TABLE OF CONTENTS

Chapter 1

MATRICES AND LINEAR EQUATIONS

A familiarity with matrices is necessary nowadays in many areas of mathematics and in a wide variety of other disciplines. Areas of mathematics where matrices occur include algebra, differential equations, calculus of several variables, probability and statistics, optimization, and graph theory. Other disciplines using matrix theory include engineering, physical sciences, biological sciences, economics and management science.

In this first chapter we give the fundamentals of matrix algebra, determinants, and systems of linear equations. At the end of the chapter we give some examples of situations in mathematics and other disciplines where matrices arise.

1.1 Matrices and matrix algebra

A *matrix* is a rectangular array of symbols. In this book the symbols will usually be either real or complex numbers. The separate elements of the array are known as the *entries* of the matrix.

Let m and n be positive integers. An m×n matrix A consists of m rows and n columns of numbers written in the following manner.

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

We often write $A = (a_{ij})$ for short. The entry a_{ij} lies in the i-th row and the j-th column of the matrix A.

Two mxn matrices $A = (a_{ij})$ and $B = (b_{ij})$ are *equal* if and only if all the corresponding entries of A and B are equal.

i.e. $a_{ij} = b_{ij}$ for each i and j.

The *sum* of the mxn matrices $A = (a_{ij})$ and $B = (b_{ij})$ is the mxn matrix denoted A + B which has entry $a_{ij} + b_{ij}$ in the (i,j)-place for each i,j.

Let λ be a scalar (i.e. a real or complex number) and let $A = (a_{ij})$ be an mxn matrix. The *scalar multiple* of A by λ is the mxn matrix denoted λA which has entry λa_{ij} in the (i,j)-place for each i,j.

1.1.2 **Proposition**

The following properties hold.

(i) A + B = B + A for all mxn matrices A and B, i.e. addition of matrices is commutative.

(ii) (A + B) + C = A + (B + C) for all mxn matrices A,B, and C, i.e.addition of matrices is associative.

(iii) $\lambda(A + B) = \lambda A + \lambda B$ for all scalars λ and all mxn matrices A and B.

(iv) $(\lambda_1 + \lambda_2)A = \lambda_1 A + \lambda_2 A$ for all scalars λ_1, λ_2 and all mxn matrices A.

(v) $(\lambda_1 \lambda_2)A = \lambda_1(\lambda_2 A)$ for all scalars λ_1, λ_2 and all mxn matrices A.

Proof

These properties follow at once from the properties of the real and complex number systems.

1.1.3 **Remark**

If we write -A for the matrix whose entries are $-a_{ij}$ for each i,j then -A = (-1)A, i.e. the multiple of the matrix A by the scalar -1. Also if we denote by O the nxn matrix with zero as each entry then A + (-A) = O.

1.1.4 Matrix multiplication

A $1 \times n$ matrix will be called a *row vector of length n* and an $m \times 1$ matrix will be called a *column vector of length m.*

Let $A = (a_1 \; a_2 \; a_3 \; . \; . \; . \; a_n)$ be a $1 \times n$ matrix and $B = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$ be an $n \times 1$ matrix.

We define the *product* AB to be the 1×1 matrix with the single entry

$$a_1b_1 + a_2b_2 + \; . \; . \; . \; . \; . + a_nb_n.$$

Now we will define matrix multiplication in general. We say that the product AB of the two matrices A and B is defined if and only if the number of columns of A equals the number of rows of B.

(i.e. AB is defined if and only if A is an $m \times n$ matrix and B is an $n \times p$ matrix for some integers m,n,p.)

We define the *matrix product* AB to be the $m \times p$ matrix which has as its (i,j)-entry

$$a_{i1}b_{1j} + a_{i2}b_{2j} + a_{i3}b_{3j} + \; . \; . \; . \; . \; . + a_{in}b_{nj}$$

(In shorthand notation the (i,j)-entry is $\sum_{k=1}^{n} a_{ik}b_{kj}$.)

In other words the (i,j)-entry of AB is the product of the i-th row of A with the j-th column of B, this product being as in the special case of $1 \times n$ and $n \times 1$ matrices defined above.

1.1.5 Example

Let $A = \begin{pmatrix} 1 & -1 & 6 \\ 4 & 1 & -2 \\ 3 & 2 & 0 \end{pmatrix}$, $B = \begin{pmatrix} -1 & 6 & 1 & 2 \\ 0 & -2 & 2 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix}$.

Then $AB = \begin{pmatrix} 1 & -1 & 6 \\ 4 & 1 & -2 \\ 3 & 2 & 0 \end{pmatrix} \begin{pmatrix} -1 & 6 & 1 & 2 \\ 0 & -2 & 2 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 5 & 14 & 5 & 7 \\ -6 & 20 & 4 & 7 \\ -3 & 14 & 7 & 8 \end{pmatrix}$.

1.1.6 **Proposition**

The following properties hold.

(i) (AB)C = A(BC) whenever these products are meaningful.

(i.e. matrix multiplication is associative).

(ii) A(B + C) = AB + AC for all m×n matrices A,B and all n×p matrices C.

(iii) (A +B)C = AC + BC for all m×n matrices A and B and all n×p matrices C.

Proof

(i) Let A,B,C be of sizes m×n, n×p, p×q respectively.

The (i,k)-entry of AB is $\sum_{r=1}^{n} a_{ir}b_{rk}$ and hence the (i,j)-entry of (AB)C is $\sum_{r=1}^{n}\sum_{k=1}^{p} a_{ir}b_{rk}c_{kj}$. An examination of the product A(BC) shows that exactly the same expression occurs as the (i,j)-entry of A(BC).

(ii) Let A be of size m×n, B and C of size n×p.

Then the (i,j)-entry of A(B + C) is $\sum_{k=1}^{n} a_{ik}(b_{kj} + c_{kj})$ and this is easily seen to equal the (i,j)-entry of AB + AC.

(iii) This follows in a similar manner to (ii).

1.1.7 **Remark**

Matrix multiplication is not in general commutative. Note first that for AB and BA to both be defined it is necessary that A and B are each n×n matrices for some integer n, i.e. square matrices of the same size. However AB and BA will be different in general.

1.1.8 **Exercise**

Let $A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix}$. Show that $AB \neq BA$.

1.1.9 **Remark**

A matrix of especial importance is the n×n *identity matrix*, denoted I_n, which is defined to have entries $a_{ii} = 1$ for all i and $a_{ij} = 0$ for $i \neq j$. Often when we are dealing with n×n matrices for a fixed value of n we will simply write I for the identity matrix omitting the suffix n.

For any m×n matrix A it is easy to see that $AI_n = A$ and that $I_m A = A$.

1.1.10 **The transpose of a matrix**

Let A be an m×n matrix.

The *transpose* of A is the n×m matrix with entry a_{ji} in the (i,j)-place. The transpose of A is denoted by A^t.

Note that the rows of A become the columns of A^t and vice versa.

1.1.11 **Proposition**

The transpose satisfies the following properties.

(i) $(A + B)^t = A^t + B^t$ for all m×n matrices A and B.

(ii) $(A^t)^t = A$ for all m×n matrices A.

(iii) $(AB)^t = B^t A^t$ whenever the product AB is defined.

Proof

Easy exercise.

Let A be an m×n matrix whose entries are complex numbers.

The *conjugate transpose* of A is the n×m matrix with entry $\bar{a}_{ji}$ in the (i,j)-place. The conjugate transpose is denoted $\bar{A}^t$.

The conjugate transpose satisfies the same three properties as those for the transpose given in (1.1.11).

1.1.12 **The trace of a square matrix**

Let A be an n×n matrix.

We define the *trace* of A by trace $A = \sum_{i=1}^{n} a_{ii}$.

The trace of A is a single real or complex number.

1.1.13 **Proposition**

The trace has the following properties.

(i) trace (A + B) = trace A + trace B for all n×n matrices A and B.

(ii) trace $(\lambda A) = \lambda$ trace A for all n×n matrices A and all scalars λ.

(iii) trace A^t = trace A for all n×n matrices A.

(iv) trace AB = trace BA for all n×n matrices A and B.

Proof

Easy exercise to prove (i),(ii), and (iii). To prove (iv) note that the (i,i)-entry of AB is $\sum_{j=1}^{n} a_{ij}b_{ji}$ which yields that trace AB $= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}b_{ji}$. Since both i and j are being summed from 1 to n this last double sum is symmetric in A and B and thus it must also give the value of trace BA.

Problems 1A

1. Let $A = \begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix}$, $B = \begin{pmatrix} 4 & -1 \\ 4 & 0 \end{pmatrix}$, $C = \begin{pmatrix} -1 & 2 \\ 2 & -1 \\ 1 & 3 \end{pmatrix}$, $D = \begin{pmatrix} 3 & 2 & 1 \\ 4 & -6 & 0 \\ 1 & -2 & -2 \end{pmatrix}$.

Calculate each of the following matrix products;

$$AB,\ CA,\ DC,\ DCAB,\ A^2,\ D^2,\ A^3B^2$$

2. Let $A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}$. Prove by induction that $A^n = \begin{pmatrix} 1 & n & n(n-1) \\ 0 & 1 & 2n \\ 0 & 0 & 1 \end{pmatrix}$.

3. Let A be an m×n matrix and B an n×p matrix. Let $B_1, B_2, \ldots, B_p$ denote the columns of B. Show that $AB_1, AB_2, \ldots, AB_p$ are the columns of AB.

If $A_1, A_2, \ldots, A_m$ denote the rows of A show that $A_1B, A_2B, \ldots, A_mB$ are the rows of AB.

4. Let A be an n×n matrix with entries in F. If AB = BA for all n×n matrices B with entries in F show that $A = \alpha I_n$ for some $\alpha \in F$, i.e. A is a scalar multiple of the identity matrix.

5. Let A be an n×n matrix with complex entries. If trace $\bar{A}^tA = 0$ show that A is the zero matrix.

(*Hint* - show that trace $\bar{A}^tA = \sum_{i=1}^{n}\sum_{j=1}^{n} |a_{ij}|^2$ where $|z|$ denotes the modulus of the complex number z.)

6. Let E_{ij} denote the n×n matrix with entry 1 in the (i,j)-place and zero elsewhere. Show that any n×n matrix $A = (a_{ij})$ is expressible in the form

$$A = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} E_{ij}.$$

Show also that $E_{ij}E_{kl} = 0$ if $j \neq k$, and $E_{ij}E_{jl} = E_{il}$.

7. Let the n×n matrix X be partitioned as follows ;

$X = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ where A is a p×p matrix, B is a p×q matrix, C is a q×p matrix, and D is a q×q matrix where p + q = n.

Let $Y = \begin{pmatrix} E & F \\ G & H \end{pmatrix}$ be an n×n matrix partitioned in a similar way. (i.e. E is a p×p matrix etc.)

Show that the product XY is partitioned as follows.

$$XY = \begin{pmatrix} AE + BG & AF + BH \\ CE + DG & CF + DH \end{pmatrix}$$

1.2 Systems of linear equations

A system of simultaneous linear equations

$$a_{11}x_1 + a_{12}x_2 + \ldots\ldots + a_{1n}x_n = b_1$$

$$a_{21}x_1 + a_{22}x_2 + \ldots\ldots + a_{2n}x_n = b_2$$

$$\vdots$$

$$a_{m1}x_1 + a_{m2}x_2 + \ldots\ldots + a_{mn}x_n = b_m$$

in n unknowns $x_1, x_2, \ldots, x_n$ can be rewritten as a single matrix equation $Ax = b$ where $A = (a_{ij})$ is an m×n matrix, $b = (b_i)$ is a column vector of length m, and $x = (x_i)$ is a column vector of length n.

We assume that the entries of A and b are real.

A *solution* of the system is an n-tuple of real numbers $(\alpha_1, \alpha_2, \ldots \alpha_n)$ such that $x_i = \alpha_i$ for each $i = 1,2,\ldots,n$ satisfies each of the m equations.

The *solution set* of the system is the set of all solutions of the system. It is a subset of $\mathbb{R}^n$. There are three possibilities for the solution set of the system;

(i) there is a unique solution, i.e. the solution set consists of a single point,

(ii) there are infinitely many solutions,

(iii) there are no solutions at all, i.e. the solution set is empty.

(In this case we say that the equations are *inconsistent.*)

For $m < n$ only possibilities (ii) and (iii) can occur whereas for $m \geq n$ all three possibilities can occur.

We illustrate this with a few simple examples;

1.2.1 **Example**

$$2x_1 + 3x_2 = 8$$
$$3x_1 - 3x_2 = 2$$

This system of two equations in two unknowns has the unique solution $x_1 = 2$, $x_2 = 4/3$.

Geometrically the two equations each represent a line in the plane and the solution set of the system is the point of intersection of the two lines.

1.2.2 **Example**

$$2x_1 + 3x_2 = 8$$
$$4x_1 + 6x_2 = 16$$

This system of two equations in two unknowns has infinitely many solutions. Specifically $x_1 = \alpha$, $x_2 = (8 - 2\alpha)/3$ for any $\alpha \in \mathbb{R}$ will be a solution.

Geometrically the two equations each represent the same line in the plane and the solution set of the system is the infinite set of all points on this line.

1.2.3 **Example**

$$2x_1 + 3x_2 = 8$$
$$4x_1 + 6x_2 = 3$$

This system of two equations in two unknowns has no solutions, the two equations being inconsistent.

Geometrically the two equations represent two parallel lines and so there are no points common to the two lines.

1.2.4 **Example**

$$x_1 + 2x_2 + x_3 = 3$$
$$x_1 - x_2 - x_3 = 2$$

Adding these two equations yields $2x_1 + x_2 = 5$. This gives $x_2 = 5 - 2x_1$.
Substituting into the first equation of the system then gives
$x_3 = 3 - x_1 - 2x_2 = 3 - x_1 - 2(5 - 2x_1) = 3x_1 - 7$.

Thus x_1 is free to take any real number value and x_2 and x_3 are then given in terms of x_1.

The solution set is $\{ (\alpha, 5 - 2\alpha, 3\alpha - 7) \; ; \; \alpha \in \mathbb{R} \}$.

Geometrically the two equations of the system each represent a plane in $\mathbb{R}^3$ and the solution set is the line of intersection of the two planes.

1.2.5 **Remark**

In this last example $m = 2$, $n = 3$, i.e. there are more unknowns than equations. In that situation a unique solution to the system cannot be expected. There is insufficient information to be able to obtain a unique value for the unknowns so that possibility (i) cannot occur.

Geometrically two equations in three unknowns represent two planes in $\mathbb{R}^3$. These two planes can either intersect in a line as in example (1.2.4) or else be parallel and so have no points of intersection, i.e. the solution set of the corresponding system is the empty set. Similar geometric considerations preclude the possibibility of there being a unique solution in the case of a system of m equations in n unknowns with $m < n$ for larger values of m and n.

1.2.6 Some special kinds of system

An $m \times n$ matrix is called an *echelon matrix* if it satisfies the following two conditions;

(i) the first non-zero entry in each row is one,

(ii) for any two consecutive rows either the lower of the two rows contains only zeros or else the first non-zero entry of the lower row is to the right of the first non-zero entry of the upper row.

Below are some examples of echelon matrices;

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

When A is an echelon matrix the system of equations $Ax = b$ is very easily solved by a method called *back substitution*. We illustrate this method by a few examples;

1.2.7 Example

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

Our equations are

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

Starting with the last equation $x_3 = 5$, we substitute into the equation above to obtain $x_2 = -1$, and then substitute into the equation above that to obtain $x_1 = -4$. This system thus has the unique solution $x_1 = -4$, $x_2 = -1$, $x_3 = 5$.

1.2.8 **Example**

$$\begin{pmatrix}1&2&1&2\\0&0&0&1\\0&0&0&0\\0&0&0&0\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}=\begin{pmatrix}4\\2\\0\\0\end{pmatrix}$$

Our equations are $x_1 + 2x_2 + x_3 + 2x_4 = 4,$

$$x_4 = 2.$$

Substituting $x_4 = 2$ into the equation above yields that $x_1 + 2x_2 + x_3 = 0$. Thus two of the variables x_1,x_2,x_3 can take any real number value and the third one is then determined in terms of these. If we take $x_1 = \alpha$, $x_2 = \beta$ where $\alpha \in \mathbb{R}$, $\beta \in \mathbb{R}$, then we can write $x_3 = -\alpha - 2\beta$. Our solution set may thus be written in the form $\{(\alpha,\beta,-\alpha - 2\beta,2);\ \alpha \in \mathbb{R},\beta \in \mathbb{R}\}$.

1.2.9 **Example**

$$\begin{pmatrix}1&3&4&1\\0&0&1&2\\0&0&0&0\\0&0&0&0\end{pmatrix}\begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix}=\begin{pmatrix}5\\8\\3\\0\end{pmatrix}$$

Our equations are $x_1 + 3x_2 + 4x_3 + x_4 = 5,$

$$x_3 + 2x_4 = 8,$$

$$0 = 3.$$

Clearly the last equation is impossible and so the system has no solution, i.e. the solution set is the empty set.

1.2.10 **Remark**

The solution sets in some of the above examples contain *free variables*, i.e. variables which may take any real number value, the other variables being given in terms of these free variables. In example (1.2.8) there are two free variables , in example (1.2.4) there is one free variable.

The number of free variables in the solution set is called the *dimension* of the solution set. (It is also sometimes called the *number of degrees of freedom.*) If there is a unique solution, i.e. the solution set is a single point of $\mathbb{R}^n$, then the dimension of the solution set is said to be zero.

1.2.11 **Elementary operations**

The following three kinds of operation on a set of simultaneous linear equations are called *elementary operations*;

(1) Interchange of a pair of equations.

(2) Multiplication of one equation by a non-zero scalar λ.

(3) Addition of λ times one equation to another equation, λ being a scalar.

1.2.12 **Definition**

Two systems of m simultaneous linear equations in n unknowns are *equivalent* if one system is obtainable from the other by a finite sequence of elementary operations.

1.2.13 **Proposition**

Two equivalent systems of simultaneous linear equations will have exactly the same solution set.

Proof

If the system $Cx = d$ is obtained from the system $Ax = b$ by a single elementary operation then any solution of $Ax = b$ will also be a solution of $Cx = d$. Since each of the three kinds of elementary operation is reversible it follows that $Ax = b$ and $Cx = d$ will have exactly the same solution set. The result follows immediately.

1.2.14 **Remark**

Our approach for solving a system of linear equations will be to transform the system into an equivalent one of the form $Cx = d$ where C is an echelon matrix. This system is then solved by back substitution.

1.2.15 The augmented matrix

Let $Ax = b$ be a system of m linear equations in n unknowns. The $(m + 1)\times n$ matrix $[A,b]$ obtained by adjoining the column b to the $m\times n$ matrix A is called the *augmented matrix* of the system.

1.2.16 Elementary row operations

The following three kinds of operation on a matrix are called *elementary row operations*;

(1) Interchange of two rows of a matrix.

(2) Multiplication of one row of the matrix by a non-zero scalar λ

(3) Addition of λ times one row of the matrix to another row, λ being a scalar.

Performing the elementary row operations of kind (1),(2), or (3) on the augmented matrix $[A,b]$ of a system of equations $Ax = b$ corresponds exactly to performing the elementary operations of (1.2.11) on the set of equations.

1.2.17 Comment

Each operation in (1.2.16) amounts to left multiplication by a certain matrix.

An operation of kind (1) where row r and row s are interchanged amounts to left multiplication by the $m\times m$ matrix E which has entries $e_{ii} = 1$ for all $i \neq r$, $i \neq s$, $e_{rs} = e_{sr} = 1$ and $e_{ij} = 0$ otherwise.

For $m = 3$, $r = 1$, $s = 3$, $E = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}$.

An operation of kind (2) where row r is multiplied by λ amounts to left multiplication by the $m\times m$ matrix E which has entries $e_{ii} = \lambda$ for $i = r$, $e_{ii} = 1$ for all $i \neq r$, and $e_{ij} = 0$ otherwise.

For $m = 3$, $r = 2$, $E = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

An operation of kind (3) where λ times row s is added to row r amounts to left multiplication by the matrix E with entries $e_{ii} = 1$ for all i, $e_{rs} = \lambda$, and $e_{ij} = 0$ otherwise.

For m = 3, r = 1, s = 2, $E = \begin{pmatrix} 1 & \lambda & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$.

The above three kinds of matrices E are known as *elementary matrices.*

1.2.18 **Gaussian elimination**

To solve a system of simultaneous linear equations Ax = b we perform a sequence of elementary row operations on the augmented matrix [A,b] to obtain [C,d] where C is an echelon matrix. It is not difficult to see that any m×n matrix A can be reduced to an echelon matrix by a sequence of elementary row operations. The argument goes as follows.

We may assume that the first column of A does not consist entirely of zeros since if it did the first variable would be redundant. By interchanging rows if necessary and multiplying by a suitable scalar we can ensure that 1 appears in the (1,1)-place. Then subtracting suitable multiples of row 1 from the other rows makes all other entries in column 1 equal to zero. Now look at the (n-1)x(n-1) matrix obtained by omitting row 1 and column 1. If the first column of this (n-1)x(n-1) matrix does not consist entirely of zeros we can, in the same manner as above, make the (1,1)-entry of this matrix equal to one and all the entries below it equal to zero. If the first column of the (n-1)x(n-1) matrix does consist entirely of zeros move to the second column. If this column does not consist entirely of zeros we can, as above, make it into a column with 1 on top and all zeros below. If the column does consist entirely of zeros move to the third column. Proceeding in this fashion we eventually must obtain an echelon form.

By (1.2.13) the solution set of the system Ax = b will be identical to the solution set of $Cx = d$ and since C is an echelon matrix we can easily find its solution set by back substitution.

The method we have described is called *the method of Gaussian elimination.*

We illustrate the method by a couple of examples;

1.2.19 **Example**

$$x_1 + 2x_2 + x_3 - x_4 = 0$$

$$2x_2 + 3x_3 + 3x_4 = 8$$

$$x_1 - x_2 - 3x_3 - 4x_4 = -8$$

$$x_1 + x_2 + 5x_3 - 2x_4 = -8$$

We write the augmented matrix as follows;

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

Subtracting row 1 from rows 3 and 4 yields

$$\left(\begin{array}{cccc|c} 1 & 2 & 1 & -1 & 0 \\ 0 & 2 & 3 & 3 & 8 \\ 0 & -3 & -4 & -3 & -8 \\ 0 & -1 & 4 & -1 & -8 \end{array}\right)$$

Adding twice row 4 to row 2 and subtracting three times row 4 from row 3 yields

$$\left(\begin{array}{cccc|c} 1 & 2 & 1 & -1 & 0 \\ 0 & 0 & 11 & 1 & -8 \\ 0 & 0 & -16 & 0 & 16 \\ 0 & -1 & 4 & -1 & -8 \end{array}\right)$$

Dividing row 3 by -16, subtracting 11 times the new row 3 from row 2, multiplying row 4 by -1 ,and interchanging rows yields the required echelon form.

$$\left(\begin{array}{cccc|c} 1 & 2 & 1 & -1 & 0 \\ 0 & 1 & -4 & 1 & 8 \\ 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 1 & 3 \end{array}\right)$$

Back substitution quickly gives the unique solution $x_4 = 3$, $x_3 = -1$, $x_2 = 1$, $x_1 = 2$.

1.2.20 **Example**

$$2x_1 - x_2 - x_3 = 2$$
$$x_1 + x_3 - 4x_4 = 1$$
$$x_2 - x_3 - 4x_4 = 4$$

The augmented matrix for this system is as follows;

$$\left(\begin{array}{cccc|c} 2 & -1 & -1 & 0 & 2 \\ 1 & 0 & 1 & -4 & 1 \\ 0 & 1 & -1 & -4 & 4 \end{array}\right)$$

Subtracting twice row 2 from row 1 yields

$$\left(\begin{array}{cccc|c} 0 & -1 & -3 & 8 & 0 \\ 1 & 0 & 1 & -4 & 1 \\ 0 & 1 & -1 & -4 & 4 \end{array}\right)$$

Adding row 3 to row 1 yields

$$\left(\begin{array}{cccc|c} 0 & 0 & -4 & 4 & 4 \\ 1 & 0 & -1 & -4 & 1 \\ 0 & 1 & -1 & -4 & 4 \end{array}\right)$$

Dividing row 1 by -4 and interchanging rows yields the desired echelon form

$$\left(\begin{array}{cccc|c} 1 & 0 & 1 & -4 & 1 \\ 0 & 1 & -1 & -4 & 4 \\ 0 & 0 & 1 & -1 & -1 \end{array}\right)$$

Back substitution quickly shows that there is one free variable and the solution set is $x_4 = \alpha$, $x_3 = \alpha - 1$, $x_2 = 5\alpha + 3$, $x_1 = 3\alpha + 2$, where $\alpha \in \mathbb{R}$.

1.2.21 **Comment**

Let C be an echelon matrix. If in the system $Cx = d$ we find that one row, say row r, of C consists entirely of zeros while the corresponding entry d_r of d is non-zero then the system will have no solution, i.e. the equations are inconsistent.

The method of Gaussian elimination we have described is equivalent to operating directly on the set of equations and successively eliminating variables. For computational purposes it is better to operate on the augmented matrix rather than on the equations themselves. When the number of equations and unknowns is large the procedure would be implemented on a computer.

1.2.22 <u>Comment</u>

We stated earlier that the coefficients in our systems of equations belonged to the field $\mathbb{R}$ of real numbers. In fact the above method for solving a system of linear equations works equally well if the coefficients belong to the field $\mathbb{C}$ of complex numbers or to the field $\mathbb{Q}$ of rational numbers. The solutions will of course then have values in $\mathbb{C}$ or $\mathbb{Q}$ respectively.

It should be noted that if all of the coefficients of a system are integers then the solutions will be rational numbers but not necessarily integers since fractions may occur in the reduction to echelon form.

Problems 1B

1. Solve by back substitution each of the following ;

(i)
$$\begin{aligned} 3x_1 + x_2 + 2x_3 - x_4 &= 2 \\ 2x_2 + 3x_3 + 4x_4 &= -2 \\ x_3 - 6x_4 &= 9 \\ x_4 &= 4 \end{aligned}$$

(ii)
$$\begin{aligned} 3w + 2x + 2y - z &= 2 \\ 2x + 3y + 4z &= -2 \\ y - 6z &= 6 \end{aligned}$$

2. Use Gaussian elimination to solve

$$\begin{pmatrix} 1 & 4 & 1 & 3 \\ 0 & 2 & 1 & 3 \\ 1 & 3 & 1 & 2 \\ 0 & 2 & 1 & 6 \end{pmatrix} \begin{pmatrix} w \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

3. Solve each of the following ;

(a)
$$w - x + y - z = 1$$

$$w + x - y - z = 1$$

$$w - x - y + z = 2$$

$$4w - 2x - 2y = 1$$

(b)

$$x + y + z = 1$$

$$x - y + 2z = 1$$

$$2x + 3z = 2$$

$$2x + 6y = 2$$

4. Show that the following system has a unique solution for all $k \neq -1$ but has no solution for $k = -1$.

$$x + y + kz = 1$$

$$x - y - z = 2$$

$$2x + y - 2z = 3$$

5. Solve each of the following ;

(a)

$$x - y - z = 1$$

$$2x - y + 2z = 7$$

$$x + y + z = 5$$

$$x - 2y - z = 0$$

(b)

$$x + 2y - z = 2$$

$$2x - y - 2z = 4$$

$$x + 12y - z = 2$$

6. Let A be an m×n matrix, x a column vector of length n, and let O denote the column vector of length m consisting entirely of zeros. Show that the system Ax = O is always consistent, i.e. the solution set is non-empty. (A system of the form Ax = O is called a *homogeneous system*.)

If the vector v satisfies Av = b show that the vector v + w will be a solution of Ax = b whenever w is a solution of the homogeneous system Ax = 0 Show that every solution of Ax = b is expressible in this form.

1.3 The inverse of a square matrix

Let A be an n×n matrix. If there exists an n×n matrix B such that $AB = BA = I_n$, the n×n identity matrix, then B is said to be an *inverse* of A.

1.3.1 Remark

If B_1 and B_2 are each inverses of A then $B_1 = B_2$. This is because $B_2 = B_2(AB_1) = (B_2A)B_1 = B_1$. Thus if an inverse of A exists it is unique and so we may talk about *the* inverse of A.

1.3.2 Notation and terminology

The inverse of A is denoted A^{-1}. We say that A is *non-singular* (or *invertible)* if the inverse of A exists. Otherwise we say that A is *singular*.

1.3.3 Proposition

The inverse has the following properties;

(i) $(A^{-1})^{-1} = A$ for all invertible n×n matrices A.

(ii) $(AB)^{-1} = B^{-1}A^{-1}$ for all invertible n×n matrices A and B.

(iii) $(A^t)^{-1} = (A^{-1})^t$ for all invertible n×n matrices A.

Proof

Easy exercise.

1.3.4 Lemma

Each of the three kinds of elementary matrix in (1.2.17) is invertible.

Proof

The inverse of an elementary matrix will be an elementary matrix of the same kind.

For an elementary matrix E of the first kind $E^2 = I$ and so $E^{-1} = E$. For an elementary matrix E of the second kind which has the non-zero scalar λ in the (i,i)-place E^{-1} will have λ^{-1} in the (i,i)-place.

For an elementary matrix of the third kind which has the scalar λ in the (r,s)-place E^{-1} will have $-\lambda$ in the (r,s)-place.

1.3.5 **Proposition**

Let A be an n×n matrix. The following three statements about A are equivalent ;

(i) A is invertible.

(ii) The system of linear equations Ax = b has a unique solution for some vector b.

(iii) A is expressible as a product of elementary matrices.

Proof

If A is invertible then the system Ax = b has the unique solution $x = A^{-1}b$. Thus (i) implies (ii).

To prove that (ii) implies (iii) note first that if the system of equations Ax = b has a unique solution then A must reduce to an echelon form with entries 1 at each point on the diagonal. (The *diagonal* of a square matrix is the set of all (i,i)-entries.) By a further set of elementary row operations we may reduce this echelon form to the identity matrix. Hence if the system Ax = b has a unique solution then there exist elementary matrices $E_1, E_2, \ldots E_r$ such that $E_1E_2 \cdot \cdot \cdot E_rA = I$. Multiply this equation on the left successively by $E_1^{-1}, E_2^{-1}, \ldots, E_r^{-1}$ and we obtain A as a product of elementary matrices by (1.3.4).

Suppose $A = E_1E_2 \cdot \cdot \cdot E_r$ where E_i, i = 1,2,....,r are elementary matrices. Then A is invertible by (1.3.3)(ii) and (1.3.4). This shows that (iii) implies (i).

1.3.6 **Remark**

If the system Ax = b has a unique solution for some vector b then it has a unique solution for every vector b. (The unique solution is of course given by $x = A^{-1}b$.)

1.3.7 Calculation of the inverse

Proposition (1.3.5) yields an effective way to calculate A^{-1}. Perform a sequence of elementary operations on A to reduce it to the identity matrix I. Then performing exactly the same sequence on I will yield A^{-1}. The following example illustrates this ;

1.3.8 Example

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

We write A and I alongside each other and perform the elementary operations on them together.

$$\begin{pmatrix} 1 & 3 & 1 \\ -1 & 1 & 2 \\ 2 & 1 & -2 \end{pmatrix} \Bigg| \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

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

$$\begin{pmatrix} 1 & 3 & 1 \\ 0 & 20 & 15 \\ 0 & -20 & -16 \end{pmatrix} \Bigg| \begin{pmatrix} 1 & 1 & 0 \\ 5 & 5 & 0 \\ -8 & 0 & 4 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 3 & 1 \\ 0 & 20 & 15 \\ 0 & 0 & -1 \end{pmatrix} \Bigg| \begin{pmatrix} 1 & 0 & 0 \\ 5 & 5 & 0 \\ -3 & 5 & 4 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 3 & 1 \\ 0 & 4 & 3 \\ 0 & 0 & 1 \end{pmatrix} \Bigg| \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 3 & -5 & -4 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 3 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{pmatrix} \Bigg| \begin{pmatrix} -2 & 5 & 4 \\ -8 & 16 & 12 \\ 3 & -5 & -4 \end{pmatrix}$$

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

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

We performed the following sequence of operations;

Add row 1 to row 2 and subtract twice row 1 from row 3

Multiply row 2 by 5 and multiply row 3 by 4

Add row 2 to row 3

Divide row 2 by 5

Subtract row 3 from row 1 and subtract 3 times row 3 from row 2

Divide row 2 by 4

Subtract 3 times row 2 from row 1.

We have thus found that $A^{-1} = \begin{pmatrix} 4 & -7 & -5 \\ -2 & 4 & 3 \\ 3 & -5 & -4 \end{pmatrix}$.

As a check on the calculation of A^{-1} the reader should always verify that $AA^{-1} = I$.

1.3.9 **Comment**

If we replace the word "row" by the word "column" everywhere in (1.2.16) we can define *elementary column operations*. Performing an elementary column operation on a matrix A amounts to multiplying A on the **right** by an elementary matrix. Elementary row operations occurred naturally in (1.2) in the manipulation of systems of linear equations whereas column operations would not have been appropriate there. However all we have said in this section (1.3) on calculating inverses can be done equally well using elementary column operations instead of row operations. It is important though not to mix row and column operations, i.e. to find A^{-1} we either reduce A to the identity matrix I using only elementary row operations or else reduce A to I using only elementary column operations.

1.4 Determinants

For each square matrix A with entries in F, F = $\mathbb{R}$ or $\mathbb{C}$, we can associate a single element of F called the *determinant* of A and denoted det A for short.

1.4.1 Definition

Let $A = (a_{11})$ be a 1x1 matrix. We define $\det A = a_{11}$.
Let $A = (a_{ij})$ be a 2x2 matrix. We define $\det A = a_{11}a_{22} - a_{12}a_{21}$.
Now assume that the determinant has been defined for all $(n-1)\times(n-1)$ matrices and let $A = (a_{ij})$ be an nxn matrix.
We denote by M_{ij} the determinant of the $(n-1)\times(n-1)$ matrix obtained from A by deleting row i and column j of A. We call M_{ij} the *minor* of A corresponding to the entry a_{ij} of A.

We now define the determinant of the nxn matrix A by

$$\det A = \sum_{j=1}^{n} (-1)^{j+1} a_{1j} M_{1j}.$$

Note that for n = 2 this reduces to the definition already given and for n = 3 we have the formula

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

The main properties of determinants are contained in the following proposition;

1.4.2 Proposition

(1) $\det A = \sum_{j=1}^{n} (-1)^{j+1} a_{j1} M_{j1}$.

(2) $\det A = \det A^t$ for all square matrices A.

(3) If A is upper triangular, i.e. $a_{ij} = 0$ for all $i > j$,

then $\det A = a_{11}a_{22}a_{33}......a_{nn}$.

If A is lower triangular, i.e. $a_{ij} = 0$ for all $i < j$, then det A = $a_{11}a_{22}a_{33}......a_{nn}$.

(4) If one row of A consists entirely of zeros then det A = 0.

(5) If B is the matrix obtained from A by multiplying one row of A by the scalar λ then det B = λ det A.

(6) If B is the matrix obtained from A by interchanging two rows of A then det B = -det A.

(7) If two rows of A are identical then det A = 0.

(8) If B is the matrix obtained from A by adding λ times one row of A to another row of A then det B = det A.

(9) Properties (4), (5), (6), (7),and (8) remain valid if the word "row" is replaced everywhere by the word "column".

(10) det AB = (det A)(det B) for any pair of n×n matrices A and B.

Proof

The proof of all these properties is rather long and is left until the appendix to this chapter.

1.4.3 **Remark**

The formula in our definition of det A is often called *the expansion along the first row* because the entries a_{1j}, j = 1,2,...n of row 1 appear in the formula.

Property (1) shows that det A can equivalently be obtained via an *expansion down the first column.*

Evaluation of det A from the basic definition involves a lot of calculation even if n is as small as 4. It is easier to reduce A to upper or lower triangular form by row operations and utilize the properties given in (1.4.2). We illustrate this by the following example;

1.4.4 **Example**

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

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

after subtracting 3 times row 2 from row 1, subtracting row 2 from row 3, and adding twice row 2 to row 4. This uses property (8).

Adding row 1 to row 4 and interchanging row 1 and row 2 then yields , using properties (6) and (8), that

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

Subtracting 5 times row 3 from row 2 and then interchanging row 2 with row 3 yields, using (8) and (6) again, that

$$\det A = \begin{pmatrix} 1 & 2 & 0 & 1 \\ 0 & -1 & 2 & -2 \\ 0 & 0 & -9 & 9 \\ 0 & 0 & 0 & 4 \end{pmatrix}$$

Property (3) now gives det A = 36.

1.4.5 **Remark**

It is possible to give a more sophisticated definition of the determinant and some textbooks do this. Their definition is equivalent to ours. However the definition of (1.4.1) and knowledge of the properties of (1.4.2) are sufficient for our purposes.

1.4.6 **The adjugate of a matrix**

Let $A = (a_{ij})$ be an nxn matrix. The *cofactor* C_{ij} of A corresponding to the entry a_{ij} is defined by $C_{ij} = (-1)^{i+j}M_{ij}$ where M_{ij} is the minor as defined in (1.4.1).

Note that $\det A = \sum_{j=1}^{n} a_{1j}C_{1j}$.

The *adjugate of A*, denoted adj A, is the n×n matrix which has entry C_{ji} in the (i,j)-place. Thus adj A is the transpose of the matrix of cofactors of A.

To obtain adj A we first find the matrix of minors M_{ij}, then multiply alternate entries of this matrix by -1, and finally transpose the matrix. The following diagram may be helpful in remembering where to multiply by -1.

$$\begin{pmatrix} + & - & + & - & . & . \\ - & + & - & + & . & \\ + & - & + & - & . & \\ - & + & - & + & . & \\ . & . & . & . & . & \\ . & & & & & \end{pmatrix}$$

For example if $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ then adj $A = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

1.4.7 **Proposition**

Let A be an n×n matrix and I the identity n ×n matrix. Then

A(adj A) = (adj A)A = (det A)I

Proof

See the appendix to this chapter.

1.4.8 **Proposition**

The n×n matrix A is invertible if and only if det A $\neq$ 0.

Proof

If det A $\neq$ 0 then A^{-1} exists and equals $(\det A)^{-1}$(adj A) by (1.4.7).

Conversely if A^{-1} exists then property (10) of (1.4.2) applied with $B = A^{-1}$ implies that (det A)(det A^{-1}) = det I = 1 and in particular det A cannot be zero.

1.4.9 **Calculation of the inverse via the adjugate**

We have seen in (1.3.7) and (1.3.8) one method of calculating A^{-1}. Proposition (1.4.7) yields a different method which is very quick for 2×2 and 3x3 matrices. For larger matrices the calculations become very cumbersome and the method of (1.3.7) for finding inverses is more efficient from a computational viewpoint.

Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Then A^{-1} exists provided $ad - bc \neq 0$ and

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

1.4.10 **Example**

Show that $A = \begin{pmatrix} 1 & 0 & 2 \\ 2 & 3 & 1 \\ 1 & 2 & 1 \end{pmatrix}$ is invertible and find A^{-1}.

The matrix of minors of A is $\begin{pmatrix} 1 & 1 & 1 \\ -4 & 1 & 2 \\ -6 & 3 & 3 \end{pmatrix}$ and det A = 3.

Hence $A^{-1} = (1/3)\text{adj } A = (1/3)\begin{pmatrix} 1 & 4 & -6 \\ -1 & -1 & 3 \\ 1 & -2 & 3 \end{pmatrix}$.

1.4.11 **Exercise**

For the matrix A of Example (1.3.8) verify that the above method yields the same value for A^{-1} as was obtained in (1.3.8).

Problems 1C

1. Use the method of 1.3 to find the inverse of each of the following;

$$\begin{pmatrix} -1 & 2 & 1 \\ 2 & 1 & -1 \\ 1 & 9 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 2 & 2 & 2 \\ 0 & 2 & 2 \\ 0 & 0 & 2 \end{pmatrix}.$$

2. Let $X = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a & -b & -c \end{pmatrix}$. Show that $aI + bX + cX^2 = O$. Deduce that X is invertible if $a \neq 0$ and that $X^{-1} = -(1/a)(bI + cX)$.

3. Let $X = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ be partitioned as in problem 7 of problems 1A. If X is invertible and if all the relevant inverses exist show that

$$X^{-1} = \begin{pmatrix} (A - BD^{-1}C)^{-1} & A^{-1}B(CA^{-1}B - D)^{-1} \\ (CA^{-1}B - D)^{-1}CA^{-1} & (D - CA^{-1}B)^{-1} \end{pmatrix}.$$

4. Calculate the determinant of each of the following matrices;

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

5. Show that $\det \begin{pmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{pmatrix} = (a - b)(b - c)(c - a)$.

Show that $\det \begin{pmatrix} 1 & 1 & 1 & & 1 \\ a_1 & a_2 & a_3 & & a_{n+1} \\ a_1^2 & a_2^2 & a_3^2 & & a_{n+1}^2 \\ & & & & \\ a_1^n & a_2^2 & a_3^2 & & a_{n+1}^2 \end{pmatrix} = \prod_{i<j} (a_i - a_j)$,

the product over all a_i, a_j with $i < j$.

(Matrices of the above type are called *Vandermonde matrices*.)

6. Let $A = \begin{pmatrix} a & b & c \\ c & a & b \\ b & c & a \end{pmatrix}$. Show that det A can be factorized in the form

$(a + b + c)(a + b\omega + c\omega^2)(a + b\omega^2 + c\omega)$ where $\omega = \exp(2\pi i/3)$, a primitive

cube root of unity.

Obtain a similar factorization for $\det \begin{pmatrix} a & b & c & d \\ d & a & b & c \\ c & d & a & b \\ b & c & d & a \end{pmatrix}$ which involves a

primitive fourth root of unity.

(Matrices of the above type are called *circulant matrices*.)

7. Let A,B, and C be $n \times n$ matrices and suppose that, for each $i = 2, 3, . . ,n$,

we have row i of A = row i of B = row i of C.

Suppose also that row 1 of C = (row 1 of A) + (row 1 of B).

Show that det C = det A + det B.

8. Determine whether any of the following matrices are invertible. When the inverse exists use (1.4.9) to calculate it.

$$\begin{pmatrix} 1 & -1 & -2 \\ 0 & 3 & 4 \\ 0 & 0 & 2 \end{pmatrix}, \quad \begin{pmatrix} 5 & 3 & 1 \\ 8 & 1 & 2 \\ 3 & -2 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 2 & -2 \\ 2 & 1 & 1 & -1 \\ 1 & 0 & 0 & 2 \\ -1 & 1 & 1 & -1 \end{pmatrix}.$$

9. Let A be an n×n matrix whose entries are integers and suppose det A = ±1. Show that all the entries of A^{-1} are integers.

10.. Let $X = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ be partitioned as in problem 7 of problems 1A.

(i) If either B = 0 or C = 0 show that det X = (det A) (det D).

(ii) Instead suppose that A is invertible. By multiplying X on the right by a suitable matrix of the form $\begin{pmatrix} I & Y \\ 0 & I \end{pmatrix}$, I being the identity matrix of the appropriate size, deduce that det X = (det A) (det $(D - CA^{-1}B)$).

1.5 Some special types of matrices

Let $A = (a_{ij})$ be an n×n matrix with real or complex entries.

A is said to be *diagonal* if $a_{ij} = 0$ for all $i \neq j$.

A is said to be *tridiagonal* if $a_{ij} = 0$ for all i,j with $|i - j| > 1$. i.e. all the entries of A are zero except possibly for those on the main diagonal $(a_{11}, a_{22}, \ldots, a_{nn})$, the superdiagonal $(a_{12}, a_{23}, \ldots, a_{(n-1)n})$, and the subdiagonal $(a_{21}, a_{32}, \ldots, a_{n(n-1)})$.

A is said to be *upper triangular* if $a_{ij} = 0$ for all $i > j$, i.e. all entries below the main diagonal are zero.

A is said to be *lower triangular* if $a_{ij} = 0$ for all $i < j$, i.e. all entries above the main diagonal are zero.

A is said to be *upper Hessenberg* if $a_{ij} = 0$ for all $i > j + 1$, i.e. all entries above the superdiagonal are zero.

A is said to be *lower Hessenberg* if $a_{ij} = 0$ for all $j > i + 1$, i.e. all entries below the subdiagonal are zero.

A is said to be *symmetric* if $a_{ij} = a_{ji}$ for all i,j, i.e. $A^t = A$.

A is said to be *hermitian* if the entries of A are complex and $a_{ij} = \bar{a}_{ji}$ for all i,j, i.e. $\bar{A}^t = A$.

(The name hermitian comes from the French mathematician Hermite.)

A is said to be *orthogonal* if all the entries of A are real and $A^t = A^{-1}$

A is said to be *unitary* if the entries of A are complex and $\bar{A}^t = A^{-1}$.

A is said to be *block diagonal* if A is of the form

$$\begin{pmatrix} A_{11} & O & O & O \\ O & A_{22} & O & O \\ & & & \\ & & & \\ O & O & & A_{kk} \end{pmatrix}$$

where A_{ii} is an $n_i x n_i$ matrix for each i = 1,2, . . . ,k.

A is said to be a *permutation matrix* if A is obtained from the identity matrix I by a permutation of the rows of I.

Problems 1D

1. If the matrix A is symmetric and invertible show that A^{-1} is symmetric. If A is hermitian and invertible show that A^{-1} is hermitian.
2. If the matrix A is upper triangular and invertible show that A^{-1} is upper triangular. If A is lower triangular and invertible show that A^{-1} is lower triangular.

1.6 More on systems of linear equations

1.6.1 Cramer's rule

Let A be an invertible nxn matrix and Ax = b be a system of linear equations. Cramer's rule is an attractive way of describing the solution set of the system. It says that the system has a unique solution given by

x_i = (det A_i)/det A for each i = 1, 2, . . . , n where A_i is the n×n matrix obtained from A by replacing the i-th column of A by the column vector b.

For a proof of Cramer's rule see problem 1 at the end of this section.

Cramer's rule is mainly of theoretical importance. It is not used much in practice for determining the solution of systems of linear equations because the numerical computations involved are much greater than those required for Gaussian elimination.

1.6.2 The LU-decomposition of a square matrix

Let A be an n×n matrix. An expression of A in the form A = LU where L and U are respectively lower and upper triangular n×n matrices is called *an LU-decomposition of A*.

An LU-decomposition of A need not necessarily exist and even if it does exist it is not unique. However for any n×n matrix A there do exist permutation matrices P and Q, a lower triangular matrix L and an upper triangular matrix U such that A = PLUQ. Further it can be shown that if A is invertible then A = PLU, i.e. Q can be chosen equal to I.

Observe that if we have A = LU then the system Ax = b can be very easily solved in the following manner;

First solve Ly = b for unknowns y = (y_i) by forward elimination, (i.e. by solving equation 1 first, then equation 2 and so on). Then solve Ux = y for unknowns x = (x_i) by back substitution as in (1.2).

More generally if A = PLUQ then we solve LUz = $P^{-1}b$ in the above manner, z = Qx being a re-arrangement of the unknowns (x_i) and $P^{-1}b$ being a re-arrangement of the co-ordinates of b.

The method of Gaussian elimination as in (1.2.18) can be interpreted as yielding decompositions of the form A = LU or more generally A = PLUQ. We will explain this in more detail. Suppose first that we are able to reduce A to an echelon form U by a finite sequence of elementary row operations of kind (2) and (3) only and that the only ones of kind (3) involve adding a scalar multiple of row i to row j where $i < j$. Then the corresponding elementary matrices, see (1.2.17), will all be lower triangular and we have $E_r E_{r-1} \ldots\ldots E_1 A = U$ where U is an echelon matrix which is upper triangular and each E_i is an lower triangular elementary matrix.

Thus A = LU where $L = E_1^{-1} E_2^{-1} \ldots E_r^{-1}$ is lower triangular.

L is lower triangular because the inverse of an lower triangular matrix is lower triangular, (see problem 2 of problems 1D), and the product of lower triangular matrices is lower triangular.

For a general $n{\times}n$ matrix A the reduction to echelon form by elementary row operations cannot be accomplished without using the interchange operations of kind (1). If A is invertible it can be shown that, by permuting the rows of A in suitable fashion, the resulting matrix QA, Q being a permutation matrix, is expressible in the form QA = LU. Hence A = PLU where $P = Q^{-1}$. If A is singular an LU-decomposition cannot be achieved without also permuting the columns of A. Permuting the columns of A amounts to right multiplication by a permutation matrix and so in this case the best we can do is to write A = PLUQ.

Problems 1E

1. Prove Cramer's rule as stated in (1.6.1).

(*Hint*-show that the i-component of (adj A)b is $\sum_{j=1}^{n} C_{ji} b_j$ where C_{ji} denotes the cofactor corresponding to a_{ji} and use the properties of determinants to show that this sum also equals det A_i.)

2. Find the unique value of k for which the following system of linear

equations does not have a unique solution and determine the solution set in this case;

$$2x + y + z = 3$$

$$3x - y - z = 7$$

$$6x + y + kz = 11$$

3. Use the method described in *(1.6.2)* to solve the following;

$$\begin{pmatrix} 3 & -2 & -2 \\ 0 & 6 & -3 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ -1 & 6 & 2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \\ 2 \end{pmatrix}$$

4. Obtain an LU-decomposition for $\begin{pmatrix} 1 & 2 & 4 \\ -1 & -3 & 3 \\ 4 & 9 & 14 \end{pmatrix}$.

5. Obtain a PLU-decomposition for $\begin{pmatrix} 0 & -6 & 4 \\ 2 & 1 & 2 \\ 1 & 4 & 1 \end{pmatrix}$.

1.7 Some places where matrices are found

(a) Matrices in linear algebra

Matrices are a vital and essential part of the area of mathematics called *linear algebra* as we shall see fully in the next chapter. For the moment let us just observe that a mapping $f : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ given by $f(x,y) = (ax + by, cx + dy)$, where $\mathbb{R}^2$ denotes the set of all ordered pairs of real numbers and a,b,c, and d are constants, may be written succinctly in matrix form as $f(v) = Av$ where $v = \begin{pmatrix} x \\ y \end{pmatrix}$ and $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

(b) The Jacobian matrix in calculus of several variables

Let $x_1, x_2, \ldots, x_r$ be a set of real variables and let $u_1, u_2, \ldots, u_r$ be a new set of variables. The $r \times r$ matrix with entries the partial derivatives $\partial x_i / \partial u_j$ is called the *Jacobian* matrix of change of variables.

This matrix is important in the differential calculus of several variables and its determinant is important in integration theory of functions of several variables.

(c) The Hessian matrix in calculus of several variables

Let $f(x_1,x_2,\ldots,x_n)$ be a function of n real variables. The *Hessian* matrix of f is the matrix which has entry $\partial^2 f/\partial x_i \partial x_j$, the second order partial derivative, in the (i,j)-place. This matrix is used in the analysis of the critical points of f.

(d) The Wronskian in differential equations

Consider the differential equation of order n

$$f^{(n)} + a_{n-1}f^{(n-1)} + \ldots\ldots + a_1 f^{(1)} + a_0 f = 0$$

where a_i, $i = 0,1,\ldots,n-1$ are constants and $f^{(i)}$ denotes the i-th derivative of the function f.

Let $f_1,f_2,\ldots,f_n$ be a set of functions each of which satisfies the above differential equation.

The determinant of the n×n matrix which has entry $f_j^{(i-1)}$ in the (i,j)-place for each $j = 1, 2,\ldots,n$, $i = 0, 1,\ldots, n-1$ is called the *Wronskian* of the set and provided it is non-zero the functions $f_1,f_2,\ldots f_n$ form an independent set of solutions to the differential equation.

(e) Matrices in control theory

Engineers of various kinds (electrical, mechanical, chemical etc.) deal with feedback control systems. A large class of such systems can be described mathematically by the following equations involving matrices;

$$\dot{x} = Ax + Bu, \; y = Cx$$

Here $x = (x_i)$, $u = (u_i)$, $y = (y_i)$ are column vectors of length n, m, p respectively, and A,B,C are matrices of size n×n, n×m, p×n respectively. These three vectors are each functions of time and $\dot{x}$ denotes the column vector of derivatives of x.

The vectors x,u,y are known as the *state vector*, the *input vector*, and the *output vector* respectively. The matrices A, B, C are known as the *system matrix*, the *input distribution matrix*, and the *measurement matrix* respectively. These three matrices are assumed to be independent of time.

(f) Matrices in the theory of graphs and networks

The abstract mathematical notion of a graph has applications in many practical situations, (e.g. electrical circuits, communications networks, transportation problems in management science, molecular structure in organic chemistry etc).

A *graph* consists of a finite set of vertices (also called *nodes)* and a finite set of edges (also called *branches*) such that each edge consists of a distinct pair of distinct vertices.

A *directed graph* is a graph such that each edge consists of an ordered pair of distinct vertices, i.e. each edge has a preferred direction.

(The graphs modelling electrical circuits have the connection points as vertices and pieces of the circuit consisting of an e.m.f source and a resistance as the edges. The graphs modelling chemical molecules have atoms as vertices and chemical bonds as the edges. The graphs modelling transportation problems could have different cities as vertices and routes between these cities as the edges.)

To any graph there is associated a matrix known as its *adjacency matrix* defined as follows;

We label the vertices V_1, V_2 etc. and the edges E_1, E_2 etc.

If the graph has n vertices then its adjacency matrix is the $n \times n$ symmetric matrix with entry 1 in the (i,j)-place whenever vertex V_i and vertex V_j form an edge and entry zero in all other places.

To any directed graph there is another matrix known as its *incidence matrix* defined as follows;

If the graph has n vertices and p edges then its incidence matrix is the $n \times p$ matrix whose k-th column has -1 in the i-th place and +1 in the j-th place where edge E_k is the ordered pair of vertices (V_i, V_j). All other entries of the k-th column are zero.

The directed graph in Figure 1.1 below has adjacency matrix

$$\begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix} \text{ and incidence matrix } \begin{pmatrix} 1 & -1 & 0 & 0 & 0 \\ -1 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & -1 & 0 & -1 \end{pmatrix}$$

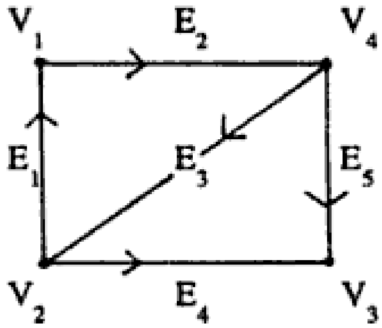

Figure 1.1

The adjacency matrix and the incidence matrix are both useful in the study of graphs and networks.

(g) Matrices in probability theory

The republic of Lowland has regular elections once a year and has three political parties labelled X,Y, and Z. The inhabitants of Lowland behave in a consistent fashion at elections and never vote for the same party at two successive elections. Those who vote for party X or party Y at one election are equally likely to vote next time for either of the two other parties. Those who vote for party Z at one election are twice as likely to vote next time for party X as to vote for party Y.

We construct a 3x3 matrix A as follows;

In the first column of A we place the probabilities that a person who voted for party X at one election will vote for parties X,Y,Z respectively at the next election. In the second (resp. third) column of A we place the probabilities that a person who voted for party Y (resp. Z) at one election will vote for parties X,Y,Z next time. The sum of the entries in each column of A equals one. The matrix A is called a *transition matrix*.

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

Suppose that in the first ever election in Lowland each of the three parties gets the same number of votes, i.e. they each get one third of the total vote. The column vector $\begin{pmatrix} 1/3 \\ 1/3 \\ 1/3 \end{pmatrix}$ is called the *initial state vector*. The state of the parties after n elections will be described by the vector $A^n v$ where v is the initial state vector.

The above is an example of a *Markov process*. In chapter 5 we will see more about Markov processes.

(h) The covariance matrix in statistics

Let $X_1, X_2, \ldots, X_n$ be a set of random variables and let $E(X_i) = \mu_i$ for each i where E denotes expectation. Statisticians make use of the *covariance matrix* of the vector of random variables $(X_1, X_2, \ldots, X_n)$. This is the nxn matrix which has entry $E[(X_i - \mu_i)(X_j - \mu_j)]$ in the (i,j)-place.

APPENDIX TO CHAPTER 1

A proof of the properties of determinants

In this appendix we derive the properties of determinants listed in Proposition (1.4.2) and also prove Proposition (1.4.7). First we state and give the proof of each of the ten properties in (1.4.2).

Proposition 1.4.2

(1) $\det A = \sum_{j=1}^{n} (-1)^{j+1} a_{j1} M_{j1}$ where M_{ij} denotes the minor of the n×n matrix A corresponding to the entry a_{ij} of A. (See (1.4.1) for the definition of minor.)

i.e. det A can equivalently be obtained by expansion down the first column.

Proof

The proof is by induction on n.

It is easy to see that the result is true for n = 2 so let A be an n×n matrix and assume the result is true for all (n-1)×(n-1) matrices. Expansion of det A along the first row gives

$$\det A = \sum_{i=1}^{n} (-1)^{j+1} a_{1j} M_{1j}.$$

By the inductive assumption we can expand M_{1j} down the first column to get $M_{1j} = \sum_{i=2}^{n} (-1)^{i-1+1} a_{i1} d_{ij}$ where d_{ij} denotes the determinant of the (n-2)×(n-2) matrix obtained from A by deleting row 1, row i, column 1, and column j. Hence the only term in det A involving $a_{1j}a_{i1}$ is $(-1)^{i+j+1} a_{1j} a_{i1} d_{ij}$

Also by the inductive assumption $M_{i1} = \sum_{j=2}^{n} (-1)^{i+j+1} a_{1j} d_{ij}$. Thus, examining the expansion $\sum_{i=1}^{n} (-1)^{i+1} a_{i1} M_{i1}$, we see that the only term involving $a_{i1}a_{1j}$ is $(-1)^{i+j+1} a_{i1} a_{1j} d_{ij}$. This proves (1).

(2) $\det A = \det A^t$ for all $n{\times}n$ matrices A.

Proof

Observe that the minor M_{ij} of A will equal the (j,i)-minor of the matrix A^t. Hence, after expanding along the first row, we see that $\det A^t = \sum_{j=1}^{n}(-1)^{j+1}a_{j1}M_{j1}$ and by property (1) we see that det A is given by exactly the same expression. This proves (2).

(3) If A is upper or lower triangular then $\det A = a_{11}a_{22}....a_{nn}$.

Proof

The proof is by induction on n.

The result is clearly true for n = 1 so let A be an $n{\times}n$ matrix and assume the result is true for all $(n-1){\times}(n-1)$ matrices. If A is lower triangular expanding along the first row yields that $\det A = a_{11}M_{11} = a_{11}a_{22}......a_{nn}$ by the inductive assumption. If A is upper triangular expanding down the first column yields a similar result. This proves (3).

(4) If one row of A consists entirely of zeros then det A = 0.

Proof

The proof is by induction on n.

If n = 1 the result is clearly true so let A be an $n{\times}n$ matrix and assume the result is true for all $(n-1){\times}(n-1)$ matrices A.

Now $\det A = \sum_{j=1}^{n}(-1)^{j+1}a_{1j}M_{1j}$. If row 1 of A consists entirely of zeros then clearly det A = 0. If row i for some i > 1 consists entirely of zeros then each minor M_{1j} will be zero by the inductive assumption so that det A = 0. This proves (4).

(5) If B is the matrix obtained from A by multiplying one row of A by the scalar λ then $\det B = \lambda \det A$.

Proof

The proof is by induction on n.

The result is clearly true for $n = 1$ so let A be an $n \times n$ matrix and assume the result is true for all $(n-1)\times(n-1)$ matrices. If row 1 of A is multiplied by λ then each term in the expansion along the first row is multiplied by λ so that $\det B = \lambda \det A$.

If row i for some $i > 1$ is multiplied by λ then each minor appearing in the expansion along the first row will be multiplied by λ because of the inductive assumption. This gives $\det B = \lambda \det A$ again and so (5) is proved.

(6) If B is the matrix obtained from A by interchanging two rows of A then $\det B = -\det A$.

Proof

Note first that the interchange of row i and row j can be achieved by successively interchanging an odd number of *adjacent* rows. (Specifically moving row i to below row j takes $i - j$ interchanges of adjacent rows and then shifting row j into the original position of row i takes a further $i - j - 1$ interchanges.) Thus it suffices to prove the result in the case when two adjacent rows are interchanged.

Now suppose we interchange row i and row $i + 1$ of A. Write M_{ij} for the minors of A and R_{ij} for the minors of B. We observe that $M_{k1} = - R_{k1}$ for all $k \neq i, i + 1$, that $a_{i1}M_{i1} = b_{(i+1)1}R_{(i+1)1}$, and that $b_{i1}R_{i1} = a_{(i+1)1}M_{(i+1)1}$.

Examining the expansions of det A and det B down the first column we see that $\det B = - \det A$. This proves (6).

(7) If two rows of A are identical then det A = 0.

Proof

Interchanging the two identical rows yields det A = -det A by property (6) and hence det A = 0.

(8) If B is the matrix obtained from A by adding λ times one row of A to another row of A then det B = det A.

Proof

Let B be the $n \times n$ matrix having the same rows as A except for having λ times row i of A plus row j of A as its j-th row.

Let C be the $n \times n$ matrix having the same rows as A except for having λ times row i of A as its j-th row.

Observe first that det B = det A + det C. (This can be seen by interchanging row j with row 1 for each of the three matrices A, B, C, finding det A, det B, det C by expansion along the first column, and using property (5).)

But det C = 0 by properties (5) and (7) so that det B = det A.

(9) Properties (4), (5), (6), (7), and (8) remain valid if the word "row" is replaced everywhere by the word "column".

Proof

The columns of A are the rows of A^t and vice versa. The result now follows because of property (2).

(10) det AB = det A det B for any pair of $n \times n$ matrices A and B.

Proof

Assume first that A is non-singular. Then the argument in the proof of (1.3.5) shows that A is reducible to the identity matrix I by a sequence of elementary row operations. If we use only operations of type (1) and (3) we can reduce A to a diagonal matrix D with all of the diagonal entries of D being non-zero. Let r be the number of operations of type (1) used. (i.e. r is the number of row interchanges). Then, using properties (6) and (8), $\det A = (-1)^r \det D$.

Also $\det AB = (-1)^r \det DB$ since exactly the same sequence of elementary row operations transforms AB into DB.

Now let the diagonal entries of D be $\alpha_1, \alpha_2, \ldots, \alpha_n$. Multiplying B on the left by D amounts to multiplying row i of B by α_i for each $i = 1,2, \ldots, n$. Then by property (5) we see that $\det DB = \alpha_1\alpha_2 \ldots \alpha_n \det B = \det D \det B$. Thus $\det AB = (-1)^r \det D \det B = \det A \det B$.

Now suppose A is singular, i.e. A does not have an inverse. If also B is singular then the system of equations $Bx = 0$ has a non-zero solution. ($x = 0$ is always one solution of $Bx = 0$ and by (1.3.6) the solution of $Bx = 0$ is unique if and only if B is non-singular.) Hence the system $ABx = 0$ has a non-zero solution which implies that AB must be singular.

If B is non-singular then again AB must be singular for if AB is non-singular then (1.3.5) gives that $AB = E_1E_2 \ldots E_r$ for elementary matrices E_i, $i = 1,2, \ldots, r$ which implies that A is non-singular. ($A = E_1E_2 \ldots E_rB^{-1}$ which is non-singular by (1.3.3)(ii) and (1.3.4).)

The above argument shows that AB is singular whenever A is singular.

Now any singular matrix X must have zero determinant. (If X is singular then X is reducible to an echelon form E with a zero somewhere on the diagonal. Properties (5), (7), and (8) imply that $\det E = \alpha \det X$ for some non-zero scalar α, and det E = 0 by property (3).)

Thus whenever A is singular the equation det AB = det A det B is valid because each side of the equation is zero. This completes the proof of (10).

Proposition 1.4.7

Let A be an $n \times n$ matrix and let I be the $n \times n$ identity matrix.
Then A (adj A) = (adj A) A = (det A) I.

Proof

Let $A = (a_{ij})$ have minors M_{ij} and cofactors $C_{ij} = (-1)^{i+j}M_{ij}$. Since adj A is the transpose of the matrix of cofactors of A the (i,j)-entry of the product A (adj A) will be $a_{i1}C_{j1} + a_{i2}C_{j2} + \ldots\ldots + a_{in}C_{jn}$.

For i = j = 1 the above sum will equal det A as it is precisely the definition of det A by expansion along the first row. Now suppose i > 1 and i = j. Let B be the matrix obtained from A by moving row i above row 1 and leaving all other rows fixed. B is obtained from A by successively interchanging row i with rows i-1, i-2, . . .,1. Hence det B = $(-1)^{i-1}$det A using (6) of (1.4.2). Note that the (1,k)-minor of B will equal the minor M_{ik} of A. Expanding along the first row yields

$$\det B = a_{i1}M_{i1} - a_{i2}M_{i2} + \ldots \quad \ldots + (-1)^{n-1}a_{in}M_{in}$$

Equating this with det B = $(-1)^{i-1}$det A and using $C_{ij} = (-1)^{i+j}M_{ij}$ we see that $a_{i1}C_{i1} + a_{i2}C_{i2} + \ldots + a_{in}C_{in}$ = det A.

Now suppose i ≠ j. Let G be the matrix obtained from A by moving row i above row 1, leaving the other rows fixed, and then replacing row j by a second copy of row i. Then det G = 0 by (7) of (1.4.2) as G has two identical rows. Note that the (1,k)-minor of G will equal $(-1)^{i-j-1}M_{jk}$. (The second copy of row i in G must be moved through i-j-1 rows to return to its rightful place.) Expanding along the first row yields

$$\det G = (-1)^{i-j-1}(\ a_{i1}M_{j1} - a_{i2}M_{j2} + \dots + (-1)^{n-1}a_{in}M_{jn}\ ).$$

Using $\det G = 0$ and $C_{ij} = (-1)^{i+j}M_{ij}$ we see that

$$a_{i1}C_{j1} + a_{i2}C_{j2} + \dots\dots + a_{in}C_{jn} = 0.$$

This proves that A (adj A) = (det A) I.

Replacing A by A^t in this last equation, transposing the whole equation, and using (ii), (iii) of (1.1.11) and (2) of (1.4.2) yields that $(\text{adj } A^t)^t A = (\det A) I$. From the definition of adjugate and property (2) of (1.4.2) we see that adj $A^t = (\text{adj } A)^t$ for any matrix A.

Thus (adj A) A = (det A) I and the proof is complete.