices, compute its adjugate and then its inverse: '3 1 2 1 2 1 1 1 1 > '5 3 2" 2 3 1 7 5 3 * "l 0 o" 1 2 0 1 2 3

146 Basic Linear Algebra 8.20 For the matrix A = -4 -3 -3 1 0 1 4 4 3 show that adj A = A. 8.21 If A is an invertible nxn matrix prove that detadjv4 = (detv4)"-1. 8.22 If A and B are invertible nxn matrices prove that adj AB = adj B • adj A. 8.23 If A is an invertible nxn matrix prove that adj(adjv4)= (detv4)',-2v4. Deduce that, for n = 2, adj (adj A) = A. 8.24 If A is an upper triangular matrix prove that so also is adj A. 8.25 If A is a symmetric matrix prove that so also is adj A. 8.26 If A is an hermitian matrix prove that so also is adj A. There are other methods of evaluating determinants that are often useful, depending on the matrices involved. For example, there is the so-called 'inspection method* which is best illustrated by example. Example 8.13 Consider the matrix A = 1 jc jc2 ,2 l y y 1 z z2 Observe that if we set jc = y then the first two rows are equal and so the determinant of A reduces to zero. Thus jc - y is a factor of det A. Similarly, so are x - z and y - z. Consider now the ^-expansion of det A as in Theorem 8.5. Every term a in this expansion consists of a product of entries that come from distinct rows and columns (since we are dealing with a permutation a). Now the highest power of x,ytz appearing in this expansion is 2. Consequently, we can say that det v4 = *(jc - >)(>■ -z)(jc-z) for some constant k. To determine it, observe that the product of the diagonal entries, namely yz2, is a term in the ^-expansion (namely, that which conesponds to the a identity permutation). But the term involving yz2 in the above expression for det A is -kyz2. We conclude therefore that k - -1 and so detv4 = (*-;y)(y-z)(z-*).

Example 8.14 Consider the matrix a b c d A= a2 & c2 d2 b+c+d c+d+a d+a+b a+b+c bed cda dab abc By the 'inspection method', factors of det A are a-bt a-ct a-dt b-c, b-dt c-d. Now the product of the diagonal entries, namely a2b3c[d + a + b)> is a term in the ^-expansion. But this is only partially represented in the product a (a-b)(a-c)(a-d)(b-c)(b-d)(c-d), which suggests that we have to find another factor. This can be discovered by adding row 1 to row 3: this clearly produces the factor a +b + c + d. Thus we have det A = k(a - b){a - c)(a - d)(b - c){b -d){c - d){a + b + c + d) for some constant k. Comparing this with the product a2blc{d+a+b) of the diagonal elements, we see that * = -1. Example 8.15 Consider the matrix A = ' X yl yz2 l y z2 a 1 z b c 1 _yzt zt t 1. If x = y then the first column is y times the second column whence the determinant is zero and so x - y is a factor of det A. If z = t then the third and fourth rows are the same, so z -1 is also a factor of det A. If now y = z then we have c 1 a b~\ det A = det = det z3 z z 1 c 2 - 1 = (1-/c)det .zt zt t 1. x 1 a b z2 z 1 c z3 z2 z 1 0 0 0 1-fc ; 1 a z2 z 1 z3 z2 z = 0 since p3 = zpi-

Thus we see that v - z is also a factor. It now follows that det A = k{x -y){y -z){z-t) for some constant k> and comparison with the product of the diagonal elements gives k=\. EXERCISES 8.27 For the matrix A = 1 a a2 a3 1 X X2 *3 1 b b2 b> express det A as a product of linear factors. 8.28 Solve the equation det x a a a a x a a a a x a a a a x = 0. 8.29 Consider the real matrix x y2 yz2 a ? ?n y a ? z2 z a yzt2 zt2 t2 t Show that, whatever the entries marked ? may be, this matrix has determinant (x-ay){y-az){z-at)t. SUPPLEMENTARY EXERCISES 8.30 Let An be the n x n matrix given by f0 ifi=;; °ij = 1 [1 otherwis otherwise Prove that det An = (-1)^(/1 - 1). 8.31 Consider the n x n matrix a + b a a A = a + b a a a + b a + b Prove that det A = lf~x{na + b)

8.32 Solve the equation det 1 1 1 1 \-x 1 1 1 2-jc .1 1 8.33 Consider the n x n matrix n -x = 0. *n = b b b abb -b a b -b -b -b -b -b -b b b b b b b b b a b Prove that 6elBn = {-\)n+lb{a-b)n-1. Hence show that if An is the n x n matrix a b b ... bl -b a b ... b -b -b a ... b then Deduce that -b -b -b ... a_ det An = {a + b) det An.x -b{a-b)n~\ det/1,,= \[{a + b)n + {a-b)n]. 8.34 Let An be the n x n matrix given by 0 if|i-;|>l; 1 if|i-;|=l; 2cosi? if 1=./. atj=< If A,, = det An, prove that An+2-2cosi?An+1+A,, = 0. Hence show by induction that, for 0 < t? < 7r, sin (/i+ l)tf det Am = sin $

150 Basic Linear Algebra 8.35 Let An be the n x n matrix given by bi ifi^;; fly = < n-l fl, + />, if i = j< n\ bn if i =; = «. Prove that det >4,, = bn J] fl,. i=l If#,, is given by 7 [fl. + fc, ifi=;\ prove that det Bn - det >4,, + an det £,,_,. Hence show that det *,,=n «,+£(*,. n *>)• .=1 i=l Jfi 8.36 If >4 and B are square matrices of the same size, prove that \A #1 det B A = det (>l + #) det (>*-#). 8.37 Let A/ = * Q R S where P, Q, R, S are square matrices of the same size and P is invertible. Find a matrix N of the form 7 P_I0 A 0 5 C such that 0 S-RP~XQ Hence show that if P and /? commute then det M - det [PS - RQ)\ and that if P and g commute then det A/ = det [SP - RQ). 8.38 [Pivotal condensation] Let A € MatRXR IR and suppose that a^ f 0. Let B be the [n -1) x (n -1) matrix constructed from >4 by defining ifl<i<p-l(l<;<9-l; >,, = det det - det det "fltf°pJ "flw flw fl,; -fl/y °M . fl/« fl/« flw au apj QP9 Qiq . °PJ «J if 1 ^1^/7-1,^ +1 <;'<«; if p + 1 ^ i < n(0 + 1 ^y ^ n. Prove that det >4 = i det 5.

8. Determinants 151 [Him. Begin by dividing the p-th row by am to obtain a matrix X in which jcw = I. Now subtract suitable multiples of the q-th column of X from the other columns of X to make the elements of the p-th row 0 except for xpq= \> thereby obtaining a matrix Y. Observe how the structure of the matrix B arises. Now consider a Laplace expansion of Y via the p-th row.] • The (n -1) x [n - 1) matrix B is called the matrix obtained from A by pivotal condensation using aM as a pivot. This useful (and little publicised) result provides a simple recursive way of computing the determinants of matrices and is particularly effective when they have integer entries. The size of the matrix reduces at each step and the calculations are simple since they involve 2x2 submatrices, and are made easier if it can be arranged that a 1 is chosen as a pivot. 3 2 3" = -17. For example, det = ^ del -6 -5 -8 -1 2 4 3 .4 3 HI Compute, via pivotal condensation, the determinants of 1 2 3 1 1 -5 -f 3 1 i "l 2 3" 2 3 1 3 1 2 * -1 0 -1 -2 -1 -2 -3 ASSIGNMENT EXERCISES (1) Given a real square matrix A = [fl,;],,x,,, consider a system Ax = b of n linear equations in n unknowns. Suppose that A is invertible. Then the system has a unique solution, namely x=A~lb. If {Aj\ b) denotes the matrix that is obtained from A by replacing the /-th column of A by b, use Theorem 8.9 to prove that the components of x are given by det(>4(;b) X: = del A These equations, known as Cramer's formulae, give the solution to Ax = b in the case where A is invertible. They are of theoretical interest, though somewhat impractical in computing solutions except when n is very small. Try the method on the system f 2jc + 4> - 3z = -1; jc + y - 3z - 2; 3x + 5y + 5z= 3.

152 Basic Linear Aigeora (2) If A is an invertible matrix with integer entries and det A = ± 1, use Cramer's formulae to show that A"1 also has integer entries. (3) In IR2 consider the line pair f ax + by + c = 0; \ a'x + b'y + c' = 0. Show that the lines intersect if and only if the matrix a b^ a' b' is invertible. In this case, use Cramer's formulae to obtain the point of intersection. (4) If a triangle in IR2 has vertices (*,,?,), (^2.^2). (^3.^3) prove that its area is the absolute value of 1^2 det Three lines, given by the system of equations r a,jc + bxy + c, = 0; - a2x + b2y + c2 = 0; ^ a3x + b3y + c3 = 0, bound a triangle. Prove that its area is the absolute value of ±(det A)2 *1 y\ *2 >2 /3 >3 1 1 1 (0,62-^1)(^3-^2)(^1 -albl) where A is the coefficient matrix of the system.

9 Eigenvalues and Eigenvectors Recall that annxn matrix B is similar to an n x n matrix A if there is an invertible nxn matrix P such that B = P~XAP. Our objective now is to determine under what conditions annxn matrix is similar to a diagonal matrix. In so doing we shall draw together all of the notions that have been previously developed. Unless otherwise specified, A will denote annxn matrix over IR or C. Definition By an eigenvalue (or latent root) of A we shall mean a scalar \ for which there exists a non-zero n x 1 matrix x such that Ax = \x. Such a (column) matrix x is called an eigenvector (or latent vector) associated with \. • Note that eigenvectors are by definition non-zero. Theorem 9.1 A scalar \ is an eigenvalue of A if and only if del{A-\In) = 0. Proof Observe that Ax = \x can be written in the form M-V,,)x = o. Then \ is an eigenvalue of A if and only if the homogeneous system of equations (A-\In)x = 0 has a non-zero solution. By Theorems 3.16 and 4.3, this is the case if and only if the matrix A -\In is not invertible, and by Theorem 8.9 this is equivalent to det {A -\In) being zero. D Corollary Similar matrices have the same eigenvalues.

154 Basic Linear Algebra Proof It suffices to observe that, by Theorem 8.7, tel(P~lAP-\ln) = 6el[p-l{A-\In)P) = detP"1 6el{A-\In)6elP = tel(A-\!n). D Note that with A = [fljy]BXII we have det(4-X/B) = det an-\ a2x fl12 fl22 — ^ 02* On«-V an\ an2 and, recalling that the product of the diagonal elements is a term in the ^-expansion, a we see that this is a polynomial of degree n in \. We call this the characteristic polynomial of A. By the characteristic equation of A we mean the equation det(y*-X/B) = 0. Thus Theorem 9.1 can be expressed by saying that the eigenvalues of A are the roots of the characteristic equation. Recall that over the field C of complex numbers this equation has n roots, some

of which may be repeated. If $\lambda_1,\dots,\lambda_k$ are the distinct roots (= eigenvalues) then the characteristic polynomial factorises in the form We call $r_1,\dots, r_k$ the algebraic multiplicities of $\lambda_1,\dots,\lambda_k$. Example 9.1 Consider the matrix We have A = det(4-V2) = det o r -1 0 -\ i -i -\ = \2 + \ Since X2 + 1 has no real roots, we see that A has no real eigenvalues. However, if we regard A as a matrix over C then A has two eigenvalues, namely i and -/', each being of algebraic multiplicity 1.

9. Eigenvalues and Eigenvectors 155 Example 9.2 Consider the matrix A = "-3 1 -7 5 -6 6 -f -1 -2 Using the obvious row/column operations, we compute the characteristic polynomial of A as follows: det(A-X/3) = det = det -3-X 1 -1 -7 5-X -1 -6 6 -2-X -2-X 1 -1 -2-X 5-X -1 0 6 -2-X = -(2 + X)det = -(2 + X)det 1 1 -1 I 5-X -1 0 6 -2-X II -1 0 4-X 0 0 6 -2-X = -(2 + X)(4-X)(-2-X) = (2 + X)2(4-X). It follows that the eigenvalues are 4 (of algebraic multiplicity 1) and -2 (of algebraic multiplicity 2). EXERCISES 9.1 For each of the following matrices, determine the eigenvalues and their algebraic multiplicity: ^2-« 0 i 0 1 + i 0 i 0 2-/ 9.2 If X is an eigenvalue of an invertible matrix A prove that X f 0 and that X"1 is an eigenvalue of A"1. 9.3 Prove that if X is an eigenvalue of A then, for every polynomial p{X), p(X) is an eigenvalue of p(A). "l 0 1 2 2 2 -l" 1 3 > '0 0 1 1 0" 0 1 -3 3 *

156 Basic Linear Algebra If \ is an eigenvalue of A then the set Ex = {x G Matnxl F; Ax = \x} i.e. the set of eigenvectors associated with the eigenvalue \ together with the zero column 0, is readily seen to be a subspace of the vector space Matnxl F. This subspace of eigenvectors is called the eigenspace associated with the eigenvalue \. The dimension of the eigenspace £\ is called the geometric multiplicity of the eigenvalue \. Example 9.3 Consider the matrix A of the previous example. The eigenvalues are 4 and -2. To determine the eigenspace E4 we must solve the system (A -4/3)x = 0, i.e. -7 1 -f -7 1 -1 -6 6-6 The corresponding system of equations reduces lox = 0, y-z = 0 and so E4 is spanned by X y z = 0 0 0 x = where y f 0 since by definition eigenvectors are non-zero. Consequently we see that the eigenspace E4 is of dimension 1 with basis u J As for the eigenspace £_2, we solve (A + 2/3)x = 0, i.e. X y z — "0' 0 0 -1 1 -1 -7 7 -1 -6 6 0 The corresponding system of equations reduces to x = yt z = 0 and so E_2 is spanned by x X= X 0 where x f 0. Thus E_2 is also of dimension 1 with basis 1 1 0 V. LJ

EXERCISES 9.4 For each of the following matrices determine the eigenvalues and a basis of each of the corresponding eigenspaces: "l 0 f 0 1 0 1 0 1) "-2 5 7" 1 0 -1 -1 1 2 9.5 Show that the matrix '-2 -3 -3" -1 0 -I 5 5 6 has only two distinct eigenvalues. Determine a basis of each of the corresponding eigenspaces. The notions of eigenvalue and eigenvector can also be defined for linear mappings. Definition If/ : V —» W is linear then a scalar X is said to be an eigenvalue of/ if there is a non-zero x € V such that /(jc) = Xjc, such an element x being called an eigenvector associated with X. The connection with matrices is as follows. Given annxn matrix A, we can consider the linear mapping fA:M<HnxlF-+M<ilnxlF given by/,4(x) = Ax. It can readily be verified that, relative to the natural ordered basis of Matnx, /\ we have Mat fA = A. Clearly, the matrix A and the linear mapping fA have the same eigenvalues. Example 9.4 Consider the vector space Diff(R, IR) of all real differentiate functions. The differentiation map D : Diff(IR, IR) —» Map(IR, R) is linear. An eigenvector of D is a non-zero differentiable function/such that, for some real X, D/ = X/. By the theory of first-order differential equations we see that the eigenvectors of D are therefore the functions/ given by /(jc) = keXx, where k f 0 since, we recall, eigenvectors are by definition non-zero. Example 9.5 Consider the linear mapping/: IR3 —»IR3 given by f(x,y,z) = {y + z, x + z,x + y).

Relative to the natural ordered basis of IR3 the matrix of/ is The reader can readily verify that "0 1 1 1 0 1 f 1 0 det(v4-X/3) = det -X 1 1 1 -X 1 1 1 -X = -(X + l)2(X-2). The eigenvalues of A, and hence those of/, are therefore 2 and -1, the latter being of algebraic multiplicity 2. EXERCISES 9.6 Determine the eigenvalues, and their algebraic multiplicities, of the linear mapping/: IR3 —»IR3 given by 0) f{*,y,z) = {x + 2y + 2z, 2y + z, -x + 2y + 2z); (2)f(x,y,z) = {y + z,0, x + y). Theorem 9.2 Eigenvectors corresponding to distinct eigenvalues are linearly independent. Proof The proof is by induction. If/ : V — V has only one eigenvalue and if x is a corresponding eigenvector then since jc f 0 we know that {jc} is linearly independent. For the inductive step, suppose that every set of n eigenvectors that correspond to n distinct eigenvalues is linearly independent. Let jc , ,..., jc,,+, be eigenvectors that correspond to distinct eigenvalues X,,..., \n+l. If we have (1) fl,jc, + • • + anx,, + an+lxn+l = 0 then, applying/ and using the fact that/(jc,) = X,*,-, we obtain (2) a,*,*, + • • • + fl^X,,^ + an+l\n+lxn+l = 0. Now take (2) - Xn+I(l) to get 0|(*i ->wi)*i + ■ • ■ + A-^i)^ = 0. By the induction

hypothesis and the fact that X)f...,Xn+1 are distinct, we deduce that It now follows by (1) that an+1xn+1 = 0 whence, since xn+, f 0, we also have an+1 = 0. Hence jc, ,..., xn+1 are linearly independent and the result follows. D

9. Eigenvalues and Eigenvectors __ 159 Definition A linear mapping/: V —» V is said to be diagonalisable if there is an ordered basis (v,),, of V with repect to which the matrix of/ is a diagonal matrix. Thus/ is diagonalisable if and only if there is an ordered basis (v;),, of V such that /(v,) = V, /(v2)= X2v2 /(vj- V» in which case the X, are the eigenvalues of/. We can therefore assert the following: Theorem 9.3 A linear mapping f: V —*Vis diagonalisable if and only if V has a basis consisting of eigenvectors of f. D Equivalently, if we define a square matrix to be diagonalisable when it is similar to a diagonal matrix then as a result on matrices Theorem 9.3 translates into the following result: Theorem 9.4 Annxn matrix is diagonalisable if and only if it admits n linearly independent eigenvectors. D We now proceed to show that /: V —» V is diagonalisable if and only if, for every eigenvalue X, the geometric and algebraic multiplicities of X coincide. For this purpose we require the following results. Theorem 9.5 Let V be of dimension n. If\u...t\k are the eigenvalues of f : V —» V and if dx,..., dk are their geometric multiplicities then rf i + • • • + 4t < n with equality if and only iffis diagonalisable. Proof For each i let £, be a basis of £^. Observe first that if vi + ' • • + vk - ° where each v, € #, then necessarily each v, = 0. This follows from Theorem 9.2. k Observe next that |J Bt is linearly independent. In fact, if Bt = {en,..., eid } i=i and ^-/1,-1^1 + -- + ^/4,^

160 Basic Linear Algebra then 53 Vj = 0 gives, from the above, each v^ = 0 whence all the coefficients ntj = 0. /= i Since the B, are pairwise disjoint, it follows that 4, + .-. + 4 = k >= I <n. Finally, equality occurs if and only if V has n linearly independent eigenvectors, i.e. by Theorem 9.3, if and only if/ is diagonalisable. D Theorem 9.6 If \ is an eigenvalue off : V —» V then the geometric multiplicity of\ is less than or equal to the algebraic multiplicity of\. Proof Let {ex,..., ed} be a basis of Ey and extend this to a basis B = {ex,..., en} of V. The matrix of/ relative to B is of the form '\ld C 0 D A/ = The characteristic polynomial of M is of the form (X - X)dp(X) where p(X) is a polynomial of degree n-d. It follows that d is less than or equal to the algebraic multiplicity of X. D We can now deduce from the above the following necessary and sufficient condition for/: V -» V to be diagonalisable. Theorem

9.7 The following statements are equivalent: (1) $f: V \longrightarrow V$ is diagonalisable; (2) for every eigenvalue λ of f, the geometric multiplicity of λ coincides with the algebraic multiplicity of λ. Proof The sum of the algebraic multiplicities of the eigenvalues is the degree of the characteristic polynomial, namely $n = \dim V$. The result therefore follows from Theorems 9.5 and 9.6. ◻ Example 9.6 As observed in Example 9.2 above, the matrix A = '-3 -7 -6 1 5 6 -f -1 -2

9. Eigenvalues and Eigenvectors 161 has only two distinct eigenvalues, namely 4 and -2. The latter is of algebraic multiplicity 2. To determine the eigenspace $£_2$, we solve $(A + 2/3)x = 0$, i.e. -1 1 -l" -7 7 -1 HS 6 0 X y z = "0" 0 0 The corresponding system of equations reduces to $z = 0$ so the rank of the coefficient matrix is 2 and consequently the solution space is of dimension 3 -2 = I. Thus the eigenvalue -2 is of geometric multiplicity 1. It follows by Theorem 9.7 that A is not diagonalisable. Example 9.7 Consider the matrix fl = The reader can readily verify that $\det(fl-X/3) = (4-X)(X+2)2$, so the eigenvalues are 4 and -2, these being of respective algebraic multiplicities 1 and 2. To determine the eigenspace E_2 we solve $(B + 2/3)x = 0$, i.e. "l 3 6 -3 3" -5 3 HS 4 "3 -3 3" 3-3 3 6-6 6 X y z = "0" 0 0 The corresponding system of equations reduces to $x -y + z = 0$, so the coefficient matrix is of rank 1 and so the dimension of the solution space is 3 -1 = 2. Thus the eigenvalue -2 is of geometric multiplicity 2. As for the eigenvalue 4, since its algebraic multiplicity is 1, it follows by Theorem 9.6 that its geometric multiplicity is also 1. It now follows by Theorem 9.7 that B is diagonal isable. If A is similar to a diagonal matrix D then there is an invertible matrix P such that $P^{-1}AP = D$ where the diagonal entries of D are the eigenvalues of A. We shall now consider the problem of determining such a matrix P.

162 Basic Linear Algebra First we observe that the equation $P^{-1}AP = D$ can be written $AP = PD$. Let the columns of P be $p_,,..., p_n$ and let where $X_,,..., X_n$ are the eigenvalues of A. Comparing the i-th columns of each side of the equation $AP = PD$, we obtain ('-1 n) $Ap^\wedge X_iP_i$. In other words, the i-th column of P is an eigenvector of A corresponding to the eigenvalue $X_,$. Example 9.8 Consider again the previous example. Any two linearly independent eigenvectors in $£_2$ constitute a basis for $£_2$. For these we can choose, for example, "l" 1 0) f 0 -1 Any single non-zero vector in $£4$ constitutes a basis for $£4$. We can choose, for example, T Clearly, the three eigenvectors 1 1 0) 1 0 -1) 1 1 2 are linearly independent. Pasting these eigenvectors together, we

obtain the matrix P = "l 1 0 1 f 0 1 -1 2 and this is such that P~lBP = -2 0 0 0-2 0 0 0 4 • Note that, in order to obtain a particular arrangement of the eigenvalues down the diagonal of the matrix D, it suffices to select the same arrangement of the columns of P.

9. Eigenvalues and Eigenvectors 163 EXERCISES 9.7 For each of the matrices A given by [1 0 f 0 1 0 1 0 1 * "-3 -7 19" -2 -1 8 -2 -3 10 -2 5 1 1 0 -1 -I 1 2 * "-4 0 -3" 1 3 1 4-2 3 find a matrix P such that P~lAP is diagonal. Let us now return to the problems of equilibrium-seeking systems and difference equations as outlined in Chapter 2. In each of these, the matrix in question is of size 2 x 2, so we first prove a simple result that will allow us to cut a few corners. Theorem 9.8 If the 2x2 matrix A = a b c d has distinct eigenvalues \,, \2 tnen lt ls diagonalisable. When b f 0, an invertible matrix P such that P~lAP = is the matrix P = X, 0 o x2 b b Xi — a \j—a Proof The first statement is immediate from Theorems 9.2 and 9.4. As for the second statement, we observe that fl-X b det = \2-(a + d)\ + ad-bc c d-\ and so the eigenvalues of A are *i- l2[{a + d) + yJ(a-d)2 + 4bc], X2= I[(a+<0-V(<i-<02+4fcc] Consider the column matrix - [b

164 Basic Linear Algebra in which, by hypothesis, b f 0. We have a bU b ' c d\ [a,-a 6a, cb + d&i -a) -*i b A, -a the final equality resulting from the fact that ^1(^1 -a) -c6 -^(X, -a) = \]-(a+d)\l + ad-bc = 0. Thus X| is an eigenvector associated with A,. Similarly, we can show that b x2 = A? -a is an eigenvector associated with A2. Pasting these eigenvectors together, we obtained the required matrix P. D Example 9.9 Consider the equilibrium-seeking system as described in Chapter 2. The matrix in question is r 1 4 3 L4 1 -l 20 19 20 J The eigenvalues of A are the roots of the equation (i-wg-M-A-o. The reader will easily check that this reduces to (5A-1)(A -1) = 0 so that the eigenvalues are \ and 1. It follows by Theorem 9.8 that A is diagonalis- able, that an eigenvector associated with X, = £ is 120. "20 J L5 4J and that an eigenvector associated with A2 = ' >s 1 i?i We can therefore assert that the matrix P = . 4 j 1 1 -1 15 is invertible and such that Since, as is readily seen, P'lAP = r 16 4 0 0 1 15 -1 1 1

9. Eigenvalues and Eigenvectors 165 we can compute An - P 0 1 P"1. We have An — -L * ~ 16 ~ 16 ~ 16 1 1 -I 15 1 1 -1 15 1 + ^ 0 1 ru i 5" 5' 1 1] [" 1 I -it 1 1 .15(1-4) 15 + ^r. Example 9.10 Consider the Fibonacci sequence (o,),^q defined recursively by a0 = 0, ax = 1, and (Vn>0) fln+2 - aB+i + An- We can

write this as a system of difference equations in the following way: bn+2 = fln+|- This we can represent in the matrix form xn+2 - Axn+l where x = and A = 1 1 1 0 The eigenvalues of A are the solutions of X2 - X -1 = 0, namely ^, = 1(1 + ^), x2-i(i->/5). By Theorem 9.8, A is diagonalisable, and corresponding eigenvectors are Then the matrix P = > = 1 1-1. = " 1 1 -X2 ""^i. P" 1 X2. > 1 >2- 1 = 1 is invertible and such that lAl [> — 0 0" Now clearly -X, -1 X2 1

166 Basic Linear Algebra and so, using the fact that X,X2 = -1, we can compute A- i_f * MP- °lhx' ~r 1 v -x2 -xJLo xjJL x2 i. >J-Al i l] r-xr' -xji -x2 -x,j [*r' mj "\n+l _ \n+l \" _ \" _ 1 A2 Al A2 Al " Ai"Ai [xs-x? X5_l -X""1. Since it is given that bx = a0 = 0 and a, = 1, we can now assert that 1 _ | A2 A| 0 -*i-*i \n2-\« '«+1 and hence we see that * = sVW -*:>« *H(i+^)1" -#!<» - Vs)]"- Example 9.11 Consider the sequence of fractions 2, 2+i 2 + 1 2+J' 2 + 1 2 + 1 2+J This is a particular example of what is called a continued fraction. If we denote the n-th term in this sqeuence by -^- then we have bn °n+i _ ? , * _ 2fl,, + bn K and so we can consider the difference equations fl«+i = 2fl,, + b,, The matrix of the system is 2 1 1 0 and we can write the system as xn+1 = Axn where Now a, = 2 and bx = 1, so we can compute xn+I from

9. Eigenvalues and Eigenvectors 167 The eigenvalues of A are the solutions of X2 -2X -1 = 0, namely X, = l+v/2, X2=l-√2. By Theorem 9.8, the matrix is invertible and such that Now it is readily seen that 1 1 -l+v/2 -l-√5 1 1 -X2 -X, P~XAP = l+v/2 0 0 1-√2 Consequently, p-i « i *-& = I 2v? 1 + v/2 ll , -1 + v/2 -1 ^ X, 1 -X2 -1 l li rx- oi r x, l -x2 -x,J[o xsJ[-x2 -l l ii r xf' \i -X2 -X,J [-X2+ -x2i rxy+1-xn+1 2 AI A2 \ n _ \ n \n-l _ \ n-1 A ■ nj At Aj and so we deduce from fc::Hf] that a^^ 2[(1 + y/2)n+l -(1- n/2P'] + (1 + y/lf - (1 - y/jf bn+l 2[((1 + n/2)" - (1 - n/2)"} + (1+ n/2)""1 -(1- n/2)""1 (1+^(3 + 2^)-(1-^(3-2^) (1 + √2)"-'(3 + 2n/2) -(1 -y/Z)rl(3-2y/i)' , this can be written Since 1 - √2 = 7= and 3 -2√2 = l+v/2 "*v~ 3 + 2v/2 1 + V2 - (-1)0(1+^)2.-1(3+^)1 from which we see that I 1 (-iy-,(l+%/2)2"-2(3+2v^)2 Um22il=l+V5.

168 Basic Linear Algebra Example 9.12 If ^, is a positive rational, define n -2 + *' Then it is easy to show that P-dl<P-«?|. In other words, if qx is an approximation to \fl then q2 is a better approximation. Starting with qx = 1 and applying this observation repeatedly, we obtain the sequence I 1 1Z 11 ?? 1 V 5' 12" 29' 70 We can use the techniques described above to determine the general term in this sequence and show that it does indeed converge to \fl.

Denoting the n-th term by a —, we have bm Vh . 2+^ K _2bn + a, K and so the sequence can be described by the system of difference equations fl«+i = <*n + 1bn K*\ - on + bn. The matrix of the system is A = 1 2 1 1 and its characteristic equation is X2 -2\ - 1 = 0, so that the eigenvalues are X, = 1+√2, X2=l-√2. By Theorem 9.8, the matrix 2 2 y/2 -v/2 is invertible and such that P'XAP = l+v/2 0 0 1-n/2 Now it is readily seen that />- = * 1 \ii 1 -y/l

9. Eigenvalues and Eigenvectors Consequently, An _ I * ~* 4 = I 4 = 1 4 and so we deduce from that 2 v/2 2HX- °1 -y/5\[0 XJ. 1 \Jl 1 -√2 2 21 |"X; y/2\n 2X^ + 2X5 2√2XY-2√2X; v/2XY-√2X; 2x7 + 2x3 = >T ^, _ 2(i+^r+2(i-^r _ x i+(^)- +1 *,,+i √2(1 + n/2)"+' -n/2(1 -N/2r' " ' 1-(^)^1' from which we see that lim ^ = \fl. n-»oo frit is of course possible for problems such as the above to involve a (non-diagonal) 2x2 matrix A whose eigenvalues are not distinct. In this case A is not diagonalisable; for if X is the only eigenvalue then the system of equations (A - X/2)x = 0 reduces to a single equation and the dimension of the solution space is 2 - 1 = 1, so there cannot exist two linearly independent eigenvectors. To find high powers of A in this case we have to proceed in a different manner. If 'a b^ A = c d then the characteristic polynomial of A is /(*) = X2 -{a + d)X + ad -be. Observe now that A2 = a2 + be b(a + d) c(a + d) bc + d2 a b c d = (a + d) -{ad-be) 1 0 0 1 = {a + d)A-(ad-bc)I2 and so we see that f(A) = 0. For n ^ 2 consider the euclidean division of X" by f{X). Since/ is of degree 2 we have

170 Basic Linear Algebra (3) X"^f(X)q{X) + alX+a2. Substituting A for X in this polynomial identity we obtain, by the above observation, An = axA + a2I2. We can determine a, and a2 as follows. If we differentiate (3) and substitute X (the single eigenvalue of A) for X then, since/(X) = 0, we obtain nX'-'-o,. Also, substituting X for X in (3) and again using/(X) = 0, we obtain Xn=*a,X +a2 = n\n +ot2 and so a2 = {l-n)\n. It now follows that An = n\n-lA + {l-n)\nI2. Example 9.13 Consider the n x n tridiagonal matrix "2 1 0 0 ... 0 0 A = 12 10 0 12 1 0 0 0 0 0 0 0 0 0 0 2 1 1 2 Writing an = det An we have, using a Laplace expansion along the first row, 110 0 ... 0" on = 2fl,,_,-det 0 2 10 0 12 1 0 0 0 0 . 0 . 0 = 2an_i~an_2. Expressing this recurrence relation in the usual way as a system of difference equations K = an-\ we consider the system x,, = v4x,,_, where ." and v4= , _ K 1 0 x =

9. Eigenvalues and Eigenvectors Now det(>i-X/2) = MA-2)+l = (X-l)2, and so

A has the single eigenvalue 1 of algebraic multiplicity 2. We as in the above: [n + 1 —n An = nA + (\-n)l2 = Consequently we have n 1 —n = An~2 n-\ -n + 2 n-2 3-n and hence we see that EXERCISES n+1 n del Am = am = n + 1. 9.8 Consider the tridiagonal matrix A = 1 5 0 0 0 -4 1 5 0 0 0 . -4 . 1 . 0 . 0 . ... 0 .. 0 .. 0 . 1 . 5 0 0 0 -4 1 Prove that deM,,= Jp-^-Hr1]. SUPPLEMENTARY EXERCISES 9.9 Determine the characteristic polynomials of 1 2 3 0 1 2 0 0 1 > 1 1 0 -1 1 1 0 1 -1 1 1 -4 2 -2 [-21 1

172 Basic Linear Algebra 9.10 Suppose that A and B are n x n matrices such that /,, -AB is invertible. Prove that so also is /,, - BA with {ln-BA)-l = In + B{ln-AB)-lA. Deduce that XY and YX have the same eigenvalues. cost? sin t? w—sin i? cost? If p = compute the product P~XAP. 9.12 For each of the following matrices determine an invertible matrix P such that P~lAP is diagonal with the diagonal entries in increasing order of magnitude: 9.11 Show that [Hi] are eigenvectors of A = "0 1 o' 1 0 1 0 1 0) 1" 3 -1 0 -1 0 3 0 0 1 9.13 Consider the system of recurrence relations If the matrix a 0 1 0 has distinct eigenvalues X,, \2 prove that a" a" ",, = 7—4~'''' -~X2"o)-v \ ("i - Vo)- A | ~* a>2 I ~* 2 What is the situation when the eigenvalues are not distinct? 9.14 Determine the n-th power of the matrix "2 1 1 2 2 2 o" 1 1 9.15 Solve the system of equations xn+l = 2xn+6yn >«+i = 6x,, - 3y,, given that jc, = 0 and >, = -1. 9.16 Given the matrix A = find a matrix B such that B2 = A. D2 ^2 Z2 -2± 6i -2i ^2 °2 Z2 0 0 4

9. Eigenvalues and Eigenvectors 173 9.17 Given A G Matnx,,IR let k = max |a,y|. Prove that, for all positive integers r, For every scalar /?, associate with A the infinite series S0(A) = In + PA + p2A2+-+prAr+-- We say that Sp(A) converges if each of the series ^j+0[Alj + P2[A2]u + - .- + ^1,- + - converges. Prove that (1) S0{A) converges if |/?| < —; H nk (2) if Sq(A) converges then /,, -0A has an inverse which is the sum of the series. Deduce that if A is a real nxn matrix and X is an eigenvalue of A then |X|^nmax |a,;|. ASSIGNMENT EXERCISES (1) Determine the characteristic polynomial of the matrix "-4 0 -2 A= 0 1 0 5 1 3 Is A diagonal isable? (2) Determine the characteristic polynomial of the n x n matrix 0 a 0 ... 0 0" 0 0 a ..; 0 0 0 0 0 ... 0 0 0 0 0 ... 0 a Loo o ... o oj (3) If a, b, c, d 6 IR determine the characteristc polynomial of the matrix a b c d~ -b a -d c -c d a -b —d -c b a. [Hint. Look back at quaternions in the assignment exercises for Chapter 2.]

174 Basic Linear Algebra (3) For the matrix A = 2 0 4 3 -4 12 1 -2 5 determine an invertible matrix P such that P~lAP is diagonal. (4) Consider the continued fraction 2 2 JC, JC+ —, JC + X .a » x + JC + x + £ Determine its limit. (5) Solve the system of equations *n+i = txH + (\-t)yH given that jc, = 0 and y, = 1. (6) If v4 G Matnxn IR define the i-th row sum p,(v4) and the j-th column sum Hj(A) respectively by ;=I i=l Prove that if X is an eigenvalue of A then |X| < min{maxpj(i4),max^(i4)}.

10 The Minimum Polynomial In Chapter 9 we introduced the notions of eigenvalue and eigenvector of a matrix or of a linear mapping. There we concentrated our attention on showing the importance of these notions in solving particular problems. Here we shall take a closer algebraic look. We begin by considering again the vector space MatnxlI F which, as we know, has the natural basis {£,; ; ij- 1,..., n) and so is of dimension n2. Thus, recalling Corollary 3 of Theorem 5.8, we have that every set of n2 + 1 elements of Matnxn F must be linearly dependent. In particular, given any A € Mat nxn F, the n2 +1 powers A = In, A, A , A , ... ,A are linearly dependent and so there is a non-zero polynomial p{X) = a0 + axX + a2X2 + ■■- + a^X"2 G F[X] such that p(A) = 0. The same is of course true for any / € Lin (V, V) where V is of dimension n; for, by Theorem 7.2, we have Lin(V, V) ~ MatnxlI F. But we can do better than this: there is in fact a polynomial p{X) which is of degree at most n and such that p(A) = 0. This is the celebrated Cayley-Hamilton Theorem which we shall now establish. Since we choose to work in MatnxlI /\ the proof that we shall give is considered 'elementary*. There are other (much more 'elegant') proofs which use Lin {V, V). Recall that if A € Matnx,, F then the characteristic polynomial of A is cA{\) = 6ei(A-\In) and that cA(\) is of degree n in the indeterminate X. Theorem 10.1 [Cayley-Hamilton] cA{A) = 0. Proof Letfl = A-X/,,and cA(\) = det B = b0 + fc,X + • • • + b,,\n-

176 Basic Linear Algebra Consider the adjugate matrix adj B. By definition, this is an n x n matrix whose entries are polynomials in X of degree at most n - 1, and so we have adj B = B0 + Bx\ + ■ • • + Bn.x\n~l for some nx n matrices B0,..., Bn_x. Recalling from Theorem 8.8 that B ■ adj B = (det B)In, we have (det B)In = B adj B= (A -X/Jadj B = /ladj B-\ adj B, i.e. we have the polynomial identity b0In + b}In\ + •■ + bnIn\n = AB0 + •■• + AB^1 -B0\ Bn.x\n. Equating coefficients of like powers, we obtain Vn = ABo bxIn = ABx- B0 K-\K = ABn-l ~Bn-2 Multiplying the first equation on the left by A0 = /,,,

the second by A, the third by A2, and so on, we obtain b0In = AB0 bxA = A2BX -AB0 b^A"-1 = AnBn.x -An-lBn_2 M" = -AnBn.x. Adding these equations together, we obtain cA(A) - 0. D The Cayley-Hamilton Theorem is really quite remarkable, it being far from obvious that annxn matrix should satisfy a polynomial equation of degree n. Suppose now that k is the lowest degree for which a polynomial $p(X)$ exists such that $p(A) = 0$. Dividing $p(X)$ by its leading coefficient, we obtain a monic polynomial m{X} of degree k which has A as a zero. Suppose that m'{X} is another monic polynomial of degree k such that m'(A) = 0. Then m(X) - m'{X} is a nonzero polynomial of degree less than k which has A as a zero. This contradicts the above assumption on k. Consequently, m(X) is the unique monic polynomial of least degree having A as a zero. This leads to the following: Definition If A e Mat nxnF then the minimum polynomial of A is the monic polynomial mA(X) of least degree such that mA{A) = 0. Theorem 10.2 If $p(X)$ is a polynomial such that $p(A) = 0$ then the minimum polynomial mA{X} divides $p(X)$.

10. The Minimum Polynomial 177 Proof By euclidean division, there are polynomials q(X), r(X) such that $p(X) = mA(X)q(X) + r(X)$ with r[X] = 0 or deg r(X) < deg mA[X]. Now by hypothesis $p(A) = 0$, and by definition mA(A) = 0. Consequently, we have r(A) = 0. By the definition of mA(X) we cannot then have deg r(X) < deg mA(X), and so we must have r(X) = 0. It follows that p{X} = mA(X)q(X) and so mA(X) divides p{X}. D Corollary mA(X) divides cA{X}. D It is immediate from the above Corollary that every zero of mA(X) is a zero of cA(X). The converse is also true: Theorem 10.3 mA(X) and cA(X) have the same zeros. Proof Observe that if X is a zero of cA[X] then X is an eigenvalue of A and so there is a non-zero x G MatBX, F such that Ax = Xx. Given any h{X} =fl0 + fl,X + ... +akXk we then have h[A]x = a0x + a} Ax + • = a0x + a,Xx + ■ = h(\)x ■ + akAkx • + ak\kx whence h(\) is an eigenvalue of h(A). Thus h(\) is a zero of ch^(X). Now take h(X) to be mA(X). Then for every zero X of cA(X) we have that mA(\) is a zero of cmA(A)(X) = ^o(^) = det(-X/J = (-l)"Xfl. Since the only zeros of this are 0, we have mA(\) = 0 so X is a zero of mA(X). D Example 10.1 The characteristic polynomial of A = is cA{X} = (X -2)3. Since A -213 f 0 and (A -2/3)2 ^ 0, we have mA(X) = cA(X). 1 0 -1 0 2 0 f 1 3

178 Basic Linear Algebra Example 10.2 For the matrix A = 5 -6 -6 -14 2 3-6-4 we have cA(X) = (X - 1)(X -2)2. By Theorem 10.3, the minimum polynomial

is therefore either $(X - 1)(X - 2)2$ or $(X -l)(X- 2)$. Since $(A - I3)\{A - 2/3) = 0$, it follows that $mA(X) = (X - \backslash)(X - 2)$. Theorem 10.4 A square matrix is invertible if and only if the constant term in its characteristic polynomial is not zero. Proof If A is invertible then, by Theorem 9.1,0 is not an eigenvalue of A, and therefore 0 is not a zero of the characteristic polynomial. The constant term in the characteristic polynomial $cA(\backslash)$ is then non-zero. Conversely, suppose that the constant term of $cA(\backslash)$ is non-zero. By Cayley- Hamilton we have $cA\{A) = 0$ which, by the hypothesis, can be written in the form $Ap(A) = In$ for some polynomial p. Hence A is invertible. D Example 10.3 The matrix A = "l 0 0 1 1 0 f 1 1 is such that $cA\{X) = (X - I)3$. Thus, applying the Cayley-Hamilton Theorem, we have that $0 = (A - /3)3 -A3- 3A2 + 3A - /3$ which gives $A(A2 -3A + 3/3) = /3$ whence we see that $A\sim x = A2 - 3A + 3/, = 1 -1 0 0 1 -1 0 0 1$ EXERCISES 10.1 If A G MatnxlIF is invertible and $\deg mA(X) = p$, prove that $A\sim l$ is a linear combination of $/_{,,}$, At A2,..., Ap~l.

10. The Minimum Polynomial 179 10.2 Determine the characteristic and minimum polynomials of each of the following matrices: "l 2 3" 0 1 2 0 0 1 * " 1 1 0" -1 1 1 0 1 -1 > "0 0 2 1 0 -1 0 1 1 10.3 Prove that the constant term in the characteristic polynomial of A is det A. 10.4 Determine the minimum polynomial of the rotation matrix cos t? sin t? ' [-sintf cost? Show that if t? is not an integer multiple of tt then R$ has no real eigenvalues. The notion of characteristic polynomial can be defined for a linear mapping as follows. Given a vector space V of dimension n over F and a linear mapping/: $V \longrightarrow V$, let A be the matrix of/ relative to some fixed ordered basis of V. Then the matrix of/ relative to any other ordered basis is of the form $P\sim l AP$ where P is the transition matrix from the new basis to the old basis (recall Theorem 7.5). Now the characteristic polynomial of $P\sim l AP$ is $\det (P-»AP - X/n) = \det [P"1 (A - \backslash ln)P] = detP-! \det(A-X/_{,,})detP = \det(A-X/n)$, i.e. we have $Cp-lAP(\backslash) = CA(\backslash)$. It follows that the characteristic polynomial is independent of the choice of basis, so we can define the characteristic polynomial of/ to be the characteristic polynomial of any matrix that represents/. A similar definition applies to the notion of the minimum polynomial of a linear mapping, namely as the minimum polynomial of any matrix that represents the mapping. EXERCISES 10.5 Determine the characteristic and minimum polynomial of the differentiation map D : JRn[X] -► Rn[X].

180 Basic Linear Algebra 10.6 Determine the minimum polynomial of the

linear mapping/: IR2 -» F2 given by f{x,y)= {x+4y, \x-y). The full significance of the minimum polynomial is something that we shall not develop here. The interested reader may care to consult more advanced texts (see, for example Further Linear Algebra in the SUMS series) that deal with the notions of invariant subspace and direct sum of subspaces. To whet the appetite, let us say simply that if the minimum polynomial of/ is where the pt{X) are distinct irreducible polynomials then each of the subspaces V, = Ker[ft,{/)]* is/- invariant and bases of the V, can be pasted together to form a basis of V. The matrix of/ is then of the block diagonal form A2 This fundamental result is known as the Primary Decomposition Theorem. In the case where each Pj(X) is linear (i.e. all of the eigenvalues of/ lie in F), the matrix of/ is a Jordan matrix. In this, each of the At is a Jordan block, i.e. is of the form r/, h in which each J, is an elementary Jordan matrix, i.e is of the form TX 1 X 1 X 1 X 1 X, If the characteristic and minimum polynomials are cf(x)=n(* - iii1', «fW=n<* - w i=l i=l then in the Jordan form the eigenvalue \, appears dt times in the diagonal, and the number of elementary Jordan matrices associated with X(- is the geometric multiplicity of X,-, with at least one being of size e{ x et.

10. The Minimum Polynomial 181_ As a particular case of this, we obtain another solution to the diagonalisability problem: Theorem 10.5 If Visa non- zero finite-dimensional vector space over afield F then a linear mapping f: V —» V [respectively, a square matrix over F) is diagonalisable if and only if its minimum polynomial is a product of distinct linear factors. D Example 10.4 The matrix of Example 10.2 is diagonalisable. EXERCISES 10.7 Consider the linear mapping/: IR3 —»IR3 given by f(*,y, z) - (* + z, 2y + z, -x + 3z). Prove that/ is not diagonalisable. SUPPLEMENTARY EXERCISES 10.8 Let A, B be square matrices over C and suppose that there exist rectangular matrices P, Q over C such that A = PQ and B = QP. If h(X) is any polynomial with complex coefficients, prove that Ah{A)=Ph{B)Q. Hence show that AmB(A) = 0 = BmA(B). Deduce that one of the following holds: mA(X) = mB{X)t mA{X) = XmB{X)t mB(X) = XmA(X). 10.9 Express the r x r matrix 'I 1 ... 11 2 2 ... 2 J r ... r_ as the product of a column matrix and a row matrix. Hence find its minimum polynomial. 10.10 Let/: ^[X] -» C2[*l be linear and such that /(1)= -1+2X2 /(1+X) = 2 + 2X + 3X2 /(1+X-X2) = 2 + 2X+4X2. Find the eigenvalues and the minimum polynomial of/.

182 Basic Linear Algebra ASSIGNMENT EXERCISES (1) Determine the

minimum polynomial of the matrix cos2t? sin t?cost? sin t? cos # sin2t? (2) Let V be a vector space with dim $V > 2$. If/: V —» V is linear and of rank 1 prove that the minimum polynomial of/ is of the form $X2 - atX$ for some scalar a. (3) Let V be a vector space of dimension n over C and let/: V —» V be a linear mapping that is represented, relative to some ordered basis of V, by the matrix 0 0 ... 0 a, 0 0 0 a,,-i a,, 0 a2 0 0 0 0 0 Determine the minimum polynomial of/2. Deduce that/is diagonal isable if and only if, for each k, otk = 0 a n+l -* = o.

11 Computer Assistance Many applications of linear algebra require careful, and sometimes rather tedious, calculations by hand. As the reader will be aware, these can often be subject to error. The use of a computer is therefore called for. As far as computation in algebra is concerned, there are several packages that have been developed specifically for this purpose. In this chapter we give a brief introduction, by way of a tutorial, to the package 'LinearAlgebra' in MAPLE 7. Having mastered the techniques, the reader may freely check some of the answers to previous questions!

184 Basic Linear Algebra Having opened MAPLE, begin with the following input: > with(LinearAlgebra) : (1) Matrices There are several different ways to input a matrix. Here is the first, which merely gives the matrix as a list of its rows (the matrix palette may also be used to do this). At each stage the MAPLE output is generated immediately following the semi-colon on pressing the ENTER key. For example, we can input the matrices from Exercise 1.12 as follows: for the first, we do input: > ml:=Matrix([[3,1,-2],[2,-2,0],[-1,1,2]]); output: ml := 3 1 -2 2-2 0 -1 1 2 In order to illustrate how to do matrix algebra with MAPLE, let us input the second matrix of Exercise 1.12: > m2:=Matrix([[1,1,1],[1,-1,1],[0,1,2]]); ml := "1 1 0 1 -1 1 r i 2 One way of adding these matrices is by using the 'Add' command: > m3:=Add(ml,m2); m3 := 4 2-1 3 -3 1 -12 4 As for multiplying matrices, this can be achieved by using the 'Multiply* command. To multiply the above matrices, for example, input: > m4:=Multiply(ml,m2); mA := Now 'Add* also allows linear combinations to be computed. Here, for example, is how to obtain 3m 1 + 4m2: "4 0 0] 0 4 0 0 0 4

11. Computer Assistance 185 > Add(ml,m2,3,4); "13 7-2" 10 -10 4 -3 7 14 (2) A simpler method An more convenient way to input commands is to use

algebraic operations > ml+m2; "4 2-1" 3 -3 1 -1 2 4_ Multiplication by scalars is obtained by using a '*': > 3*ml+4*m2; "13 7-2" 10 -10 4 -3 7 14 Multiplication of matrices is obtained by using a V: > ml.m2; "4 0 0" 0 4 0 0 0 4 As for a more complicated expression: > ml.(4*m2-5*ml*2); "-219 -25 230" -110 86 60 85 -25 -74 m (3) Inverses Inverses of square matrices can be achieved by using the 'MatrixInverse' command: > MatrixInverse(ml); r i lin Alternatively we can input the following (here it is necessary to insert brackets round the -1):

186 Basic Linear Algebra > ml^(-l); Note that we can use negative powers in products: > ml"(-3).m2rt3; 23 65 25 32 64 8 i. 51. 55 16 64 32 45 65 201 64 64 64 J (4) Determinants To compute the determinant of a square matrix, use the 'Determinant' command: > Determinant(ml); -16 Of course the determinant of a product is the product of the determinants: > Determinant(ml.m2); 64 > Determinant(ml)*Determinant(m2); 64 (5) More on defining matrices We now look at other ways of defining a matrix. We start with a clean sheet (to remove all previous definitions): > restart; > with(LinearAlgebra): We can enter a matrix as a row of columns: > M0:= «a,b,ol<d,e, f>|<g,h,i»; MO := a d g b eh cfi or as a column of rows (which can also be done using the matrix palette):

11. Computer Assistance 187 > Ml := «a I b I c> , <d I e I f> , <g I h I i>>; Ml := a b c d ef gh i Then, for example, we have > Ml"2; a2 + bd + eg ab + be + ch ac + bf + ci da + ed+fg bd + e2+fh dc + ef + fi ga + hd + ig gb + hc + ih cg+fh + i2 > Determinant(Ml); aei - afh + dch - dbi + gbf - gee Particular types of matrix can be dealt with. For example, consider the upper triangular matrix given in Exercise 1.33. > A:=Matrix(4,[[0,a,a*2,a*3],[0f0fafa*2]f[0f0f0fa]f > [0,0,0,0]], shape=triangular[upper]); A := "0 a a2 a3 0 0 a a2 0 0 0 0a 0 0 0 0 r3,A"4) ; OOfl2 2a3' 0 0 0a2 0 0 0 0 0 0 0 * "0 0 0 a3' 0 0 0 0 0 0 0 0 0 0 0 i "0000 0 0 0 0 0 0 0 0 0 0 0 Consider now the matrix B. Since only the first three powers of A are non-zero, we have: > B := A-(1/2)*A"2+(1/3)*A"3; B:= 0 a 0 0 0 0 0 a 0 0 I*3 a 0

188 Basic Linear Algebra > print(B*2,B*3,B*4); OOfl2 a3] OOOo2 0 0 0 0 0 0 0 0 • "0 0 0 a3! 0 0 0 0 0 0 0 0 0 0 0 0 • ro o o o 0 0 0 0 0 0 0 0 0 0 0 0 0 We now check that the sum of the series for B gives the matrix A: > B+(1/2)*B"2+ (1/6)*B"3; 0 a a2 a3' 0 0 a a2 000 a 0 0 0 0 To input symmetric matrices we can proceed, for example, as follows: > M:=Matrix(3,3,shape=symmetric);

M:= then input, for example, >M[1,1]:=2; M[l,3]:=23; M[2,3]:=Pi; "0 0 0" 0 0 0 0 0 0 A/,., A/ M = 2 = 23 = 7T > M; > Determinant(M); 2 0 23 0 0 7T 23 7T 0 -2tt2 We now take a look at Exercise 1.26. > A:=Matrix(3,3,shape=antisymmetric);

11. Computer Assistance 189 A := 0 0 0 0 0 0 0 0 0 > A[l,2]:=-c; A[l,3]:=b; A[2,3]:=-a; > A; ^1,2 ^1,3 ^2,3 0 C -b := := := -c 0 a -c b -o b -a 0 > AA3+A; 0 -(-c2 - b2)c +ca2-c {-c2 - b2)b -ba2 + b 0 b2a - {-c2 - a2)a - a 0 (-c2 - a2)c -cb2 + c c2b - {-b2 - a2)b - b -c2a + {-b2 - a2)a + a Now the question gives us: >c:=sqrt(l-a^-b^); c:=Vl-a2-b2 > A~3+A; 0 -(-1 + a2) % 1 + % 1 a2-% 1 (-1 + a2)b-ba2 + b (-1 + b2) % 1 - % \b2 + % 1 0 b2a-{-\ + b2)a-a (l-a2-b2)b-H>2-<*2)b-b -(l-a2-b2)a + &2-a2)a + a 0 %1 := y/l-a2-b2 > simplify(A*3+A); "0 0 0 000 000 (6) Systems of linear equations To illustrate the use of MAPLE in solving systems of linear equations, let us consider Example 3.11. > restart; > with(LinearAlgebra): > A:= «1,1,3,0>l<0,l,l,l>l<l,0,l,2>l<0,0,l,l>l > <1,2,1,2>>;

190 Basic Linear Algebra A := 10 10 1 11002 3 1111 0 12 12 > GaussianElimination(A); 10 10 1 0 1-10 1 0 0-1 1-3 u0 0 0 4-8 Consider now Example 3.16. > B:= «1,2,3>I<2,4,6>I<1,4,5>I<2,8,7>|<1,4,7»; 12 12 1 B := 2 4 4 8 4 w365 77 > GaussianElimination(B); 12 1 2 1 002 42 0 00-3 2 > ReducedRowEchelonForm(B); 1200 0 ooio I 000 1 -f j Next, let us consider the system of equations in Example 3.27. > restart; > with(LinearAlgebra): > sys := [x+y+z=l,2*x-y+2*z=l,x+2*y+z=alpha]: > var := [x,y,z]: > (A,b) := GenerateMatrix(sys,var); "i l r 2-12 1 2 1 • "1" 1 Of Atb:= > GaussianElimination(<Alb>); 1111 0-3 0 -1 0 00a-f Consider separately the cases a f \ and a = \\

11. Computer Assistance 191 > assume(alpha<>4/3); > ReducedRowEchelonForm(<Alb>); '10 10" 0 100 000 1 - > alpha:=4/3: > ReducedRowEchelonForm(<AIb>); " 1 0 1 f' 0 1 0 ± 0000 (7) Writing procedures We now show how to input matrices whose (i, j)-th entry is a function of i and j. > M:=Matrix(6,6,(i,j)->i*j); "12 3 4 5 6" 2 4 6 8 10 12 _ 3 6 9 12 15 18 *"~ 4 8 12 16 20 24 5 10 15 20 25 30 6 12 18 24 30 36. More complicated functions can be produced by writing a procedure. The following example illustrates a very simple procedure in order to define the identity matrix: > f:=proc(i,j) ; > if i=j then 1 else 0 end if; > end proc; / := proc (/,;')

if i = j then 1 else 0 end if end proc > Ml:=Matrix(6,6,f); '10 0 0 0 0" 0 10 0 0 0 0 0 10 0 0 0 0 0 10 0 0 0 0 0 10 0 0 0 0 0 1 Here is a more complicated procedure:

192 Basic Linear Algebra > f:=proc(i,j); > if i>j then x else if i=j then 0 else - y end if > end if; > end proc; / := proc (/ J) if j < i then x else if i = j then 0 else - y end if end if end proc > M2:=Matrix(6,6,f); M2:= > d:=Determinant(M2) ; d := (x4 - x3y + y2x2 - y3x + y*)xy > expand(d*(x+y)); 6 . 6 x y + xy Hence d = (x6y + xy6)/(x + y). Let us examine further cases: > for n from 2 to 8 do > M2:=Matrix(n,n,f); > d:=Determinant(M2); > print(cat(' size ',n(> end do: 0 X X X X X -y 0 X X X X -y -y 0 X X X -y -y -y 0 X X -y -y -y -y -y 0 X -y -y -y -y -y -y -y 0-y x 0 gives'), expand(d*(x+y))); size 2 gives yx2 + y2x size 3 gives -yx3 + y3x size 4 gives yx* + y*x size 5 gives -yx5 + y5x size 6 gives yx6 + y6x size 7 gives -yx1 + y7jt size 8 gives yx9 + y*x The general theorem should be easy to spot. Try to prove it. Another way of inputting matrices uses the command 'BandMatrix' which we illustrate by referring to Exercise 8.34. Consider first the case where n = 6. > M3:=BandMatrix([1,2*cos(x),1],1,6,6);

11. Computer Assistance 193 M3:= 2 cos(jc) 10 0 0 0 1 2 cos(x) 10 0 0 0 1 2 cos(x) 1 0 0 0 0 12 cos(x) 1 0 0 0 0 12 cos(x) 1 0 0 1 2 cos(x) 0 0 > d:=Determinant(M3) ; d := 64 cos(jc)6 - 80 cos(x)4 + 24 cos(jc)2 - 1 > expand (sin(7*x)/sin(x)); 64 cos(x)6 - 80 cos(x)4 + 24 cos(x)2 - 1 The following procedure will verify the identity of Exercise 8.34 for n taking the values from 2 to 10, but may be used for any range of values of n. Try it out! > for n from 2 to 10 do > mat:=BandMatrix([1,2*cos(x), 1], 1, n, n); > d:=Determinant(mat); > dl:=expand(sin((n+1)*x)/sin(x)); > print(cat ('case of ',n,'x',n,' matrix')); > print(cat(' det = '), d); > print(cat(' sin(',n+1,'x)/sin(x) = '), dl); > print(' '); > end do: (8) More on determinants We now consider a generalisation of Exercise 8.28. > restart; > with(LinearAlgebra): > f:=proc(i,j); > if i=j then x else a end if; > end proc; / := proc (/J) if / = j then jc else a end if end proc > m:=Matrix(3,3, f); m := x a a a x a a a x

194 Basic Linear Algebra > d:=Determinant(m); </:=x3-3;ur + 2a3 > factor(d); (2a + x)(a-x)2 > solve(d=0,x); -2a,a,a We can now do this for any range of values of n. For example: > for n from 2 to 12 do > m:=Matrix(n,n,f); > d:=Determinant(m); > print(solve(d=0,x)); > end do: -a,a -2a, a, a -3a, a, a, a -4a,a,a,a,a -5a,a,a,a,a,a -da,a,a,a,a,a,a —la,a,a,a,a,a,a,a —Sa, a, a, a, a, a, a, a,

a -9a,a,a,a,a, a, a, a, a, a -Ma, a, a, a, a, a, a, a, a, a, a, a Consider now the matrix A,, of Exercise 8.33. This we can investigate, first for n = 6, as follows. > f:=proc(i,j); > if i=j then a else if i<j then b else -b end if > end if; > end proc; / := proc(/, j) if / = j then a else if i <) then b else - b end if end if end proc > m:=Matrix(6,6,f); " a b b b b b' -bo b b b b -b-b a b b b b m: -b-b -b a b b -b -b -b-b a b -b-b-b-b -b a

11. Computer Assistance 195 > d:=Determinant(m); d:=a6+\5a4b2+\5b4a2 + b6 The following procedure does the same job for n = 2 to n = 8: > for n from 2 to 8 do > m:=Matrix(n,n,f); > d:=Determinant(m); > print(d); > end do: a2 + b2 a3 + lab2 a4 + 6a2b2 + b4 a5 + \0a3b2 + 5ab4 a6 + \5a4b2 + \5a2b4 + b6 fl7+21fl562 + 35flV + 7flfc6 a8 + 2Za6b2 + 70a4b4 + 2%a2bh + b* Now carry out the following procedure to obtain the same display: > for i from 2 to 8 do > expand((((a+bri+(a-bri)/2); > end do; (9) Matrices with subscripted entries Consider now a matrix of the type given in Example 8.13. > m:=Matrix(3,3,(i,j)->b[i]* (j-1)) ; m := 1 bx b2^ 1 bj bj > d:=Determinant(m); d := bjby —bj b$ + b$b\ —bib$ +6^2 — ^2^1 > factor(d); -(-^ + ^2)^1-^2)(^1- ¾ Of course MAPLE can find the determinant of matrices of this type for much larger dimensions. For example, try the following: > factor(Determinant(Matrix(6,6,(i,j)->b[i]* (j-1))));

196 Basic Linear Algebra (10) Circulant matrices We now show how to define the entries of a matrix by using a function that depends on certain preset global variables. > restart; > with(LinearAlgebra): > a:=5; b:=8; > f:=proc(i,j) global a, b; > if i>j then a else if i=j then a+b else b end if > end if; > end proc; a := 5 b:=% /:=proc(/,j) global a, 6; if) < i then a else if i = j then a + b else b end if end if end proc > m:=Matrix(6,6,f); '13 8 8 8 8 8" 5 13 8 8 8 8 5 5 13 8 8 8 m: 5 5 5 13 8 8 5 5 5 5 13 8 _ 5 5 5 5 5 13. By a circulant matrix we shall mean a square matrix A = [a,y],,xn with the property that akj = 0,+1^+1, the subscripts being reduced modulo n. For example, the matrix P of Exercise 3.1 is circulant. We shall now describe a procedure that defines annxn circulant matrix with three non-zero entries in each row, namely 1,1,-1, whose positions in the first row of the matrix are at columns atbtc respectively. > a:=l: b:=2: c:=3: n:=6: > g:=proc(i,j) global a, b, c, n; > if i mod n = (j-a+1) mod n then 1 else > if i mod n = (j-b+1) mod n then 1 else > if i mod n = (j-c+1) mod n then -1 else 0 end if > end if end if; > end proc;

1 0 0 0 -1 1 1 1 0 0 0 -1 -1 1 1 0 0 0 0 -1 1 1 0 0 0 0 -1 1 1 0 0 0 0 -1 1 1 g := proc(i j) global a, btct n\ if i mod n=(/'-a + l) mod n then 1 else if i mod n = (j -b+\) mod nthen 1 else if i mod n= (j -c + \) mod nthen - 1 else 0end if end if end if end proc Here is the 6 x 6 circulant matrix that this defines: > ml:=Matrix(6,6,g); ml := We now compute the sequence of absolute values of the determinants of the above matrices from size 1 x 1 to 20 x 20. > si: = [seq(abs(Determinant(Matrix(n,n,g))), n=1..20)]; s\ := [1, 0, 4, 5, 11, 16, 29, 45, 76, 121, 199, 320, 521, 841, 1364,2205,3571,5776,9349, 15125] Can this sequence be suitably described? For this purpose, consider the sequence (£,,)n>1 of Lucas numbers defined recursively as follows: £, = 1, £2 = 3, (n£3K = 4,.1+4-2. The Lucas sequence has the same recursive definition as the Fibonacci sequence except for the second term, which is 1 for Fibonacci. We can get MAPLE to remember the Lucas sequence as follows: > lucas:=proc(x) option remember; > if x = 1 then 1 else if x = 2 then 3 else > lucas(x-l)+lucas(x-2) end if end if; > end proc; lucas := proc(jt) option remember; if x = 1 then 1 else if x = 2 then 3 else lucas(x -1)+ lucas (x - 2) end if end if end proc Here then are the first 20 Lucas numbers:

> s2:=[seq(lucas(i),i=l..20)]; $2:=(1,3,4,7, 11, 18,29,47,76, 123, 199,322,521,843, 1364, 2207, 3571, 5778, 9349, 15127] By comparing the sequences s\ and s2, make a guess at the value of the determinant of the above circulant matrix for arbitrary n. Then try to prove your answer. (11) Vector spaces We shall now illustrate how MAPLE can tackle vector space problems. First we compute bases for subspaces A, B spanned by given row vectors, and for A D B. > restart; > with(LinearAlgebra): > vl := <2I2I1|3>; > v2 := <7|5I5I5>; > v3 := <3|2|2|1>; > v4 := <2|1|2I1>; vl v2 v3 v4 > Basis([vl,v2,v3,v4]); [[2,2, 1,3], [7, 5, 5, 5], [3, 2, 2,1]] > xl := <2I2I3|1>; > x2 := <5|7|5|5>; > x3 := <2I3I2|1>; > x4 := <3I5|2|4>; x\ := [2, 2, 3, 1] *2 := [5, 7, 5, 5] x3 := [2, 3, 2, 1] x4 := [3, 5, 2, 4] > Basis([xl,x2,x3,x4]); [[2,2,3, 1], [5,7, 5, 5], [2, 3, 2,1]] > IntersectionBasis([[vl,v2,v3,v4], [xl,x2,x3,x4]]); r[_l I _2 _I1 [3 5 I 231, 1 I 4» 4» 8» 8J » U» 4' 8» 8 * :=(2,2,1,3] := [7, 5, 5, 5] = (3, 2, 2, 1] = [2,1,2,1]

Now let us consider the matrix of a linear mapping. We start with an example in which we determine the matrix of a linear mapping relative to a 3-element basis. The method follows that of hand calculation, in that we obtain the linear equations that express the images of the

basis vectors as linear combinations of the basis vectors and construct a matrix whose columns are the coefficients in each of the equations. The example we use is the mapping and the basis of Exercise 7.5. > restart; > with(LinearAlgebra): > f:=(x,y,z) -> <z,y,x>; > bl:=l,2,0 > b2:=2,l,0 > b3:=0,2,l > B:=«bl>l<b2>l<b3»: > Ml:= <LinearSolve(B,f(bl))ILinearSolve(B,f(b2)) I > LinearSolve(B,f(b3))>; /:=(*, y, z) -»(z, y, x) "0-: Ml := "0 0 1 -2 1' 1 0 20 Of course, in this example we have/ o/ = id and so the square of its matrix should be the identity matrix: > Ml"2; "1 0 01 0 1 0 0 0 1 As for the mapping g of Exercise 7.5, we compute its matrix similarly: > g:=(x,y,z) -> <2*x,y+z,-x>; > bl:=l,2,0: > b2:=2,1,0: > b3:=0,2,1: > B:=«bl>l<b2>l<b3»: > M2:= <LinearSolve(B,g(bl))ILinearSolve(B,g(b2)) I > LinearSolve(B,g(b3))>; 8-= (x,y,z)->(2x,y+zt-x) " 2 2 2 Ml := 0 1 -1 -1 -2 0 The composite mapping/ o g o/ then has matrix A/I -A/2 -A/1:

200 Basic Linear Algebra > M1.M2.M1; 2 2-1 -1 -1 0 0 0 2 Here now is an example in which we verify that, relative to a given ordered basis, the matrix of/2 is the square of the matrix of/. > f: = (x,y,z,w) -> <2*x-y+2*z-3*w,3*x-7*y+z-w,-x+y-z+5*w,x+y-z-w>; bl:=0,1,1,0: b2:=l,0,1,0: b3:=l,1,0,0: b4:=l,0,0,1: B: =«bl> I <b2> I <b3> I <b4»: > M:= <LinearSolve(B,f(bl))ILinearSolve(B,f(b2)) I > LinearSolve(B,f(b3))ILinearSolve(B,f(b4))>; f'-= (x, y, z, w) —* (2x-y+2z-3w,3x-7y+ z-w, -x + y-z + 5w,x + y-z-w) M := 71 -f-l-f ^ "I 5 -f -| 0 0 2 0 We now consider the composite/o/. > fl:=(x,y,z,w) -> <f(f(x,y,z,w))>: > bl:=0,1,1,0: > b2:=l,0,1,0: > b3:=l,1,0,0: > b4:=l,0,0,l: > B:=«bl>l<b2>l<b3>l<b4»: > Ml:= <LinearSolve(B,fl(bl))ILinearSolve(B,fl(b2)) I > LinearSolve(B,fl(b3))ILinearSolve(B,fl(b4))>; Ml := 25 _i 29 _211 2 J 2 2 -T 5-T T 32 2 65 2 -5 12 2 29 B2 —IS £L—5 2 1J 2 2 10 -5 -3 We invite the reader to use MAPLE to check that A/1 = A/2.

11. Computer Assistance 201 (12) Eigenvalues and eigenvectors We now illustrate how to find eigenvalues and eigenvectors. The following is self-explanatory. > a:=Matrix<[[-2,-3,-3],[-1,0,-1], [5,5,6]]); -2 -3 -3 a := > e:=Eigenvalues(a) ; e := -1 0-1 5 5 6 > print(e[l],e[2],e[3]); 2,1,1 > > > > ch:=CharacteristicPolynomial(a,X); ch :=5X-AX2-2+X3 factor(ch); (X -2)(X-\)2 solve(ch=0,X); 2,1,1 evl:=Eigenvectors(a); ev\ := 2 1 1 • 3-1 -1 1 1 0 -5 0

1 In this display the column vector gives the eigenvalues of a and the matrix has eigenvectors as its columns. Although MAPLE always finds correct eigenvectors it may find different ones from the same code. For example, consider the following input which provides more information:

202 Basic Linear Algebra > ev2:=Eigenvectors(a, output='list'); 2.1.- r r -i . 1.2.- r "-1" 1 0 I "-1" 0 1 > ▶ ev2 := Here MAPLE returns a list of lists, each of the form [eitmh {v[l, i),..., v[nh i]}] where el is an eigenvalue, tnt is its algebraic multiplicity, and {v[l, i],..., v[n,-, i)} gives nt linearly independent eigenvectors where n, is the geometric multiplicity. Thus, for example, the eigenvalue 1 has algebraic multiplicity 2 with "-I" 1 0 . "-1" 0 1 as two linearly independent eigenvectors. The corresponding eigenspaces can be determined by using the 'Nullspace' command: > id:=IdentityMatrix(3): > kl:=NullSpace(a-l*id); *1:= > k2:=NullSpace(a-2*id); *2:= r < r -i 0 . " 0" -i 1 > > 3 1 -5 > p:=Matrix([kl[1],kl[2],k2[1]]); P-= 0 1 3 -1 -1 1 1 0 -5 We can now verify that p reduces the matrix a to diagonal form: > p^(-l).a.p; 1 00 0 1 0 00 2 Finally, we can illustrate Theorem 10.3 for a particular 5x5 matrix

11. Computer Assistance 203 > A:=Matrix([[5,-1,-3,2,-5],[0,2,0,0,0],[1, 0,1,1,-2], > [0,-1,0,3,1],[1,-1,-1,1,1]]); A := 5 -1 -3 2 -5 0 2 000 10 11-2 0-1 0 3 1 1-1-11 1 > factor(CharacteristicPolynomial(A,X)); (-3 + X)2(X-2)3 > factor(MinimalPolynomial(A,X)); (X-2)2(-3 + X)2 EXERCISES 11.1 In the description of a circulant matrix, try changing the values of a, bt c and compute the corresponding determinants. 11.2 Find the eigenvalues and corresponding eigenvectors of 1-2 3-4 -5 6-7 8 9-10 11 -12 -13 14-15 16 11.3 Find the determinant and the inverse of the matrix 12 3 4 5 6 11.4 Consider the matrix A = 2 : J 4 5 6 1 3 4 5 6 12 4 5 6 12 3 5 6 12 3 4 6 12 3 4 5 "1 2 0 x 12 3 3 10 11 x 1 1 3 Determine the values of x for which A is invertible.

204 Basic Linear Algebra 11.5 In the diagram 205 204 20 206 193 208 200 192 4 207 k 6 196 400 each column and each row and each main diagonal sums to 1000. Complete the square with positive integers and find what range of values may replace k. ASSIGNMENT EXERCISES (1) Consider the n x n version An of the matrix in Exercise 11.3. Devise a procedure similar to that employed for circulant matrices to determine the sequence of values of det An for n = 1 to n = 10. By observing that, for each of these, det An has a factor \n2{n + 1), guess the general form of A~l. (2) Write a procedure which, for a

given positive integer n, returns the n x n matrix A whose (i, /)-th entry is -—: . i+j-1 Incorporate your procedure in a program which, for n = 1 to n = 6, carries out the following calculation: It finds c = Ab where b is the n x 1 matrix all of whose entries are 1. It then calculates the matrix X which is defined as A except that the entries are 10-digit floating point numbers rather than rational numbers. Finally, it calculates x = X_,c and displays the answer x. Comment on the output of the program.

12 Solutions to the Exercises 1.1 1.5 "2 3 4" 3 4 5 4 5 6 . 1.2 "i o r 0 1 0 1 0 1 . u 2 3 2 6 6 3 4 12 10 15 20 6 6 1 -1 -1 1 1 -1 f -1 1 . 1.4 0 -1 -1 -1 1 0 -1 -1 1 . 1 . 0 . -1 . 1 ■ . 0 4 5 4 10 12 15 4 20 20 5 12 30 6' 6 6 12 30 6 • 111111 12 12 12 3 1 1 1 4 1 11115 12 3 2 1 1 1 1 2 1.6 a\n ... an an a2,, ... a-n a2\ oM ... a,,2 a,,\ 1.7 Real numbers are 1 x 1 matrices. 1.8 Theorem 1.3. 1.10 X = J/\ + %B. UIX = A + 2B = 12 3 4 2 4 6 8 3 6 9 12 4 8 12 16 '•» [:: — y x -~z y — w y x —z y — w 1.13 "3 2 r 2 2 1 1 1 1 . 1.14 [= "3 2 2" 2 3 2 2 2 3 . 1.12 . [30]. 1.15 Each is "4 0 0' 0 4 0 0 0 4 • '3 45 9" 11 -11 17 7 17 13 1.16 The product is the 1 x 1 matrix [t] where t= ax2+ Ihxy + by1 + 2gx + 2fy + c. Each of = [0] with, respectively, M the matrices 0 0 -2a 0 1 0 -2a 0 0 the equations can be written in the form [x y \]M P f 'I 2 i 5 2 J 2 L4 f 2. 1.17 A2 = • 0 £ 0 .0 0 -1. "0 0 a2' 0 0 0 0 0 * "0 J 0 " \ 0 0 0 0 -a2 x~ y i * , A3 = 0.

206 Basic Linear Algebra 1.18 We have/IB = 0,A2 = A,B2 = -B, and (A+B)2 - I2. On the other hand, A2+2AB+B2 = 1 2l . Since {A + B)2 = 12 we have (A + £)3 = (A + B)(A + B)2 = A + B; and since AB = 0. A* = A, By = B we have Ay + 3A2B + 3AB2 + B* = A + B. 1.19 We have ay = 0 and bti - 0 whenever i f j and so [ABlj = £ a,kbkj = 0 if if j. k 1.20 By Exercise 1.19 and induction, if A is diagonal then so is Ap. 1.21 If A and £ commute then a simple inductive argument shows that Am and £ commute for every positive integer m. Fixing m, the same induction shows that B" commutes with Am for every positive integer n. n 1.22 The proof is the same as for the binomial theorem (jc + y)" = Y, 0*Y'" and is by /=0 induction, using properties of the binomial coefficients and the hypothesis that xy = yx. 1.23 [A(B + O]' = [AB\ + [AC]' = [B'A'] + [CA'\. 1.24 The result is trivial for n = 1. Suppose it holds for n. Then (An+l)' - (A"A)' = A'{A")' = A'{A'f = (/\')"+l. 1.25 If AB = BA then A'B' = (£/*)' = (>\fl)' = B'A'. a2 —I ab ac ab b2-\ be ac be c2 — 1 on the left by A then gives A1 = —A. Finally, A* = —A2 where d2 is as above. 1.27 If A is both symmetric and skew-symmetric then ai} = ait and al} — -a,, whence fly = 0. US A' = A and £'

= -B. Thus (/IB + BA)' = #/1' + /l'# = -BA -AB= -(AB + BA) and so AB+BA is skew-symmetric. Similarly, AB-BA is symmetric. Next, (/I2)' = (/I')2 = A2 so i42 is symmetric. Similarly, B2 is symmetric. Finally, {AP&AP)' = (A'WfWK = /*'(-£)*/*' and so APB9AP is symmetric if a is even, and skew-symmetric if a is odd 1.26 Using a2 + fc2 + c2 = I we have A2 = = X'X—I2. Multiplying 1.29 A' = W -yx1)' = yx' -x/ = -/*. If x = V *«. andy = >r jy then x'y = £>,>>, = y1* 1=1 If now x*x = y'y = [1] and x*y = y*x = [it] then A2 = (xy7 -yx'Xxy' -yx') = xyV -yxV -xy'yx' + yx'yx' = Jtxy1 -yy* -xx' + Jfcyx' and hence A> = (Jtxy1 -yy'-xx' + Jfcyx')(xy'- yx') = kxyxy1 - yy'xy' - xx'xy' + fcyx'xy' - fa/yx' + y^yx' + xx'yx* - Jfcyx'yx* = *2xy'-ltyy'-xy' + ltyy'-Jtxx' + yx' + itxx'-it2yx' = it2(xy'-yx')-xy' + yx' = (it2 -1)/1. 130 [AB - BA]U = al2b2l - a2lbl2 and [AB - BA]22 = a2lbi2 - ax2b2\ so the sum of the a b , , \a2 + bc 0 then E2 = „ , c -a 0 a2 + to- diagonal elements is 0. If E = = (fl2 + bc)I2.

12. Solutions to the Exercises 207 For the last part, observe from the above that (AB -BA)2 = \l2 and therefore commutes with every C. a b] , . ThenX2 = /2ifandonlyif c d 131 LetX = a2 + bc=\, b(a + d) = 0, c(a + d) = 0, cb + d2=\. Suppose that b = 0. Then these equations reduce to a2 = 1. c(a + d) = 0, d2 = 1 from which we see that either a — d— \.c= O',ora = d = —1, c = 0\ora = \,d= —\,c arbitrary; or a = -\.d= \,c arbitrary. If b ^ 0 then we must have a + d = 0 whence d = -a and c = (1 - a2)/b. Thus the possibilities for X are 1 0 0 1) -1 0 0 -1) 1 0 c -1) -1 0 c 1 * a b ,21 1.32 If (x, y) lies on the curve)-2 - x2 = 1 then A - x y -y -x (\-a2)/b -a is such that A2 = -/2, so there are infinitely many such matrices. 1.33 Simple calculations reveal that AA = 0 whence A" = 0 for n ^ 4 and so B = A-LA2 + lA* = Likewise, B* = 0 and so B" = 0 for n > 4. Thus "0 a {a2 0 0a 0 0 0 0 0 0 a a2 ay 0 0 a a2 0 0 0a 0 0 0 0 W W a 0 ^ B+ ±B2+±B> = 134 The result follows from the standard formulae cos(i? + <p) = cos ■d cos yj — sin 1? sin <p and sin(t? + <p) = sin t?cos ip + cos t?sin p. 135 For the inductive step, use the previous exercise: A"*1 = A" A = cos ifd sin m? —sinm? cosnd cos t? sin t? -sint? cost? cos(n+1)t? sin(/i+1)t? -sin(n + 1)t? cos(n+ l)i? 1.36 Let B,, be the matrix on the right. Then clearly B\ = A so the result is true for /1 = 1. For the inductive step, observe that B,,A = B,,+\. 1.37 pBJC] = (AB - BA)C - C(AB - BA) = ABC + CBA - BAC - CAB; and similarly [[BC]A] = BCA + ACB - CBA - ABC, [[CA]B] = CAB + BAC - ACB - BCA. The first result follows by adding these expressions together. As for the second, we have [(A + B)C] = (A + B)C - C(A + B) = AC -CA + BC - CB = [AC] + [BC]. The third result follows by expanding as in the first. Take

A = B = 0 1 f 0 X = 1 0 0' 0 Then [[AB]C] = 0, [A[BC]] = 2 0 0 -2 1.38 Substitute for y in the expression for x and compare with the expression for x in terms of z. We have C\ = a\b\ + a2bit etc.

208 Basic Linear Algebra 2.1 If A is orthogonal then AA' = /,, = A'A. Since A = (A')'t we then have A'(A')' = I,, = (A')'A' so that /T is orthogonal. 2.2 (AB)'AB = ffA'AB = B!lnB = l,, and similarly AB(AB)' = /,,. 2.3 See Example 4.3 of the text 2.4 Let the top sheet be the (x, y)-plane and the bottom sheet the (jc', /)-plane. If i? is the angle of anti-clockwise rotation of the top sheet, we have 5. 12 LI3J = /?- cost? —sini? sin i? cos i? cosi? sin i? and so cos i? = ^ and sin t? = {3. The point (jc',/) above which the point (2,3) lies is then T2l given by 5. _I2 13 13 il L 13 13 .i.e. (*',/)= (-2,3). 2.5 Clearly jc' = jc and / = -y, so 1 0 0 -1 The other matrix is -1 0 0 1 2.6 To obtain (xL,yL), rotate the axes through 1?, take the reflection in the new jc-axis, then rotate through -1?. The matrix in question is R-^MRt where M is the matrix of the previous [cos 2i? sin 2i?" exercise. A simple calculation shows that this product is . _ v sin2i? - cos 2d 2.7 Rotate the axes through 1?, project onto the new jc-axis, then rotate the axes through -1?. The required matrix is cosi? — sin t? sin 1? cos 1? 1 0 0 0 cos 1? sin 1? -sini? cosi? cos21? sin 1? cos 1? sini? cosi? sin2 i? 2.8 A{pxi + qx2) - pA%\ + (1 -p)Ax2 = pb + (1 -p)b = b a b 2.9 If X - " '. is such that X2 = 0 then we have a2 + be = 0, b(a + d) = 0, c(a + d) = 0, To ol be + d1 = 0. If b = 0 then clearly a = d — 0 and X = A which is of the required form. If b ^ 0 then d = -a and a2 + be = 0 which gives X = ^/t " . Writing z= y/band a b -a2/b -a w — ajyjby we see that X is again of the required form. The result fails for real matrices since. To ol for example, A = 1 0 is such that A2 = 0 but there is no b such that -b2 = 1. 2.10 [A lj = zji= [A%. That A + B = A + B and AB= BA follow immediately from similar properties of complex numbers. 2.11 (A + #)' = A' + (X)' = A'+A = W + A. Hence A + /F is hermitian. Similarly, A - A is skew-hermitian. For the last part, follow exactly the proof of Theorem 1.10. 3.1 pi becomes pi ;p4 becomes p2« Pi becomes P3ip2 becomes P4. 3.2 ps becomes pi; pi becomes p2;p2 becomes p3;p4 remains the same. 3.3 pi becomes7P3ip2 becomes 6p4;P3 becomesapi;p4 becomesPp2. 3.4 pi becomes ap2 + P3; P2 becomes (3p2; pi becomes pi + np2.

12. Solutions to the Exercises 209 3.5 3.6 "1 2 31 3 1 2 5 5 8 ■V* 3 3 1 2 0 1 2 3 10 113 1112 1 1 1 2 3 0 -5 -7 0 -5 -7 3 1 2 0 3 1 0-1 1-10 1 2 3 0 -5 -7 0 0 0 12 0 3 1 0 0 3 0 2 0-21-22 0-11-10 1 2 » ri 2 3i 3 1 2 2 3 I ■V* 1 1 2 3 0 -5

-7 0 -1 -5 2 0 3 1 1 2 0 -5 0 0 3 -7 _IK 5 . 0-11-10 0-21-22 2 0 -1 0 3 0 2 0 2 0 -1 0 0-1 0 0 0 w0 0 0 "3 1 1 -1 0 0 2 0 8 3 0 3.7 a =1. 3.8 10 0-2-30 0 10-1 21 0 0 1-1-52 0 0 0 0 0 0 3.9 Each is L. 3.10 Each is 10 0 1 0 110 0 0 0 0 0 0 0 0 3.11 \\p\ + >2P2 + ^iPi — 0 gives \i + >3 = 0 = 2X| + >2 = ^2 + ^3. the only solution of which is \\ = >2 = ^3 = 0. 3.12 The maximum number of independent rows (and columns) is 3. 3.13 The row rank is 2. 3.14 1 1 1 1 \ 1 1 1 \ 2 2 2 1 4 1 4 3-\ 6 \ 6 1 1 1 0 X-1 0 0 0 \-\ 0 0 0 1 4 0 0 2-X 2 \-2 -2 soifX ^ 1,2 then the row rank is 4. When X = 1 the row rank is 2, and when \ = 2 the row rank is 4. 3.15 The Hermite form of each matrix is /3 so they are row-equivalent. 3.16 The row and column rank are each 3. 01 3.17 Proceed exactly as in Example 3.26. P — 1 1 J f 4 4 -2 -3 andfi = 1 0 0 0 -1 0 f 1 1 do the trick. (The solution is not unique, so check your answer by direct multiplication.) 3.18 '* °Q forn= 1,2,3,4. 3.19 The number of zero rows in the normal form of A is n - rank A and the number of zero rows in the normal form of B is m — rank B. Since the latter must be less than the former we have n — rank A ^ m - rank B whence rank B ^ m — n + rank A. 3.20 By Theorem 3.10, row-equivalent matrices have the same rank and so are equivalent. 3.21 Both A and A' have the same normal form. 3.22 Each has normal form [/3 0]. 3.23 The coefficient matrix has rank 3 whereas the augmented matrix has rank 4. There is therefore no solution (Theorem 3.17). 3.24 A row-echelon form of the augmented matrix is 110 1 4 0-2 0-6 -1 0 0 1-1 1 0 0 0 0 a-1 Thus the coefficient matrix and the augmented matrix have the same rank (i.e. the system has a solution) only when a = 1. In this case, since the rank is 3 and the number of unknowns is 4 we can assign 4-3=1 solution parameter. Taking this to be t, the general solution is x = It + |, y = -3/ + J. z = 1 +1.

210 Basic Linear Algebra 3.25 In the augmented matrix interchange the first two columns and the first two rows. This has no effect on the rank. A row-echelon form of the resulting matrix is 1 2 1 -6a 0 0-6 -1 20a 0 0 /? 4 +6a Thus if (3 ■£ 6,0 the rank of the coefficient matrix is 3, as is that of the augmented matrix. Hence a unique solution exists (Theorem 3.18). When ^ = 0 the last line of the above matrix becomes [0004 + 6a] so a solution exists only if a = —3. In this case the solution is x = —^z + y and y = —\z — \ where z is a parameter. When ^ = 6 the above matrix becomes 1 2 1 -6a 0 0-1 20a 0 0 6 4 + 6a *Vf 1 2 0 0 0 0 1 —6a -1 20a 0 4+126a In this case a solution exists if and only if a = —^ and the general solution is x = —^y — |, z= 42 1 6.V 3.26 [10 0] 0 1 0 0 0-2 i [1 0 0] 0 0 1 0 1 0) "1 -2 0" 0 1 0 0 0 1 3.27 It is readily seen

that rows 1, 3, 4 are linearly independent, so r ^ 3. If r = 3 then the second row must be a linear combination of rows 1.3.4. This is the case if and only if a = 0 = d andbc= 1. 3.28 If AX = B then necessarily X is of size 3 x 2. Let the columns of X be X| and x2 Then AX = B is equivalent to the two equations Ax\ = 0 (homogeneous) and Ax2 = b2 (non- homogeneous). By the usual reduction method the latter is easily seen to be consistent if and only if a = -I. 3.29 1-10- 1 -5 a 2 1-1-41/9 II I-4-67 14 2-8-5 6 "1 -1 0 -1 -5 0 1-21 12 0 0 5-5 -25 0 0 0 0 0 a /?-a-7 a -20 + 37 8a-/?-ll7 + 56 so a solution exists if and only if 8a —0 -117 + 56 = 0, and in this case the rank of the matrix is3. Whena = /3=-1.7 = 3,6 = 8thesolutionisjc = 2tt + 3/,y= 1 +u-2t,z = 2+u+5t where u, t are parameters. 3.30 Interchange the second and third columns, and the first and second rows. The matrix becomes 1X01 -2 -\ /i + 3 -3 2 3* 4 -\ 1 \ 0 1 0 \ p + 3 -1 0 0 1-p -\-\ If/x 7^ 1 then the rank is 3 and a unique solution exists. If/i = 1 then a solution exists if and only if \ = -1. In this case the rank is 2 and the general solution involves one parameter. 3.31 min{m,rt}. 3.32 r 3.33 Consider A = 1 -1 -1 1 -1 1 andfl = 1 0 0 0 The normal form of A is B so these matrices are equivalent. But they are clearly not row-equivalent.

12. Solutions to the Exercises 211 334 Let A and B be equivalent. Then A can be transformed into normal form N by a sequence of row and column operations, and then N can be transformed into B by a sequence of row and column operations (namely the 'inverses' of those that transform B into N). Hence A can be transformed into B by a sequence of row and column operations. Conversely, if A can be obtained from £ by a sequence of row and column operations then A and B must have the same normal form. 3.35 a = 0,b= l.c = 2. 336 Treat the system as a system of equations in which all calculations are done modulo S. For example, adding the first row to the second gives [0 2 1 2 2 2]. 4.1 The given matrix has rank 2. By Theorem 4.1, it therefore has neither a left inverse nor a right inverse. 4.2 Proceed as in Example 4.3 1 0-1 1 -1 2 -1 1 -1 "1 1 1" 1 2 3 0 1 1 -i = 1 2 1 1 3 2 1 0 1 -t -i I -1 J L 2 0 - n i i 2 3 1 2" 1 3 has rank 2 so is not invertible; 1111 1 2-12 1-12 1 13 3 2 T "I i _! _2n 3 3" 4 _1 _4 "99 _2 1 9 9 2 2 3 3 3J 11 11 13 12 12-11 5 9 16 has rank 3 so is not invertible. 43 The product Ax can be written as x\&\ + x2a2 + • • + jc,,a,,, i.e. as a linear combination of the columns of A. Then Ax = 0 has only the trivial solution if and only if ai,..., a,, are linearly independent, which is the case if and only if A has rank n, which is so if and only if A is invertible. a b\ which is of rank 2 (i.e. A is invertible) if and only if a f 0, 4.4 If

c = 0 we have A = 0 d d^O which is equivalent to ad ^ 0. If a = 0 then A \fbcf0. If a f 0 and c f 0 then c d 0 b a b c d a b 0 ad-bc which is of rank 2 if and only which is of rank 2 if and only if a d -be ^0. In this case we have A'1 = ^^ apt~cp\ d -b —c a 4.5 1 1 0 1 0 0 1 2 a "1 0 0 0 1 0 0 0a which is of rank 3 if and only if a ^ 0. In this case, by the process of Example 4.3, A~l = 0 1 0 1 -1 0 .111 La a a. 4.6 The proof is by induction. The result is trivial for p = 1. By Theorem 4.4, the inductive stepis(^,.-MPVi)_I = A£Mi "Ar^^iV'-V-

212 Basic Linear Algebra 4.7 That A3 = 0 is routine. Using this fact, we have AxAy = (h + xA+ ljX2A2)(h + yA+ iyM2) = /3 + xA + \x2A2 +yA + xyA2 + \y2A2 = h + (x+y)A + \(x + y)2A2 = Ax+y. It follows that AXA-X = A0 - /3 whence Ax is inveitible with A~x = A-x. 4.8 That A,,A,, - A,,+m is routine. It follows that AnA-H = A0 = /2 so An is invertible with A;1 = A_,,. Similarly. BnBm = £n+m and B? = £_,,. Finally, by Theorem 4.4, (A,,B,,)~l = B-mA-,,. 4.9 Since (/,, + A)(I,, -A) = /,, -A2 = (/,, - A)(ln + A) we have that (/,, + A)(/,, - A)(l,, + A)"1 = (/,, - A)(ln + A)(ln + A)"1 = /,, - A and so (/,, -A)(In + A)"1 = (/,, + *)-■(#,, -A). If A is skew-symmetric then A' = -A and we have PF» = (/, ^)(/^)^1(/,-^(/,, +A)"1]' = <#■+*)-'</■-*)<#■ + AV(In-A') = (/fl+A)-'(/fl-A)(/ ,,-^(/^ +A) = In- It follows that P is invertible with P_I = P\ i.e. P is orthogonal. For the given matrix A it is readily seen that /,,"A = 1 -cos tf 0 cost? 1 —sin 1? 0 sin t? 1 whence P is the matrix stated. 4.10 I -1 0-1 0 -\ 0 0 1 I I 1 5*55 2 12 1 5 2 5 5J , (fn+A)-l = "i(l+sin2tf) j cost? i sin t? cos t? -jCosiS 1 2 4 sin ^ sin v cos -h sin tf i(l+cos2t?) 4.11 If B is row-equivalent to A then there is an elementary matrix P such that B = PA. Thus, if A is invertible, we have that Bis a product of invertible matrices and so is also invertible. 4.12 If AB is invertible then there exist elementary matrices P, Q such that PABQ = /,,. It follows that PA and QB are invertible, and from (PA)~lPA = /,, and BQ(BQ)'1 = /,, we see that A and £ are invertible. 4.13 (A + B)A-l(A -B)= (/,, + BA-!)(A -B) = A-B + B- BA~lB = A - BA"^. and similarly (A - £)A_I(A + B) - A — BA~lB whence we have the required equality. 4.14 Expand the right-hand product using the distributive law. The resulting sum is the left-hand side. If now A* = 0 then A*+I = 0 and the equality gives /i,= (/i,-A)(/fl+A + A2 + ... + A'"1) whence I,, — A is invertible.

12. Solutions to the Exercises 213 For the last part, follow the instructions; A~l = 4.15 We have 0 -1 0 0" 10 2 1 1-13 1 0-111 (BA-ln)[B(AB-I,,)-lA-I,,] = BAB(AB - l,,)~lA -BA- B(AB - I,,)-lA +1,, = [B(AB -1,,) + B](AB - In)~lA -

BA- B(AB - I„)~lA + ln v v ' = BA + B(AB-l„)-lA -BA-B(AB-I„)~lA + !„ = /
„. Hence BA - ln is also invertible. 5.2 (1) and (2) are subspaces; (3) is not since it does not contain (0,0,0,0); (4) is not a subspace since, for example. (1, -1,1, -1) and (0,0,0,1) belong to the set but their sum (1,-1,1,0) does not. S3 The sum of two continuous functions is continuous, and every scalar multiple of a continuous function is continuous. 5.4 The sum of two differentiable functions is differentiable. and every scalar multiple of a differentiable function is differentiable. 5.5 (1) If A and B are symmetric nx n then so is A + B\ and so is \A for every \. Hence the set of symmetric nx n matrices is a subspace. (2) The set of invertible nxn matrices is not a subspace since every subspace of Mat„x„ R must contain the zero matrix, and this is not invertible. (3) The matrices 1 0 o' 0 and 0 0 o" 1 are not invertible. but their sum is h which is invertible. Hence the set of non-invertible matrices is not a subspace. 5.6 If Ax = 0 and Ay = 0 then A(x + y) = Ax + Ay = 0. and A(\x) = \Ax = \Q = 0. Thus the solutions oMx = 0 form a subspace of MatmxI F. 5.7 Every subspace must contain the zero of the parent space. 5.8 The rank of the coefficient matrix is 3 and so for all given a,b,c the system is consistent. a € Mat 3,(1 R there exist scalars ^1,^2,^3 such that Thus, for every a b c = >i 1 1 1 + >2 1 2 3 + >3 1 3 2 whence the three column matrices span Mat 3*1 JR. 5.9 No; for example, the constant polynomial 1 cannot be expressed as a linear combination of the two given polynomials. 5.10 Let £, = 1 1 0 0 .£2 = 0 0 1 1 a b c d .£3 = 1 0 0 1 .£4 = 0 1 1 1 . Then we have = xE\ + y£2 + zEy + f£4

214 Basic Linear Algebra "1 1 0 0 0 0 1 1 1 0 0 1 0] 1 1 1 if and only if x + z = a,x +t = b,y +t = c,y +z + t = d. The coefficient matrix of this system of equations is A = which is of rank 4 (hence invertible). The system therefore has a solution (which is in fact unique). Hence the given set is a spanning set. k 5.11 (1) Suppose that ^Xfx; = 0 where each x, E S{. Since 5, C 52 we have that each X; £ S2 and so, since 52 is linearly independent (by hypothesis), every X,- = 0. Hence 5j is linearly independent. (2) If 5| is linearly dependent then by Theorem S.S at least one element of S\ can be expressed as a linear combination of other elements in 5|. But 5| C 52. so all of these elements belong to 52. By Theorem S.S again, therefore, 52 is linearly dependent. 5.12 The sets (1) and (2) are linearly independent since the 3 x 3 matrices formed from them are each of rank 3. As for (3), this set is linearly dependent; the third column matrix is the sum of the first two. 5.13 (1) Linearly independent; consider the entries in the (2,1) position. (2) Linearly independent; take a

linear combination of the four matrices to be equal to the zero matrix and solve the corresponding equations (only the zero solution possible). (3) Linearly dependent; we have 2 3 4 3 _ i ~~ 2 1 0 0 2 + i i i 2 1 + 5- 0 3 2 1 5.14 Follow the process in Example S.22. (1) and (2) are linearly independent. (3) is linearly dependent; we have 13 + X=3(1+X + 2X2) - 2(-5 + X + 3X2). 5.15 Every p(X) E IR,,[X] can be written uniquely in the form p(X) = a0 + fl|* + a2X2 + ■ ■ - + anX". m n 5.16 Every A £ Matmxn IR can be written uniquely in the form A = £ £ o^Ep^. 5.17 The sum of two diagonal matrices is a digonal matrix, and every scalar multiple of a diagonal matrix is a diagonal matrix. Hence the diagonal matrices form a subspace. A basis for this subspace is the set of diagonal matrices Epp of the previous exercise. 5.18 The set of Toeplitz matrices is clearly closed under addition and multiplication by scalars, and so forms a subspace. A basis consists of the Toeplitz matrices Ew where p f q and the Toeplitz matrix /,,. 5.19 Since cos 2x = cos2 x-sin2* we have that W= Span {/,#}. Now/ = cos2andg= sin2 are linearly independent. To see this, let X|Cos2x + >2sin2Jf = 0. Differentiate to get (\\ - X2)sin jccosx = 0. Since this must hold for all x we must have \\ = X2, and since the original equation holds for all x this means that \t = \2 - 0. Hence a basis for W is {cos2, sin2}. 5.20 Apply the process of Example S.2S. (1) is a basis; (2) is not.

12. Solutions to the Exercises 215 5.21 Apply the process of Example 5.25. The matrix A = is invertible so the given set is a basis. Since A-l = we deduce that 1 1 0 0 ce 2 0 -1 -1 -1 1 0" -1 1 1 1 1 -1 2 3 -3 -1 5 -2 0 3-1 1 -2 1 1 0 0 (fl,*,c,<0= Ml, 1,0,0)+ M-1.-1,1.2) +Mi.-i. 1,3) +MO. 1,-1,-3) where \t = 2a -6 + 5c-d,\2 = 3c-d, X3 = -a + b-2c + d. X4 = -a + 6. 5.22 WritingX= {jt|,Jt2} and Y = {.V|,V2} we observe that V| = j(jc2—jci) and >2 = 2.t2-.r|. It follows that Y C Span X and therefore Span Y C Span X. Similarly, jci = y2 - Ay} and JC2 = V2 — 2>'i whence we have the reverse inclusion Span X C Span Y. 5.23 (1) =>■ (2): Since {V|, v2} is a basis we can write w, and w, as unique linear combinations of w\, wi and so there is a matrix A such that = A Since Span {w,, h>2 } = V, we can write V|, v2 as linear combinations of wx, w2 and so there is a matrix B such that = B Consequently we have A is invertible. (2) =>(!): If ve w, .W2. Vl yi = BA = A V\ ?2 V| V2 wil Since {V|, v2} is a basis we must have BA = h whence with A invertible then we have = /T' whence {vi»v2} C Span{H>|,H>2}. Since {vi,v2} is a basis of V it follows that V= Span{u'|,H'2}. 5.24 Suppose that x € A n B. Then x = \\Q\ + ^2 where ai,a2 are the elements of the given spanning set of A, and similarly x = \$b\ + X4&2

where 61,62 are the elements of the given spanning set of B. Consequently we have \xa\ + ^2 - ^361 - ^462 = 0. But, as can readily be verified, the elements ^1,^2,61,62 are linearly independent. Hence each X-, = 0 and consequently x = 0. Thus we see that A n B = {0} and so has dimension 0. 5.25 If {vi,..., v„} is a basis of V as a vector space over C then every x £ V can be written n uniquely as x = £a4v4 where each ak € C. Writing ak - ak + 164 where ak ,bk G Rwe *=i have x = £ akvk + £ 64(iv4). Thus {vi vn, 1 vi,..., iv„} is a spanning set of V over IR. *=i k=\ This spanning set is linearly independent, for if £ akvk + 52 bk(ivk) = 0 then 53 a*v* = 0 and *=i *=i *=i 1 J] bkvk = 0 whence ak = 0 = 64 for each it since {vi,..., vn) is a basis over C. Hence the *=i above set is a basis for V over IR in which case V is a real vector space of dimension In.

216 Basic Linear Algebra 5.26 We have that (a, b, c, d) € Span X if and only if the system [2 7 3 2] 2 5 2 1 15 2 2 3 5 11 X y z t = a b c d is consistent. Now, by the usual reduction method. "2732a 2 5 2 \ b 15 2 2c 3 5 1 1 d 15 2 2 c 0 10 1 -2a + b + 2c 0 0 2-2 \0a-6b-Sc 0 0 0 0 -\0a + \0b + 6c-2d For consistency, we therefore require 5a-5b-3c + d=0. The first three vectors of X form a basis for Span X. Since R4 is of dimension 4, any vector not dependent on these three vectors (i.e. any vector not satisfying the above condition) may be added to obtain a basis for R*. e.g. the vector (0,1,1,0). "0' will do. 5.27 As a third basis vector. 5.28 Recall from Exercise 5.15 that V\,,[X] is of dimension n + 1. A possible basis is obtained by adding the monomials X2 and X3. 5.29 (1) is true. The standard conditions are x + y e W and \x € VV. These together imply \x+ny £W. Conversely, if the latter holds, take \ = /i = 1, then p = 0, to obtain the former. (2) is false; the subspace has dimension 1 with basis {(1,1,1)}. (3) is true. If the given spanning set is also linearly independent then it forms a basis and there is nothing to prove. If not, then at least one element is a linear combination of others. Removing this element, we still have a spanning set. Continuing in this way we discard elements one by one, obtaining a smaller spanning set each time. Do this until the remaining set is also linearly independent (in the worst case this will have only one element); it will then form a basis. (4a) is true. We have a{ 1,2,1) + b(2,2,1) = (a + 2b, 2a + 2b, a + b). Taking a'+ b = y and b = x we see that so also is (4b). (5a) is true. We can extend a basis of P to a basis of Q. (5b) is false. For example, consider P = {(x, x,x)\ x E R}, Q - {(jc, y, 0); x, y e IR}. We have dim P = 1 (see (2) above) and dim Q = 2, but P £ Q. (6) is false. For example, take w = —x and recall

that 0 cannot belong to a basis. 5.30 (1) Yes. (2) No. For example (1,0,0,0) and (0,1,0,0) belong to the set but their sum does not. (3) Yes. The set is {(0,0,c,rf); c,d e JR). (4) No. For example. (1,0,0,0) and (0,1,0,0) belong to the set but their sum does not. 5.31 Both sets are closed under addition and multiplication by scalars, so are subspaces. 5.32 Given a,/?,7 e C, to determine \,n, v e C such that (0,)0,7)=^3-1,2 + 21,4) + ^2,2 + 41,3) + 1/(1-1,-21,-1), solve the resulting equations for \,\k, v to obtain the unique solution '\~ 1 12 + 12i -2 + 2i 5-3i -6-61 2-6i -7 + 5i 6 + 61 -2-10i -l + -3i 6 + 61 This shows that the given set is a basis. Moreover, taking a= 1,/?=7 = 0we obtain (1,0,0) in terms of the basis, and similarly for the others.

12. Solutions to the Exercises 217 5 .33 We have "12 2 0 2 2 2 6 2 1 4 0 1 -1" -1 -2 1 -4 0 -3 'V* "10 0 0 1 0 0 0 1 0 0 0 2 -* 0 0 1 -1 0 0 The general solution is therefore '-It - w ±t + w 0 t w = t '-2' i 2 0 1 .0. + w '-]' 1 0 0 1 x = The solution space is therefore of dimension 2. 534 As defined, A + B is closed under addition and multiplication by scalars, so is a subspace. If C is a subspace such that A C C and B C C then for a £ A and b £ B we have a,beC and so a + b e C. Hence A + BCC. To establish the equality we show that LHSDRHS and LHSCRHS. For the former, observe that LHS D L n A/ and LHS D L n N whence, by the above, LHSDRHS. As for the latter, if x e RHS then x - y + z where y e L n A/ and z € LnN. Since y,z e L we have Jt = y + z e L. Moreover, since y G A/ we have jc = y + z € A/ + (L n N). It follows that jceLn[M + (LnA/)]=LHS. For the last part, take L={(x,x,0);x,yen}, M= {(O.y.z); y,z G H}, JV= {(Jt.0,0); jc e K}. Then A/ + N = fl3. L n (A/ + /V) = L, Ln A/= {(0,0,0)}, LOW = {(0,0,0)}. and the stated inequality holds. 535 It is readily verified that E,, is closed under addition and multiplication by scalars and so is a subspace of Map(IR, IR). Suppose now that/ is the zero map in E\. Then we have (Vjc e fl) a0+aicosx+b\s\nx = 0. Taking x = 0 we obtain a0 + a\ = 0, and taking x = it/2 we obtain a0 + b\ = 0. Thus o\ = b\ - -aQ. Taking x = 7r/4 we obtain a0 + -jjfli + ^^i = 0 whence a0 = ai = b\ - 0. Suppose now, by way of induction, that the zero map of £,,_i (with n ^ 2) has all its coefficients 0 and let/be the zero map of En. It is readily verified that tff+rPfis given by the prescription n-l (DV+ n2f)(x) = n2a0 + Efa2 -*2)(a* cos ifcje + bk sin itje) and since/ is the zero map of £,, we have that D*f+n2f\slbe zero map of £,,_j. By the induction hypothesis, therefore, we have that all the coefficients aQ, ax,..., a,,-t, b\,..., bH-\ are 0 and the formula for/ reduces to (Vjc gR) 0 = /(jc) = a,, cos jc + b,, sin jc. Taking jc = 0 we obtain an = 0, and

taking jc = 7r/2n we obtain b,, = 0. Thus all the coefficients of/ are 0 and the result follows by induction. It is clear that the 2n +1 functions generate £,,. Moreover, by what we have just proved, the only linear combination of these 2n + 1 functions that is zero is the trivial linear combination. Hence these functions constitute a basis for £,,.

218 Basic Linear Algebra 5.36 Let M be the set of rational functions described. Then clearly M is closed under addition and multiplication by scalars, so is a subspace of Map(IR, IR). Following the hint, observe that each/ e M can be written uniquely in the form / = fla/o + • • • + Or+s-lfr+t-l and so the/; form a basis of M. For the last part, it suffices by the hint (and by Corollary 2 of Theorem 5.8) to show that B is linearly independent. For this purpose, suppose that _£■_ + ...+ a' +JL. + ...+ bl =0. x-a +(x-*)' x-p (x~PY Multiplying both sides by (jc -a)r(x -fi)' we obtain ax{x -a)r-](x-pY + a2(x - a)'-2(x -/?)* + ■ ■ + a,(x -PY + bt(x "<*)'(* -PY-* +b2(x -qY(x ~PY~2 + •■■ +M* -a)' = 0. Taking the term ar(x~PY over to the RHS, what remains on the LHS is divisible by x -a and, since a f /?, we deduce that a, = 0. Similarly, we see that bs = 0. Extracting a resultant factor (x -a)(x -p), we can repeat this argument to obtain ar-\ = 0 = bs-\. Continuing in this way. we see that every coefficient is 0 and therefore that B is linearly independent. Hence B is a basis. 5.37 The result is trivial if n = 1 since f\ is non-zero. By way of induction, suppose that (/1)1 i«-i 's linearly independent whenever r,,..., r,,-\ are distinct. Consider(/,),<i<n and suppose that r\ r,, are distinct. If \ rfx + • • • + \,Jn = 0 then (VxeR) ^,^ + .- + ^^ = 0. Dividing by er"x (which is non-zero) and differentiating, we obtain h(ri - /•><"-'">* + ..- + >,,_,(/•,,_, -r^'---'")* = 0. Since the n - 1 real numbers n - r,, r,,_, - r,, are distinct, the induction hypothesis shows that \t = • • ■ = \,,-i = 0. Consequently. \,Jn = 0 and hence \n = 0 (since e'"* ^ 0). Thus /,,....,/,, are linearly independent. Hence the result by induction. Conversely, if the r, are not distinct then r, = r7 for some ij whence f, = fj and the/; are dependent. n 5.38 Suppose that £ X^M = 0. Since degp;(X) = i we have, on differentiating n times, i=0 n-l X,, = 0 whence £ ^iM = 0- Differentiating n - 1 times, we deduce that \n-\ = 0. i=0 Continuing in this way we see that every coefficient \,■. = 0 and therefore the given set is linearly independent. Since dim IR,,[X] = n + 1 and there are n + 1 such functions, they therefore form a basis (recall Corollary 2 of Theorem 5.8). 5-39 Since sums and scalar multiples of step functions are also step functions it is clear that the set £ is a subspace of the real vector space of all mappings from JR to IR.

Given t? £ E. the step function t?j that agrees with t? on the interval [a;, ai+1 [and is zero elsewhere is given by the prescription t?,W = <?(fli)k(x)-efl/tl(x)]. n+i Since then t? = £ tf, it follows that {ek ; * e [0,![} spans E. Since the functions ek are i=0 linearly independent they therefore form a basis of E. Similarly, the set F of piecewise linear functions is a vector space.

12. Solutions to the Exercises 219 That C is a subspace is clear since it is closed under addition and multiplication by scalars. That {gk; k G [0,1 [} is a basis of C is similar to the above. That every f g F can be expressed uniquely in the form g + e where g G C and e G E can be seen from geometric considerations. A typical f E F can be depicted as in the diagram Now think of strips of wood that can slide up and down and thereby manufacture g via e. 6.1 (1),(2). and (5) are linear; (3) is not linear since/(1,0,0) +/(-1,00) ^/(0,0,0); (4) and (6) are not linear since in each case, for example,/[2(0,0,0)] f 2/(0,0,0). 6.2 (1) and (2) are linear. (3) is not linear since in general TB(\I,,) f \TB(I,,). 6.3 (1) and (2) are linear; (3) is not linear since in general /(\p{X)) f- V(p(X)). 6.4 fA(x + y) = A(x + y) = Ax + Ay = fA(x) + fA(y) andfA(\x) = A\x = \Ax = \fA(x). 6.5 By theorems in analysis, f(p + q) = f p+ fqand f\p= \ f p. 6.6 No. A and -A have the same Hermite form, so/(4) = /(-A) whence/(A) -/(-A) = 0. But/(A - (-A)) = /(2A) = /(A) since 2A has the same Hermite form as A. 6.7 Observe that T(0) = IB2 so if T is linear we must have B2 = 0. Conversely, if B2 = 0 then T{A) = AB + BA- 2BA + 3AB = 4AB - BA which is linear. 6.8 D-flR,,!*]) = Hi-i[X] and />-({0}) = R (the set of constant polynomials). 6.9 we have that rrrw = {/(*): * e rrw} = {/(*) ; /(*) G r(A)} = (W); /(*>) = M for some a g A} = /-(A), and/-/-/-(fl) = {a ; /(a) g rr(B)} = {fl ; /(a) = f(x) where/(x) € B} = {a ; /(a) G B) = /-(B). 6.10 D"(X) is the set of polynomials whose even coefficients are zero. D""(X) is the set of polynomials whose even coefficients, except possibly the constant term, are zero. 6.11 The image of the x-axis is the line y = jc; the image of the v-axis is the line y = — jc; and the image of the line y = mx where m f I is the line y = jzjjf •*. 6.12 Observe that if x G /-(K) then/(x) G Y and therefore we have f~{f~{Y)) C Y. Since r is inclusion-preserving, it follows that/"*[* 0/-(0] C r{X)nf~(f~{Y)) C f^(X)n Y. For the reverse inclusion observe that if a G f~{X) n Y then there exists b G X such that f(b) = a£Y whence fr G x n/-(K) and so a = /(&) G /"*[* n/-(K)]. 6.13 /(fl,fr,c) = fl(1,0,1)+6(1. 1,0)+c(0,1, 1).Im/=Span{(l,0,1),(1,1,0),(0,1,1)} = R\ and Ker / = {(0,0,0)}. 6.14 Ker/= {(a,b) G R2 ; b = 0} = Im/. 6.15 / : Ft3 -* Ft5 given by/(a,fc,c) = (c,0,0) is such that 1m / c Ker/; and g : R5 -> Fl> given

by g(a, fc, c) = (fc, c, 0) is such that Ker gC\mg. 6.16/(0,^,^^^)=0(1,1,2,0) + ^-1,0,-1,-1) + ^3,2,5,1) + 6(-1,-1,-1,0) and so

220 Basic Linear Algebra Im/ is spanned by the rows of the matrix 1 12 0 -1 0 -1 -1 3 2 5 1 -I -1 -1 0 0 0 0 0 0 which has Hermite form ri oo 0 1 0 0 0 1 000 000 n -i 0 0 oj A basis of Im/is {(l, 1,2,0),(-1,0,-1,-1),(-1,-l,-1,0)}. It is readily seen that Ker/= {(-2c,fc,c,c,0) ;fc,ce JR), so a basis for Ker/is {(0,1,0,0,0),(-2,0,1,1,0)}. 6.17 AbasisofIm/is{X2,X3};Ker/=R 6.18 Take/(Jt,>,z,r)= (jc,0,z,0)andg(je,>,z,/)= (0,y,0,r), "i i r 6.19 Argue as in Example 6.12, noting that the matrix 2 -1 -I is invertible. [l 2 -1 6.20 Suppose that £ q/(v,-) = 0. Then/(£ a.v,) = 0 whence./being injective, £ aw - 0. The fact that a\,..., a,, are linearly independent gives each a-, = 0 whence/(vi),...,/(v,,) are linearly independent. 6.21 (I) => (2) : If (1) holds then for every x G V we have/(jc) G Im/ = Ker/ whence /^jc) = 0, so that/2 = 0. If/ = 0 then Ker/ = V whence the contradiction n = dim V = dim Ker / = dim Im / = 0. Hence / f 0. By Theorem 6.4 we have n = dim V = dim Ker /+ dim Im / = 2 dim Im / so n is even and the rank of/ is ^n. (2) => (I): If (2) holds (benf(x) = 0 gives/(jc) G Ker/ whence Im/ C Ker/. By the Dimension Theorem, dim V = rank + nullity, i.e. n = in + nullity. Hence dim Ker/= |n = dim Im /. It follows by Theorem 5.9 that Ker / = Im /. 6.22 Consider the differentiation map D: R2[X] -» R2[X]. It is not possible to write X2 as the sum of an element of Im D and an element of Ker D. 6.23 Writing/(jc) = a sin x + fccos jc we have (1) tf(f) = 2a so/ € Ker tf if and only if a = 0. Hence Kerd = Span {cos} and the nullity is 1. (2) <?(0 = 0 so Ker d = W and the nullity is 2. (3) ti{f} = 0 if and only if a = 0 in which case the nullity is 1. 6.24 If tf = af + bg + ch then T{ti}{x) = -ex so rank T = I. Ker T= {af+bg; a,b£ IR} so the nullity of T is 2. 6.25Take/(£a.xO = (flo,...,a,,). i=0 6.26 If (je,>) G Ker i? then (jc,> -/(jc)) = (0,0) and so x = 0 and y = f{x} = /(0) = 0 Hence Ker d = {(0,0)} and so d is injective. By Theorem 6.5, t? is an isomorphism. "1 1 I] 1 2 3 J 1 2 is a basis of IR3. Proceed as in Example 6.15 to obtain Ax,y,z) = (4* ~4y + z, 5jc -5> + z, 8* - 10y + 3z). 6.28 /(a + bX + cX2) = fl/(l) + bf(X) + cf(X2) = a + cX + bX2 + cX\ 6.29 If C is considered as a complex vector space, then for \ = 1 -1 we have f{\(1 + i)) = /(2) = 2 whereas >/(1 + «) = (1 — i)(l -1) = -2i, so that/ is not linear. 6.30 / is linear if and only if 6.27 Recall Example 5.25: the matrix is invertible, so {(1,1,1), (1,2,3),(1,1,2)} /(Vi + >2V2 + A3v3) = Ai/(v,) + X2/(v2) + >3/(v3),

12. Solutions to the Exercises 221 i.e. if and only if This is the case if and only

if/i = 0. Then/(jcvi +yv2 +zv3) = 0 if and only if x = 0, y+z = 0 and so Ker/ = {a(v2 - v3); a e F). Thus a basis of Ker/ is {v2 - v3}. 6.31 A basis of Im/is {(1,0,0),(1,0,1)} and a basis of Ker/is {(1,-1,-1)}. /-(A) = {(jc,-jc,z) ; jc,z e R}, abasis of which is {(1,-1,0),(0,0,1)}. 632 f{x,ytz) = 2{x-y)X+(y+z)X2 + (x+z)X\ AbasisofIm/is{2X+X3,-2X+X2}. A basis of Ker/is {(1,1,-1)}. This can be extended to the basis {(1,1,-1),(1,0,0),(1,1,0)}. r-i 1 1 is invertible so {aybyc} is a basis. We have 6.33 The matrix 1 -1 1 1 1 -1 (x,ytz)=±(y + z)a + ±(x + z)b+±(x + y)c and so f(xtytz)=(y+^x + z)t{(z-yl^(\-\)x + (\ + l)yl{[(\-\)y-x + \z)). f is injective if \ f - 1 and \ f 2. The dimension of W is 2. If \ = 2 then /(1,1,0) = (1,-1,2,0) andr{(l, 1,0,0)} = (2,-2,5). 634 If x = /(j|) and > = /($2) where $1, $2 € S then for / e [0,1] we have tx + (1 -/)> =/1/5, + (1 -/)52] ertf) and 50/-(5) is also convex. 635 By the Dimension Theorem (6.4) we have ^(-O'dimV; = -dim Im/,-dim Ker/, + dim Im /2 + dim Ker/2 - dim Im /3 - dim Ker/3 +(-1)" dim Im/, + (-l)"dim Ker/, Since the sequence is exact, dim Ker/, = 0 and Im /,, = V,,+I. Moreover, Im / = Ker/+|. The n above display therefore reduces to £(-1)' dim VJ = (-l)"dim V,,+] whence the result follows. 636 The rank is 2 so the nullity is 1. 637 (1) (g of){x) = g(y) where y = f{x) so Im {g of) C Im g. (2) If/(jt) = 0 then g[f(x)] = g(0) = 0 so Ker/ C Ker(g of). (3) By (1) we have rank {g of) ^ rank g. By (2) and Theorem 6.4, we have rank(go/) = dim V-dim Ker(go/)^ dim V-dim Ker/= rank/. Hence rank (go/)< min {rank /, rank g). Let {ei,..., e,,) be a basis of V. Let m = dim IV and 5 = rank /. As basis for IV we can take[f{ei)t...J{es),ws+l w,,). Then Img = Span{g/"(e,) gf{e,),g{ws+l) g{w,,)} - Span{gf(e,) gf{ep),g{ws+\) g(wm)} where p = rank gf. It follows that rank g ^ p + m -5 = rank g/" + dim IV - rank /

222 Basic Linear Algebra 6.38 {l,X,X2,XJ} is a basis of IR3[X] and/(1) = 0,f(X) = 0,f(X2) = 0,/(X') = 6(X - 1). So the nullity of/ is 3 and the rank is 4 - 3 = I. 6.39 If {e,,e2,e3} is the natural basis of IRJ then /v(<?,)= (0,-^3,72),^2) = (>3,0,-V|), /vfo) = (-^2,^1,0). Since y f 0 we can assume that >, ^ 0. Then we have /,(^.) = -^/,(^)-¾^) y\ y\ where /v(e2),/y(ej) are linearly independent. Hence a basis for Im /v is {^(e2),/v(e3)} so rank /v = 2. It follows that dim Ker/y =3-2=1 whence Ker/v = Span {y} since fy{y) = 0. 6.40/(1,0,0,0) = 1 0 0 -1 ,/(0,1,0,0) = 0 1 1 0 ./(0,0,1,0) = 0 i -i 0 ./(0,0,0,1) = 1 0 0 1 So/ carries a basis to a basis and therefore, by Theorem 6.5, is an isomorphism. 7.1 1 2" 2 -1 -1 0 1 -2 -3 -2 2 3 7.2 2 - 0 -10] [0 2 -2 -lj; [l - 2 -I 1 0 7.3 "2 2 1] 0 3 0 0 -4 1 7.4/(0,1,0) = /(1,1,0)-/(1,0,0) =(2,-2,6) and/(0,0,1) = /(1,1,1)-/(1,1,0) = 2 2 1" (1,-2,3) and so the matrix of/ relative to the natural basis of IRJ is 3 -2 -2

-2 6 3 7.5/(*,>>, Mat8/ = z) = (z,>\ ro -2 r 0 I 0 1 2 0 *) • Mal8g = '222 0 I -I -1 -2 0 7.6 A has rank 3 and so is invertible. The matrix of/-1 is A Mat8/$/ = r 2 2 -11 -l -l 0 0 0 2 1 is A~l, namely r-f -6 2i LI 3 -u 7.7 Use Theorem 7.3 and induction. 7.9 A = 7.8 1 0 0 1 0 0 0 1 I I 0 0 0] 1 I 0 r o i o" -10 0 0 0 1 ; B = ■ 0 0 -1] -10 0 1 1 1 The matrix X represents the identity map relative to a change of reference from the first basis to the second. Since (1,0,0) = 1(1,1,0)-1(0,1,1)+1(1,0,1) (0,1,0) = 1(1,1,0) +1(0,1,1) -1(1,0,1) (0,0,1) = -1(1,1,0)+1(0,1,1)+1(1,0,1) 1 I -f -1 1 1 1 -1 1 7.10 We have lo determine an ordered basis B = {&,, b2t b)} of IR3 such that P is the transition matrix from B to the natural ordered basis {ei,e2,ej}. B = {e2 + ehet + e^e\ + e2}. 7.11 If A is similar lo B then there is an invertible matrix P such that A = P~]BP. Since P is invertible. so is F. Then A' = PlB,{p-])' = [(/>''')']-,ff(/>"1)' and so A' is similar lo fl'. 7.12 If A = P~lBP then by induction Ak = P~]BkP. we see thai X - \

12. Solutions to the Exercises 223 7.13 cosi? -sini? _ 1 i sin 1? cos t? i I 7.14 The respective matrices are "0 1 0 0 ... 0 0 0 2 0 ... 0 0 0 0 3 ... 0 ■i-i (1) 0 0 0 0 0 0 0 0 (2) 0 e 0 0 n 0 0 n-\ 0 0 1 i 0 0 0 n-2 0 0 0 0 0 0 0 0 (3) 0 1 0 0 0 0 0 0 0 0 -2 2 0 0 0 -3 .. 0 .. 3 .. 0 .. 0 .. . —n . 0 . 0 n . 0 .0 0 0 ... 1 0. 7.15 If x f 0 is such that f~l (x) f 0 then for every k ^ p - 1 we have/*(.r) f 0. To show that {*,/(*),. •• ,r'W} is linearly independent, suppose that >u* + \if{x)+- + \p-lf-i(x) = 0. Applying/p-1 to this we obtain Xvf^^x) = 0 whence X0 = 0. Thus we have M*)+ .-.+ ^1^(^) = 0- Applying f'1 to this we obtain similarly \\ = 0. Continuing in this way we see that every \, = 0 and consequently the set is linearly independent. If/ is nilpotent of index n = dim V then {xj{x),... ,/'"''(*)} is a basis of V. The matrix of/ relative to this ordered basis is then that in the question. Conversely, if there is an ordered basis with respect to which the matrix of/ is of the given form then to see that/ is nilpotent of index n it suffices to observe that the matrix M in question is such that M" = 0 and A/"-1 ^ 0. 7.16 We have /(1) = 1 f(X) = \+X f{X2) = (\+X)2= \+2X + X2 f{X") = (1 + X)" = I + (1)X + QX2 + ■■■ + Xn and so the matrix of/ is Mil 0 12 3 0 0 13 1 © 0 0 0 0... 1 7.17 We have /(1) = /W /(1 +X) 0 1 0 0 1 0 1 0 ■ ■ 1 1 0 1 = 1 0 0 0 1 0 0 0 + 1 1 0 0 1 1 0 0 t * 1 1 0 0 + - 1 1 1 0 ^ •« 1 1 1 0 + 1 1 1 1 and so Mat f = -1 1 0 1 -1 1 0 1 -1 0 0 1 . The rank of this matrix is 3 and so dim Ker/ = 0

224 Basic Linear Algebra Im/ 7.18 = Span < 0 I 0 0 • 1 0 1 0 • 1 0 0 I }■ 10 0 0 -10-1 1 0 0 0 0 0 0 0-11 • [1 0 0 0] 0 0 0 0 0 0 10 0 0 0 0 7.19 (1),(2),(3) are

routine. If A and B are similar then there is an invertible P such that A = P~lBP. By (3), tr(A) = tr{P~lBP) = It(P~]PB) = ir{B). For the last part, observe for example that tr 1 0 0 -I = 0=tr but these matrices are not similar. 8.1 All expressions are the same. -,/12345678' W 1458 7 326 !)■ 8.2 Routine. \ /1234567 8 Y \\ 2 3 4 5 6 7 8 8.4 Even; odd. 8.5 -28. 8.6 26 + 13* + \2 -\\ 8.8 400. 8.9 If the elementary matrix is obtained by interchanging two rows (columns) then the determinant is —1; by multiplying a row (column) by \ it is \; and by adding \ times one row (column) to another it is 1. 8.10 By induction. For the inductive step, use a Laplace expansion via the first column. 8.11 x f \. 8.12 Use a Laplace expansion via the first row, then via the first column of the resulting 5x5 matrix, then by the first row of the resulting 4x4 matrix, and so on. 8.13 Add the first row to all the others; the answer is n\. 8.14 det 0 0 0 0 0 0 0 0 0 0 0 1 1 0 0 0 0 1 I I 0 0 0 0 0 0 0 0 0 0 = (-1)(--1)+(--2)+-+1 = (_|)2 i = #_i\5-(--i) 8.15 For example, take ps — 2p2 then use a Laplace expansion via the first column 8.16 I*"1 = 1-2 4 I 1 -2 4 4 1 -2 8.17 det A" = (det AY = 6 and so det A = 0. 'bc-f1 fg-hc hf-bg fg-hc ac-g2 hg-af hf — bg gh—af ab—h1 and P~lAP = '9a 0 0] 0 9b 0 0 0 9c det A = 3fiabc 8.18 d -b I 0 0 0 -I 0 -1 0 1 8.19 The first two have determinant 1 so adjugate=inverse; these are respectively I 0 -I 1 1 -2 -3 -1 5 4 1 -11 I -3 1 -I _4 9 The third has determinant 6 and inverse i 0 8.21 Since A- adj A = (det A)l,, we have 6 -3 0 0 0 3 0 -2 2 (det A)(det adj A) = det [(det A)/,,] = (det A)".

12. Solutions to the Exercises 225 det A A being invertible, det A f 0 and so det adj A = (det A)"-1. 8.22 ABadjAB = {del AB)ln = (det A det fl)/,, and so adj AB = B~x A"1 det A det B - B~l adj A det B = B~l det B adj A = adj fladj A. 8.23 adj (adj A) = det adj A- (adj A)"1. By Exercise 8.21 and the fact that (adj A)"1 = we obtain adj (adj A) = (det A)"~2A. 8.24 If A is upper triangular then ai} = 0 when i > j. Thus, for / < 7 we have det A,-, = 0. Consequently, for 1 > 7, [adj A],y = (—l)'*" det Ay,- = 0 and so adj A is also upper triangular. 8.25 For a symmetric matrix, A,j = Ay,. 8.26 Observe that [adjA],,- = (-l),+>det A,, = (-l),+'det A^ = [adj A% and so adjA* = adj A = adj A. 8.27 det A = (jc -a)(jr -&)(jc -c)(a - b)(b - c){c - a). 8.28 det = (jc - a)3(x + 3a) so the solutions are x — a and x = -3a. 8.29 Start withe 1 -yc2. 830 Begin by adding all the rows to the first row, thereby obtaining a factorn-\. Then, for 1 >1.takep, —pi. 831 Begin by adding all the rows to the first row. thereby obtaining a factor na + b. Then, for 1 >1.takep, — ap\. 832 Begin by subtracting the first column from the others. Then do Laplace expansions via the rows. The

solutions are x = 0,..., n — 1. 8.33 For det B,, first subtract the first row from the other rows, then subtract the last column from the other columns. For det A,, first add the last column to the first column. Now do a Laplace expansion via the new first column and use the result for det B,,-\. For the last part, use the previous formula and induction. 8.34 Use a Laplace expansion via the first row of A,,+2. The last part follows by induction and basic trigonometry. 835 For det A,, subtract the last column of A,, from the other columns and then use a Laplace expansion via the first row. For the recurrence formula for det Bn use a Laplace expansion via the last column. For the last part use induction. 836 We have det A B B A = det A B B-A A-B = det = det = det A + B B 0 A-B I B 0 A-B 1 B 0 A-B A + B 0 0 I A + B 0 0 I det = det (A-B) det (A + B). 837 Clearly, only if A 0 B C P Q AP AQ R S BP + CR BQ+CS A=P~\ BP+CR=0,BQ+CS = S-RP~lQ and this is of the given form if and

226 Basic Linear Algebra Since B = -CRp-* we have C{S - RP~*Q) = S-RP~XQ, so we can choose C = I and then B=-RP~X. Now clearly det/VdetA/=det 1 p-*Q 0 S-RP~lQ = det I 0 0 S-RP~lQ = det (S-RP~lQ) i.e. det P~l det A/ = det (5 - RP~]Q), so det M = det P det (5 - RP~]Q) = det {PS - PRP~l Q). If now PR = RP then PRP~l = K and so det A/ = det (#»S - RQ). Likewise, we also have det M = det (5 - RP~lQ) det P = det (S/> - RP~XQP) so if />(? = QP then det A/ = det {SP - RQ). 8.38 Follow the hint. 28; -18; n\. 9.1 (1) The eigenvalues are 1,2,3 each of algebraic multiplicity 1. (2) The only eigenvalue is 1, of algebraic multiplicity 3. (3) The eigenvalues are 2,1 + i, 2 - 2i, each of algebraic multiplicity 1. 9.2 Suppose that Ax = \x where x ^ 0. If X were 0 then we would have the contradiction x = A~iAx= A~iQ= 0. Hence\ ^0. Nowx= A~lAx = A~l\x= Xi4'"xandsOi4'"x= X'"x; i.e. X"1 is an eigenvalue of A~l. 9.3 If Ax = Xx then by induction we have Akx = X*x. Thus, for any polynomial p{X) = a0 + a\X + • • + a,,X" we have p{A)x = /?(X)x whence p{\) is an eigenvalue of fj{A). 9.4 Proceed as in Example 9.3. (1) The eigenvalues are 0,1,2, each of algebraic multiplicity 1. £o = Span £i = Span £2 = Span x 0 —jr 0 x 0 JC 0 x ; x f 0 \ so a basis is x ^ 0 \ so a basis is x ^ 0 \ so a basis is (2) The eigenvalues are 0,1, -1, each of algebraic multiplicity 1. £0 = Span £1 = Span £-1 = Span x ^ 0 > so a basis x ^ 0 > so a basis is x ^ 0 > so a basis is 9.5 The eigenvalues are 1 and 2, of algebraic multiplicities 2,1 respectively. £1 has basis " 1" 0 1_-lj • r°i 1 L~i and £2 has basis 3 1 -5

12. Solutions to the Exercises 227 9.6 (1) Relative to the natural ordered basis

of IR3 the matrix of/ is values are 1,2 of algebraic multiplicities 1,2. (2) Relative to the natural ordered basis of R3 the matrix of/is 1 0 -1 2 2 2 2 1 2 "0 0 1 1 0 1 1" 0 0 The eigen- The eigenvalues are 0,1, -1 each of algebraic multiplicity 1. 9.7 The first two matrices are those of Exercise 9.4. Suitable matrices P are therefore, respectively. 10 1] 0 1 0 -1 0 I and 1 -1 2 -1 -2 -1 1 1 1 [3 1 2] 1 2 1 1 1 1 and Proceeding similarly with the other two, we see that suitable matrices P are, respectively, 1 1-3 0 1 -1 -1 -2 5 9.8 Let det A,, = a,,. Use a Laplace expansion via the first column to obtain the recurrence relation a,, = a,,-x + 20a,,-2. Consider therefore the system of difference equations an = 0,,.,+20^-1 The matrix of the system is A = Theorem 9.8, the matrix P = 1 20 1 0 20 20 -5 4 which has two distinct eigenvalues, -4 and S. By is invertible and such that P lAP = -4 0 0 5 Now />'" = 180 4 -20 5 20 A" = P and so (-4)" 0 0 5" />-'=• 5«+i _(_4)«+i 4-5^+5-(-4)^1 4•5', + 5•(-4)', 4.5" + 5(-4)n Finally, an K = A"~2 ai bi = A"-2 21 1 - i " 9 5«+i -/-4)1+1 5«_(_4)« from which the result follows. 9.9 (1 -X)\ X3-X2-X + 3, X3 + 3X2 + 10* + 30. 9.10 We have (/,, - BA)[ln + B(I,, - AB)~lA] = I,,-BA + (/,, - BA)B(ln - AB)~lA = l,,-BA + {B- BAB){I,, - Afl)_,/\ = In-BA + B(ln-AB){ln-ABTxA = ln-BA + BA = A,- Hence (/,, - BA)~l = I,, + B{ln -AB)~XA. The last part follows from the fact that \ is an eigenvalue of XY if and only if XY - \l,, is not invertible which, in the case where X f 0. is equivalent to /,, - \~lXY being not invertible. -lAD- 9.11 P~lAP = e" 0 o «-•'* 9.12 P = 1 -1 1] -v/5 o v/5 i i l ; P- r°] _ii 0 1 1 1 0 0

228 Basic Linear Algebra -lAD — 9.13 Proceed as in Example 9.10. By Theorem 9.8 there is an invertible P such that P AP - X, 0 0 \2 and, since \ \ + \2 = a, we have P = P P ">2 ->l When the eigenvalues are not distinct, use the expression for A" given on page 170. " 4 + 2-4" 3-4" -4 + 4"] -2+2-4" 3-4" 2 + 4" . 9.15 -2 + 2-4" 3-4" 2 + 4" 9.16 Diagonalising A in the usual way we have P~lAP = D where n 9.14 \ 6 1 [6(-7)-1-6"] -^19(-7)--1 + 4-6"-']J P = Now the matrix E = ri i 0 1 -1 0 0] 1 1 />-' = i -i 11 2 2 u 0 0 1 , 0= "4 0 0' 0 9 0 0 0 4 [2 0 0 0 3 0 0' 0 2 is such that E2 = D and consequently the matrix B = PEP -i has the property B2 = A. A simple computation shows that B = r f I J 2 2 2 0 0 2 9.17 For the first part, use induction. The result clearly holds for r = 1. Suppose that it holds for r. Then IK+%I = \t^[A%\ ^ EKIIMyl ^ £**v-' = *'+ln'-' £ 1 = *'+V f=i i=i f=i f=i nk (1) If/?| < — then nk 1 + \P\Ua\\ + \P2\\IA%\ +■+ I/HIM,! + • • • ^ 1 + k\p\ + k2n\p\2 + ■ ■■ + kTtrx\P\r + • • • = 1 + k\p\{\ + kn\p\ + • ■ ■ + k'-xrrx\p\'-x + • • •) which is less than or equal to a

geometric series which converges. Thus we see that if |/?| < the S$(A) is absolutely convergent, hence convergent. (2) If Sp{A} is convergent then lim,- *, p'A' = 0 so (In-PA)(I,, +0A+^ + ---) = \\ml^0(ln-pA)(I,, + pA + .-- + b"A") = linwa.-zr'A''') = /«■ Consequently /,, - PA has an inverse which is the sum of the series. For the last part, let \ be an eigenvalue of A. Then \l,, - A is not invertible. Suppose, by way of obtaining a contradiction, that |X| > nit. Then — < —. Consequently, if we let |a| nk P=\~lwe have, by (1), that Se{A) converges and so, by (2), /,, -/?A = /,, -\'''A is invertible. It follows that \I,, -A is invertible, a contradiction. Hence we must have |X | ^ nit. 10.1 mA{X} = a0 + axX + • • • + apXp. We must have a0 f 0 since otherwise 0 = mA{A) = atA+ ■ + ApAp and therefore, since A is invertible. a,/,, + a2A + • • • + OpA?'* = 0 which contradicts the fact that mA(X) is the minimum polynomial of A.

12. Solutions to the Exercises 229 NowO = a0I,, + aiA + + apAp can be written A{at + a^A + I + <V4'-')=-fl()/n + apAr). Thus A ' is a whence, since A is invertible, we have A l = (at + a2A + linear combination of /,,, A,..., Ap_l. 10.2 For each of the matrices the characteristic and minimum polynomials coincide; they are respectively (X - 1)3; -X3 + X2 + X - 3; -X3 + X2 - X + 2. 10.3 cA(X) = det [A - XIn] = a0 + axX + • • • + A,,Xn. Taking X = 0 we obtain det A = a0. 10.4 We have cj,, (X) = X2-2X cos t?+ 1 = (X-cost?-/sint?)(X-cost? + tsint?). Then cgt[X] = mKt[X]. If t? is not an integer multiple of 7r then i sin t? ^ 0 and so rt has no real eigenvalues. 10.5 cD{X} = (-ir'^-'-.m^X) = X""1. 10.6 The matrix of/ relative to the natural ordered basis {(1,0),(0,1)} is A = mf(X)=mA(X) = X2-3. 4 -1 10.7 Consider the matrix 1 0 -1 0 2 0 1" 1 3 . The only eigenvalue is 2 (of algebraic multiplicity 3). The geometric multiplicity of this eigenvalue is 2 so/is not diagonalisable. 10.8 For all positive integers i we have A,+l = {PQ)i+l = P{QP)lQ = PffQ. If now h(X) = Zq + Z\X + • • + z«X" then we have Ah(A) = zoA + z,A2 + • • • + z«An+l = z0PQ + zlPBQ+- + z,,PBnQ = P[z0ln + ZiB+ •■ + z,,Bn)Q = Ph(B)Q. It follows immediately that AmB(A) = PmB{B)Q = 0. Similarly, we have BmA(B) = 0. Consequently, mA{X}\XmB(X) and mB{X}\XmA(X), and so we can write XmB[X] = p{X)mA{X} and XmA(X) = q(X)mB(X). Comparing the degrees of each of these equations, we deduce that deg p + deg q = 2. Thus, either deg p = 0 in which case p{X) = 1 and XmB(X) = mA(X), or deg q = 0 in which case q{X) = 1 and XmA(X) = mB(X), or degp = deg q = 1 in which case mA(X) = mB(X). 10.9 The matrix can be written as a product PQ where P is the column matrix [I 2 • - • r]' and (2 is the row matrix [1 1 • • •]]. Note

that then B = QP is the 1 x 1 matrix whose entry is \r(r+ 1). We havem8(X) = -$r(r+ 1) + X. Clearly,mA{X} f mB{X} andmB{X} f XmA(X). Thus, by the previous exercise, we must have mA{X} = XmB{X} = -\r{r + 1)X + X2. 10.10 The matrix of/ relative to the ordered basis {1,1 + X, 1 + X - X2} is A = -10 0 2 5 6 -2 -3 -4 c>,(X) = (X-2)(X + l)2 = m>,(X). 11.2 The eigenvalues are 0,0,6 ± 2VT7.

230 Basic Linear Algebra 11.3 The determinant is -27216 and the inverse is 1 126 r-2o l i i i 22 1 1 1 1 22 -20 1 1 1 22 -20 1 I 1 22 -20 1 1 1 22 -20 1 1 1 1 11.4 x = ±√6. 11.5 The MAPLE input is > eql:=205+a+208+4+b=1000; > eq2:=c+206+d+207+6=1000; > eq3:=204+e+200+f+196=1000; > eq4:=20+193+g+h+400=1000; > eq5:=i+j+192+k+l=1000; > eq6:=205+C+204+20+i=1000; > eq7:=a+206+e+193+j=1000; > eq8:=208+d+200+g+192=1000; > eq9:=4+207+f+h+k=1000; > eql0:=b+6+196+400+l=1000; > eqll:=205+206+200+h+l=1000; > eql2:=b+207+200+193+i=1000; > solve({eql,eq2,eq3,eq4,eq5,eq6,eq7,eq8,eq9,eglO,eqll,eql2}, {a,b,c,d,e,f,g,h,i,j,k,l}); The solution is given in the form h = 194, g = 193, c = 374,;" = j,b= 203, i = 197, a 380,/= 195, ^ = 221 -;',/= 179+;, it =-/ + 416. It follows that the range of values of it is 196 < it < 415.