MT2501: Linear Mathematics

Lecturer: Steve Buckland

Session: 2014-15 (semester 2)

Contents

Introduction

Linear algebra arises out of the study of matrices and vectors. In previous studies (e.g., in MT1002), you have probably have seen how matrices arise naturally in two settings, first in the context of solving systems of linear equations and second as specifying certain geometric transformations. Vectors also arise naturally in the geometric setting of 2-dimensional and 3-dimensional real space. In this course, we shall see how the concept of a vector space has been generalised from its geometric origins to an algebraic object which has applications in a range of different settings:

- pure mathematics (e.g. geometry, algebra, functional analysis, etc.),
- applied mathematics (e.g. spaces of solutions to differential equations, etc.),
- statistics (e.g. in specifying models and fitting them to data),
- physics (e.g. quantum mechanics, etc.),
- etc.

Linearity makes solving many problems much easier. The sort of information that we are able to determine concerning transformations that are linear is far beyond what we can hope to achieve for arbitrary functions. Further, vector spaces form the natural setting in which to study linearity.

In this course, we will meet the following methods:

- how to work with row and column operations;
- how to calculate the determinant and inverse of a matrix;
- how to recognise a vector space;
- how to show that a set of vectors in a vector space is linearly independent;
- how to show that a set of vectors in a vector space forms a basis for the space;
- how to recognise a linear transformation;
- how to obtain the matrix of a linear transformation;
- how to calculate the rank and nullity of a linear transformation;
- how to find eigenvalues and eigenvectors and hence diagonalise a linear transformation.

All these ideas are taken further in the module MT3501 Linear Mathematics.

Chapter 1

Matrices and systems of equations

Definition 1.1 A rectangular array A of mn entries set out in m rows and n columns is called an $m \times n$ matrix. It has the form

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

The entry a_{ij} in the ith row and jth column is called the (i, j)th entry of A. We sometimes write $A = [a_{ij}]$.

A *square* matrix is one with $m = n$, and its *main diagonal* is the set of entries $\{a_{11}, a_{22}, \ldots, a_{nn}\}$.

Example 1.2 *The $n \times n$ identity matrix I_n is a square matrix, with 1's on its main diagonal and 0's everywhere else.*

$$I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

(We will simply write I if the value of n is clear.)

Observe that, if we have a system of linear equations,

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ \vdots \qquad \vdots \qquad\qquad \vdots \quad & \quad \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned} \tag{1.1}$$

then solving the system of equations (1.1) is equivalent to solving the equation

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

for $\boldsymbol{x}$, where $\boldsymbol{x}$ is the column vector

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

and $\boldsymbol{b}$ is

$$\boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}.$$

This connection motivates much of the theory which we will develop in this course.

Where do the entries of a matrix A come from? In order for our theory to work, we require the entries come from a system, such as $\mathbb{R}$ or $\mathbb{C}$, which possesses certain nice properties. The necessary system is called a *field* F; we give a non-technical definition, sufficient for our purposes.

Definition 1.3 A *field* F is a structure in which we can add, multiply and subtract any two elements and we can divide any element by a *non-zero* element. Furthermore, all natural rules of arithmetic should hold.

We refer to the elements in our field as *scalars*.

For example, the natural numbers and the integers do not form fields, but the rational numbers, the real numbers and the complex numbers do (check!). In this course, you may assume that F is always $\mathbb{R}$ or $\mathbb{C}$.

It is assumed in this course that you have met the field of complex numbers $\mathbb{C}$ before, but a brief recap is offered here. A *complex number* is an expression of the form $z = a + bi$ where a and b are real numbers and i is thought of as the square root of -1. We refer to a as the real part of z, $Re(z)$, and b as the imaginary part of z, $Im(z)$. Each complex number has a so-called *complex conjugate*: the complex conjugate of $z = a+bi$ is defined to be $\bar{z} = a-bi$ ($Re(\bar{z}) = Re(z)$, $Im(\bar{z}) = -Im(z)$). The usual rules of algebra hold with complex numbers — always remembering that $i^2 = -1$. The sum and product of two complex numbers is again a complex number: precisely, $(a+bi)+(c+di) = (a+c)+(b+d)i$ and $(a+bi)(c+di) = (ac-bd)+(ad+bc)i$. An important property of a field is that any non-zero element has an inverse. To obtain the inverse of z, we take $\frac{1}{z} = \frac{1}{a+bi} = \frac{a-bi}{(a+bi)(a-bi)} = \frac{a-bi}{a^2+b^2} = \frac{a}{a^2+b^2} + \frac{(-b)}{a^2+b^2}i$. The quantity $z\bar{z}$ equals $a^2 + b^2 = |z|^2$, where $|z|$ is called the modulus or absolute value of z. We can represent $a + bi$ as the point (a, b) in the coordinate plane, with distance $|z|$ from the origin.

We define the *conjugate* of a complex matrix $A = [a_{ij}]$ by the matrix $\bar{A} = [\bar{a_{ij}}]$. We will return specifically to matrices with entries from $\mathbb{C}$, in the final chapter.

We denote the matrix, all of whose entries are zero, by $\mathbf{0}$. We call this matrix the *zero matrix*.

Definition 1.4 The *transpose* of an $m \times n$ matrix $A = [a_{ij}]$, denoted by A^T, is the $n \times m$ matrix whose (i, j)th entry is a_{ji}.

Matrix A^T is obtained from A by interchanging rows and columns.

Note that we have the following *reversal rule* for transposes:

Proposition 1.5 *For an $m \times n$ matrix A and an $n \times m$ matrix B,*

$$(AB)^T = B^T A^T.$$

PROOF: Recall that for $A = [a_{ij}]$ and $B = [b_{ij}]$, the formula for their product $AB = C = [c_{ij}]$ is

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}.$$

So the (i, j)th entry of $(AB)^T$ is

$$\sum_{k=1}^{n} a_{jk} b_{ki}.$$

We have $A^T = [a_{ji}]$ and $B^T = [b_{ji}]$. Then the (i, j)th entry of $B^T A^T$ is given by

$$\sum_{k=1}^{n} b_{ki} a_{jk} = \sum_{k=1}^{n} a_{jk} b_{ki}$$

as required. $\square$

We will meet various special families of matrices in this course.

Definition 1.6 An $n \times n$ matrix A is

- *symmetric* if $A^T = A$.
- *skew-symmetric* if $A^T = -A$.
- *idempotent* if $A^2 = A$.

Definition 1.7 A square $n \times n$ matrix A is called a *diagonal* matrix if all entries off its main diagonal are zero. Let $A = [a_{ij}]$; then A is diagonal if $a_{ij} = 0$ whenever $i \neq j$.

If $A = [a_{ij}]$ is diagonal, we often write it as $diag\{a_{11}, a_{22}, \ldots, a_{nn}\}$. The identity is an example of a diagonal matrix (we could write it as $diag\{1, 1, \ldots, 1\}$). Note that the all-zero matrix is also a diagonal matrix (we do not insist that the diagonal entries are non-zero).

Elementary row operations

It is often very helpful to put matrices into certain standard forms:

Definition 1.8 A matrix E is said to be in *(row) echelon form* if it has the following two properties:

(i) the zero rows, if any, occur at the bottom;

(ii) each leading (non-zero) entry in a row is in a column to the right of the leading entry of the row above it.

An echelon matrix E is said to be in *reduced echelon form* if, in addition,

(i) in each non-zero row, the leading entry is 1, and

(ii) in each column that contains the leading entry of a row, all other entries are zero.

Thus a matrix in echelon form has the following general shape:

$$E = \begin{pmatrix} * & * & * & \cdots & & & \\ 0 & \cdots & * & * & \cdots & & \\ 0 & \cdots & \cdots & * & * & \cdots \\ \vdots & & & \ddots & * & * \\ 0 & \cdots & \cdots & & & * \end{pmatrix}$$

while a matrix in reduced echelon form has the following shape:

$$E = \begin{pmatrix} 1 & * & 0 & 0 & * & 0 & * & * \\ 0 & 0 & 1 & 0 & * & 0 & * & * \\ 0 & \cdots & 0 & 1 & * & 0 & * & * \\ \vdots & & \ddots & 0 & 0 & 1 & * & * \\ 0 & \cdots & & & 0 & 0 & * & 0 \end{pmatrix}$$

A specific example of an echelon matrix is given by

$$E = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 6 & 7 & 8 & 9 \\ 0 & 0 & 0 & 10 & 11 \\ 0 & 0 & 0 & 0 & 12 \end{pmatrix}$$

while a specific example of a reduced echelon matrix (not related to the above) is

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

Given a matrix, there are three types of operation, called *elementary row operations (e.r.o.'s)*, which we may perform on its rows. These are as follows (where r_i denotes the ith row of the matrix):

- interchanging of two rows ($r_i \leftrightarrow r_j$, $i \neq j$);
- multiplying a row by a non-zero scalar λ ($r_i \to \lambda r_i$, $\lambda \neq 0$);
- adding a non-zero scalar multiple of one row to another ($r_i \to r_i + \lambda r_j$, $i \neq j$).

For matrices A and B, we say that A is *row-equivalent* to B precisely if there is a sequence of elementary row operations that transforms A into B.

Proposition 1.9 *(i) Every non-zero matrix A is row-equivalent to an echelon matrix.*

(ii) Every non-zero matrix A is row-equivalent to a reduced echelon matrix.

PROOF: (i) Systematic but lengthy use of elementary row operations.
(ii) Let E be an echelon matrix obtained from A by the process of part (i). Divide each non-zero row of E by the leading entry in that row, to make each leading entry 1. Now subtract suitable multiples of every such non-zero row from every row above it, to obtain a reduced echelon matrix. $\square$

Example 1.10 *We give an example of a matrix which is converted first to echelon and then to reduced echelon form via ero's:*

$$\begin{aligned}
A = &\begin{pmatrix} 0 & 2 & 4 & -1 \\ 1 & 1 & 1 & 3 \\ 1 & 0 & -1 & 6 \end{pmatrix} \\
\to &\begin{pmatrix} 1 & 1 & 1 & 3 \\ 0 & 2 & 4 & -1 \\ 1 & 0 & -1 & 6 \end{pmatrix} && r_1 \leftrightarrow r_2 \textit{ to make leading entry of row 1 non-zero} \\
\to &\begin{pmatrix} 1 & 1 & 1 & 3 \\ 0 & 2 & 4 & -1 \\ 0 & -1 & -2 & 3 \end{pmatrix} && r_3 \mapsto r_3 - r_1 \textit{ to clear the first column} \\
\to &\begin{pmatrix} 1 & 1 & 1 & 3 \\ 0 & 2 & 4 & -1 \\ 0 & 0 & 0 & \frac{5}{2} \end{pmatrix} && r_3 \mapsto r_3 + \tfrac{1}{2} r_2 \textit{ to clear the second column}
\end{aligned}$$

This is now in echelon form. To reach reduced echelon form, we must persevere further:

$$\begin{aligned}
&\begin{pmatrix} 1 & 1 & 1 & 3 \\ 0 & 2 & 4 & -1 \\ 0 & 0 & 0 & \frac{5}{2} \end{pmatrix} \\
\to &\begin{pmatrix} 1 & 0 & -1 & \frac{7}{2} \\ 0 & 1 & 2 & -\frac{1}{2} \\ 0 & 0 & 0 & 1 \end{pmatrix} && \begin{array}{l} r_1 \mapsto r_1 - \tfrac{1}{2} r_2 \textit{ to get 0 at the top of the second column} \\ r_2 \mapsto \tfrac{1}{2} r_2 \textit{ to obtain the leading entry 1 in row 2} \\ r_3 \mapsto \tfrac{2}{5} r_3 \textit{ to obtain the leading entry 1 in row 3} \end{array} \\
\to &\begin{pmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} && \begin{array}{l} r_1 \mapsto r_1 - \tfrac{7}{2} r_3 \textit{ to get 0 at the top of the fourth column} \\ r_2 \mapsto r_2 + \tfrac{1}{2} r_3 \textit{ to get 0 in row 2, column 4} \end{array}
\end{aligned}$$

For a given matrix A, the resulting echelon matrix is not uniquely determined by A — it depends on the sequence of e.r.o.'s chosen. However, all echelon matrices row-equivalent to a given matrix A do have the same number of non-zero rows (see later). It can be shown (we do not prove it here) that:

Theorem 1.11 *(i) For a given matrix A, there is only one reduced echelon matrix E that is row-equivalent to A;*

(ii) In the case when A is a square matrix, then either E has an all-zero row or $E = I_n$.

Observe the strong link between these elementary row operations and the familiar operations used on systems of linear equations during the process of Gaussian elimination (multiplying an equation by a non-zero factor, interchanging two equations, adding some multiple of one equation to any other equation).

For example, suppose we have the system

$$\begin{aligned} x + y - z - 3t &= 7 \\ x + 3y + 2z + 4t &= -2 \\ 2x + y + z + 4t &= 5. \end{aligned}$$

We can represent this system using the augmented matrix:

$$\begin{pmatrix} 1 & 1 & -1 & 3 & | & 7 \\ 1 & 3 & 2 & 4 & | & -2 \\ 2 & 1 & 1 & 4 & | & 5 \end{pmatrix}$$

Applying Gaussian elimination, we obtain

$$\begin{array}{ccccccccc} x & & & & + & t & = & 4 \\ & & y & & - & t & = & 0 \\ & & & z & + & 3t & = & -3 \end{array}$$

This is precisely equivalent to reducing the augmented matrix so that the left-hand matrix is in reduced echelon form:

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

Hence we obtain the general solution, with one parameter α:

$$x = 4 - \alpha, y = \alpha, z = -3 - 3\alpha, t = \alpha \quad (\alpha \in F)$$

Observe that our reduced system, comprising 3 equations in 4 unknowns, has a general solution involving $1 = 4 - 3$ parameters. This is an example of a general result (which we shall state, but not prove, here). Recall that a system is called *consistent* if it has one or more solutions (and inconsistent otherwise).

Theorem 1.12 *Suppose that a consistent system of equations, in reduced echelon form, consists of r equations (where redundant "0=0" equations have been deleted) in n unknowns. Then the general solution (obtained by Gaussian elimination) involves $n - r$ parameters.*

Recall that a system is called *homogeneous* if all the right-hand sides are zero (it is certainly consistent, as it has the trivial solution, at least).

Corollary 1.13 *A homogeneous system of linear equations, with fewer equations than unknowns, has a non-trivial solution (a solution in which at least one of the unknowns is non-zero).*

Determinants and inverses

Next, we turn our attention specifically to square matrices (which correspond to systems of n equations in n unknowns).

Definition 1.14 Let A be a square $n \times n$ matrix. We say that A is *invertible* with inverse A^{-1} if

$$AA^{-1} = A^{-1}A = I_n$$

where I_n is the $n \times n$ identity matrix. Note that A^{-1} is necessarily also $n \times n$.

Definition 1.15 A square matrix P whose inverse is equal to its transpose is called *orthogonal*.

Lemma 1.16 *If A and B are invertible $n \times n$ matrices, then AB is also invertible, with*

$$(AB)^{-1} = B^{-1}A^{-1}.$$

PROOF: Calculate

$$AB \cdot B^{-1}A^{-1} = A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I$$

and similarly $B^{-1}A^{-1} \cdot AB = I$. Hence AB has $B^{-1}A^{-1}$ as its inverse. $\square$

It is not the case that all matrices have inverses, but if a matrix is invertible then its inverse is unique. The following result (seen in MT1002) is useful for determining whether a matrix has an inverse, and finding it if so:

Proposition 1.17 *Let A be an $n \times n$ matrix. If a sequence of row operations, when performed on A, reduce it to the identity I_n, then exactly the same row operations applied in the same order to I_n will change I_n into A^{-1}.*

Theorem 1.18 *Let A be an $n \times n$ matrix. Then the following statements are equivalent:*

(i) A is invertible;

(ii) For each $n \times 1$ (column) vector $\boldsymbol{b}$, the system of equations $A\boldsymbol{x} = \boldsymbol{b}$ has a unique solution for $\boldsymbol{x}$;

(iii) A is row-equivalent to I_n.

PROOF: (i) implies (ii): If A is invertible, then given $A\boldsymbol{x} = \boldsymbol{b}$ we obtain:

$$A^{-1}\boldsymbol{b} = A^{-1}(A\boldsymbol{x}) = (A^{-1}A)\boldsymbol{x} = I_n\boldsymbol{x} = \boldsymbol{x}.$$

(ii) implies (iii): From (ii), by specializing to $\boldsymbol{b} = \boldsymbol{0}$, we see that the system of equations $A\boldsymbol{x} = \boldsymbol{0}$ has a unique solution for $\boldsymbol{x}$, which must be the trivial solution $\boldsymbol{x} = \boldsymbol{0}$. Expressing this system in full, we have:

$$\begin{aligned}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= 0 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= 0 \\
\vdots \qquad \vdots \qquad\qquad \vdots \quad &\;\; \vdots \\
a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n &= 0
\end{aligned} \tag{1.2}$$

By Corollary 1.13, this system (in reduced echelon form) cannot have fewer equations than unknowns (otherwise it would have a non-trivial solution). So the reduced echelon matrix to which the $n \times n$ coefficient matrix A is converted cannot have any zero rows, and so must be I_n.
(iii) implies (i): this follows immediately from Proposition 1.17. $\square$

We mention that it is of course also possible to define *elementary column operations*: simply replace "row" by "column" in the definition. These cannot be used in the same way as row operations to solve systems of equations — they do not produce an equivalent system. However we will revisit them later. Note that column operations are essentially row operations on the transpose of the matrix. We also have the notion of two matrices being *column-equivalent*.

Next, we consider a number which arises naturally when solving systems of n equations in n unknowns — the *determinant*. Let $A = [a_{ij}]$ be an $n \times n$ matrix with entries from a field F. The determinant of A was defined, inductively, in MT1002. We recall this.

First associate to each position in the $n \times n$ matrix a sign from the following grid:

$$\begin{pmatrix} + & - & + & \cdots \\ - & + & - & \cdots \\ + & - & + & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

So, for the (i, j)th entry, the sign is $(-1)^{i+j}$.

Definition 1.19 Fix a particular entry in the $n \times n$ square matrix A. The *cofactor* of this entry is defined to be:

> (the determinant of the $(n-1) \times (n-1)$ matrix obtained by deleting the row and column containing this entry) $\times$ (the sign associated with this entry).

The *determinant* of the $n \times n$ matrix $A = [a_{ij}]$ is then obtained by multiplying each entry in the top row by its cofactor and adding the resulting products. We denote this determinant by $\det A$ or $|A|$.

In order to calculate a determinant of a 3×3 matrix, we first need to calculate that of a 2×2 matrix, and so on.

- For a 1×1 matrix, $A = (a_{11})$, the determinant is given by

$$\det A = \det(a_{11}) = a_{11}.$$

- For a 2×2 matrix,

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

  the determinant is

$$\begin{aligned} \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} &= a_{11} \det(a_{22}) - a_{12} \det(a_{21}) \\ &= a_{11}a_{22} - a_{12}a_{21}. \end{aligned}$$

- For a 3×3 matrix,

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

  the determinant is

$$\begin{aligned} \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} &= a_{11} \det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{12} \det \begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} \\ &\qquad + a_{13} \begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} \\ &= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) \\ &\qquad + a_{13}(a_{21}a_{32} - a_{22}a_{31}) \\ &= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} \\ &\quad + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}. \end{aligned}$$

These expressions possess a lot of symmetry. For example, if we take the expression for $\det A$ in the 2×2 or 3×3 case, and we swap a_{ij} for a_{ji} throughout, then the expression remains unchanged overall. Since this swapping is equivalent to taking the transpose of A, it tells us that here, the determinant of A^T is the same as that of A. In fact, this is true in general:

Proposition 1.20 *Let A be an $n \times n$ matrix. Then* $\det A = \det A^T$.

A key consequence is that any property of determinants that is expressed in terms of rows also holds when expressed in terms of columns.

We may check that, in the 2×2 and 3×3 case, the same expression for $\det A$ can be obtained by taking the sum of "entry times cofactor" for *any* row of A, not just the first. This is also true in general — and by Proposition 1.20, any column, too:

Proposition 1.21

- *For each $i = 1, \ldots, n$,* $\det A$ *is the sum of all (entry)× (cofactor) products for the ith row. This is called* expanding the determinant along the ith row.
- *For each $j = 1, \ldots, n$,* $\det A$ *is the sum of all (entry)× (cofactor) products for the jth column. This is called* expanding the determinant down the jth column.

Example 1.22 *Calculate the determinant of*

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

Solution: An expansion in terms of the top row gives:

$$\det A = 1 \times \begin{vmatrix} -1 & 1 \\ 1 & -1 \end{vmatrix} - 2 \times \begin{vmatrix} 1 & 1 \\ 2 & -1 \end{vmatrix} + 1 \times \begin{vmatrix} 1 & -1 \\ 2 & 1 \end{vmatrix}$$

$$
\begin{aligned}
&= 0 - 2 \times (-3) + 3 \\
&= 0 + 6 + 3 = 9.
\end{aligned}
$$

An equally valid solution would be to expand in terms of the first column:

$$
\begin{aligned}
\det A &= 1 \times \begin{vmatrix} -1 & 1 \\ 1 & -1 \end{vmatrix} - 1 \times \begin{vmatrix} 2 & 1 \\ 1 & -1 \end{vmatrix} + 2 \times \begin{vmatrix} 2 & 1 \\ -1 & 1 \end{vmatrix} \\
&= 1.(1-1) - 1.(-2-1) + 2.(2+1) \\
&= 0 + 3 + 6 = 9.
\end{aligned}
$$

We now connect with our previous work, by asking: what is the effect on the determinant of a matrix A of applying each of the elementary row operations to A?

Theorem 1.23 *Let A be an $n \times n$ matrix with entries from F.*

- *Let B be the matrix obtained from A by interchanging two rows. Then $\det B = -\det A$;*
- *Let B be the matrix obtained from A by multiplying a row of A by a non-zero scalar λ. Then $\det B = \lambda \det A$;*
- *Let B be the matrix obtained from A by adding to one row a non-zero scalar multiple of another row. Then $\det B = \det A$.*

Example 1.24 *Calculate the determinant of*

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

Solution: This is an alternative solution to Example 1.22, but this time we shall apply row operations:

$$
\begin{aligned}
\det A &= \begin{vmatrix} 1 & 2 & 1 \\ 1 & -1 & 1 \\ 2 & 1 & -1 \end{vmatrix} \\
&= \begin{vmatrix} 1 & 2 & 1 \\ 0 & -3 & 0 \\ 0 & -3 & -3 \end{vmatrix} && \begin{matrix} r_2 \mapsto r_2 - r_1 \\ r_3 \mapsto r_3 - 2r_1 \end{matrix} \\
&= \begin{vmatrix} -3 & 0 \\ -3 & -3 \end{vmatrix} && \text{(expanding along first column)} \\
&= 9.
\end{aligned}
$$

We have a useful corollary:

Corollary 1.25

- *If the matrix A has a zero row, its determinant is zero.*
- *If the matrix A has two identical rows, its determinant is zero.*

PROOF: The first part is immediate on expanding along the zero row. For the second part, suppose rows r_i and r_j are identical ($i \neq j$). For matrix A, swap $r_i \leftrightarrow r_j$ to get matrix B; then $\det B = -\det A$. But $A = B$, so $\det A = -\det A = 0$. $\square$

We can now prove a key theorem about determinants and inverses:

Theorem 1.26 *For an $n \times n$ matrix A, $\det A = 0 \Leftrightarrow A$ is non-invertible.*

PROOF: The matrix A can be converted, by a sequence of elementary row operations, to a reduced echelon matrix E. From Theorem 1.23, these row operations do not affect whether the determinant is zero. By Theorem 1.18, if A is invertible, then $E = I_n$, and since $\det I_n = 1 \neq 0$, $\det A \neq 0$. If A is non-invertible, then E has at least one zero row; expanding along this, we see $\det E = 0$, and so $\det A = 0$. □

From this, we may obtain (but do not prove here) the multiplicative property of determinants:

Theorem 1.27 *If A and B are $n \times n$ matrices, then $\det(AB) = \det A \cdot \det B$.*

We may now restate a useful earlier result in terms of determinants:

Theorem 1.28 *The system of equations $A\boldsymbol{x} = \boldsymbol{b}$ has a unique solution if and only if $\det A \neq 0$.*

Finding inverses

We end this chapter by describing a direct method for finding inverses. The first ingredient is the following:

Definition 1.29 Let A be an $n \times n$ matrix. The *adjugate* (or *adjoint*) of A, denoted by $\operatorname{adj} A$, is constructed from A by the following two steps:

(i) Replace each entry in the matrix by its cofactor;

(ii) take the transpose of the resulting matrix.

Example 1.30 *Find the adjugate of the matrix*

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

and calculate the product $A \cdot \operatorname{adj} A$.

Solution: Recall that the distribution of the signs associated with the cofactors is:

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

Each cofactor is the determinant of a 2×2 matrix adjusted by the above signs, so we calculate the matrix of cofactors as

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

Hence

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

Thus

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

We can easily check that: $\det A = 1.0 + 2.1 + 0(-3) = 2$. Hence $2I = (\det A)I$.

We will now show that what was observed in this example is in fact true generally.

Theorem 1.31 *If A is a square matrix, then*

$$A \cdot \operatorname{adj} A = (\det A)I.$$

PROOF: Let $A = [a_{ij}]$ be an $n \times n$ matrix. Write

$$B = [b_{ij}] = \operatorname{adj} A$$

for the adjugate of A. Here b_{ij} is the cofactor of the (j,i)th entry of A.

Now consider the (i,i)th entry of the product AB. It equals

$$\sum_{k=1}^{n} a_{ik} b_{ki}.$$

This is the sum of each entry in the ith row multiplied by its cofactor (b_{ki} is the cofactor of the (i,k)th entry), so this sum is simply the determinant of A calculated by expanding along the ith row. Hence the diagonal entries in AB all equal $\det A$.

Now consider the $(1,2)$ entry of the product AB. It equals

$$\sum_{k=1}^{n} a_{1k} b_{k2}.$$

This is much like expanding the determinant of A using the cofactors of the second row, but we are multiplying them by the entries of the first row instead of the second. Hence this sum equals the determinant

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{11} & a_{12} & \cdots & a_{1n} \\ a_{31} & a_{32} & \cdots & a_{3n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

when we expand along the second row. This matrix has repeated rows, so its determinant is zero. Hence the $(1,2)$ entry of AB equals 0. The same holds for all other off diagonal entries by the same argument. Hence

$$A \cdot \operatorname{adj} A = AB = (\det A)I.$$

□

This enables us to calculate the inverse of an invertible matrix directly from its adjugate:

Theorem 1.32 *If A is invertible, then*

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

PROOF: If A is invertible, then $\det A \neq 0$. Hence $B = (1/\det A) \operatorname{adj} A$ exists and

$$AB = \frac{1}{\det A} A \cdot \operatorname{adj} A = \frac{1}{\det A} \cdot (\det A) I = I.$$

Similarly, it can be checked that $BA = I_n$, and so $B = A^{-1}$. □

Example 1.33 *Find the inverse of*

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

Solution: The determinant of A is

$$\begin{aligned} \det A &= \begin{vmatrix} 1 & 2 & 1 \\ 3 & 0 & 2 \\ 1 & 3 & 2 \end{vmatrix} \\ &= \begin{vmatrix} 1 & 2 & 1 \\ 0 & -6 & -1 \\ 0 & 1 & 1 \end{vmatrix} = -6 + 1 = -5. \end{aligned}$$

Taking cofactors and the transpose, we find that the adjugate of A is

$$\operatorname{adj} A = \begin{pmatrix} -6 & -1 & 4 \\ -4 & 1 & 1 \\ 9 & -1 & -6 \end{pmatrix}.$$

Hence

$$A^{-1} = \begin{pmatrix} \frac{6}{5} & \frac{1}{5} & -\frac{4}{5} \\ \frac{4}{5} & -\frac{1}{5} & -\frac{1}{5} \\ -\frac{9}{5} & \frac{1}{5} & \frac{6}{5} \end{pmatrix}.$$

Recall from earlier that the system of linear equations $A\boldsymbol{x} = \boldsymbol{b}$ has a unique solution if and only if A is invertible, i.e. if and only if $\det A \neq 0$. If A is non-invertible, the system has either no solutions or more than one solution. If A is invertible, the system has a unique solution given by $\boldsymbol{x} = A^{-1}\boldsymbol{b}$. We finish this chapter with a specific example of solving a system using matrix inverses.

Example 1.34 *Consider the system of equations*

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

Determine whether it has a unique solution, and find it if so.

Solution: In matrix form, this system is equivalent to

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

or

$$A\boldsymbol{x} = \boldsymbol{b},$$

where

$$A = \begin{pmatrix} 1 & 3 & 3 \\ 1 & 4 & 3 \\ 1 & 3 & 4 \end{pmatrix} \qquad \text{and} \qquad \boldsymbol{b} = \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}.$$

We check whether matrix A is invertible: $\det A = 1 \neq 0$, so A is invertible. Proceeding via the adjugate, we calculate that

$$A^{-1} = \begin{pmatrix} 7 & -3 & -3 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \qquad \text{and so} \qquad \boldsymbol{x} = A^{-1}\boldsymbol{b} = \begin{pmatrix} -7 \\ 1 \\ 2 \end{pmatrix}.$$

Hence $\{x = -7, y = 1, z = 2\}$ is the unique solution to this system of equations.

Alternatively, we could proceed via Gaussian elimination as follows: the augmented matrix is:

$$\left(\begin{array}{ccc|c} 1 & 3 & 3 & 2 \\ 1 & 4 & 3 & 3 \\ 1 & 3 & 4 & 4 \end{array}\right)$$

In echelon form this is:

$$\left(\begin{array}{ccc|c} 1 & 3 & 3 & 2 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 2 \end{array}\right)$$

In reduced echelon form, the solution can now be read off directly (notice that A has been converted to the identity, as expected):

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